

Cairo University
Institute of Statistical Studies and Research

**The 52nd Annual Conference on Statistics, Computer
Sciences and Operations Research**

Computer Science

25-27 Dec. 2017

Index

Computer Sciences

1	An Improvement and Implementation of the DPLL Satisfiability Algorithm Munira A. El-Maksoud , Areeg Abdalla	1-12
2	A proposed Architecture for Shopping Association Mining Essam Eldin Mosallam, Reda Abd Elwahab, Alkhoribi, Mohamed Ragaie Sayed Osman	13-23
3	Mining IoT Data Streams Zaki. A. M.	24-41
4	Arabic Named Entity Recognition Mariam Muhammad	42-62
5	Image Encryption Scheme with Hashed Biometric Key Ali M. Meligy ¹ , Hossam A. Diab ² and Marwa S. ElDanaf ³	63-74
6	Cauchy Based fuzzy neural network with Mutual Subsethood Product Inference Nelly S. Amer ¹ , Hesham A. Hefny ¹	75-85
7	Survey on Land Change Modeling Khalid A. Eldrandaly ^a , Mamdouh M. Abdeen ^b , Safa A. Abdelkareem ^c	86-105

¹ Dept. of computer sciences Institute of statistical studies and researches ,Cairo University.

² Dept. of computer sciences Institute of statistical studies and researches ,Cairo University.

An Improvement and Implementation of the DPLL Satisfiability Algorithm

Munira A. El-Maksoud , Areeg Abdalla *

Abstract

This paper introduces a modification of the well known Satisfiability solver, MINISAT. A new measure of the activity of the variable, determined by its occurrences in the not yet satisfied clauses, is introduced. Hence, variable activities are changed dynamically (increased and decreased) by adding and removing clauses. It is applied on an interesting mathematical problem, in finding the van der Waerden numbers, which are known to be very difficult to compute.

keywords: Satisfiability, DPLL, MINISAT, Branching Rules, van der Waerden numbers.

1 Introduction

Boolean Satisfiability Problem (SAT) is the problem of deciding whether a given boolean formula is satisfiable or not. Since SAT is very important in many applications, there has been many algorithms for testing the satisfiability. The most well-known one is introduced in 1962 by “ M. Davis, H. Putnam, G. Logemann and D. Loveland ”(DPLL [6]). It is considered the basis for almost all modern SAT solvers. It is known that SAT can be used (utilized) for solving problems of various applications after encoding them into SAT. For mention one example, some of mathematical problems can be encoded as SAT then solved using SAT solvers. A SAT solver(based on DPLL [6]) is a software, and many of the SAT solvers are written in C or C++ [25]. One may design his (or her) own SAT solver for a specific application, and in this case he (or she) has to choose between two ways, either to build their solver from scratch which may be hard work and time consuming

*Department of Mathematics, Faculty of Science, Cairo University.

or to modify an existing solver to meet the requirements (goals). For example, one can create new ideas in the branching rules resulting in various versions of already existing SAT solvers.

In this research we choose to modify the MINISAT solver in terms of calculation method of the activity of a variable and apply this modification in finding the van der Waerden numbers as a mathematical application. The rest of this paper is organized as follows: section2 concerns the Satisfiability Problem (SAT), the relation between SAT and combinatorics and introduces the problem of van der Waerden numbers as an example of this relation. Section3 presents a comparison between the results of both of the solvers VANSAT, MINISAT on a number of examples. Section4 summarizes the work we have done and what we intend to in the future.

2 SAT and Combinatorics

Areas of satisfiability and combinatorics may help advance each other. On one hand, thanks to the significant efficiencies of modern SAT solvers it became possible to encode many of combinatorics problems as boolean formulas and then solve their corresponding satisfiability problems. In this scenario, novel results in combinatorics are obtained. On the other hand, combinatorics problems can be utilized as a rich source of structured formulas for developing new generations of SAT solvers [7].

2.1 The Satisfiability Problem (SAT)

SAT is a central problem in theoretical computer science studied by computer scientists and mathematicians, which can be formulated as [1]:

A *truth assignment* is a mapping f that assigns 0 (interpreted as False) or 1 (interpreted as True) to each variable in its domain, let us enumerate all the variables in the domain as x_1, x_2, \dots, x_n . The complement \bar{x}_i of each variable x_i is defined by: $f(\bar{x}_i) = 1 - f(x_i)$ for all truth assignments f .

Both x_i and \bar{x}_i are called *literals*, if $u = \bar{x}_i$ then $\bar{u} = x_i$. A *clause* is a set of (distinct) literals, and a *formula* is a family of (not necessarily distinct) clauses. For example, $\{x_1, \bar{x}_2, x_3\}$ is a clause with three distinct literals and $\{\{x_1, x_2\}, \{x_1, \bar{x}_2\}, \{\bar{x}_1, x_2\}, \{\bar{x}_1, \bar{x}_2\}\}$ is a formula with four clauses over two variables.

A truth assignment *satisfies* a clause if it maps at least one of its literals to 1, the assignment *satisfies* a formula if and only if it satisfies each of its clauses. A formula is called *satisfiable* if it is satisfied by at least one truth assignment, otherwise it is called *unsatisfiable*. The problem of recognizing

satisfiable formulas is known as the *satisfiability problem*, or SAT for short. A SAT *instance* is any formula for which we want to solve the SAT problem.

2.2 The DPLL Algorithm

DPLL [6] was introduced in 1962 as a refinement of its earlier M. Davis, H. Putnam (DP) algorithm. Essentially, it is a (complete - “depth-first” - backtracking) search algorithm. More recently, João P. Marques-Silva and Karem A. Sakallah introduced Generic seaRch Algorithm for the Satisfiability Problem (GRASP) [19] as an extension of DPLL [6] with learning and non-chronological backtracking. In recent decades, GRASP prompts research on conflict-driven clause learning (CDCL) solvers. Most state-of-the-art SAT solvers as Glucose, CryptoMiniSat and MINISAT follow CDCL.

2.3 Branching Rules

The rule by which a solver selects an unassigned (free) variable and assigns it a value (branching on) is called a branching rule. It is known that choosing a suitable branching rule is an important since it affects the performance. Selecting different branching rules for the same algorithm may produce search trees with quite different sizes. Over the years, various rules have been proposed by researchers who try to find new ideas for improving the performance of SAT solvers. Some examples are Jeroslow-Wang [13], BOHM’s [5] and MOMS [10].

Here, we can talk about two kinds of branching rules: State-dependent and Variable State Independent Decaying Sum (VSIDS) rules.

State-dependent rules Where the occurrences of literals differ by different assignments of the variables. Hence to maintain the occurrences, the solver should pay a high price where occurrences should be recalculated each time a solver assigns (or un-assigns) a variable. Some examples are Maximum Occurrence in clauses of Minimum Size (MOMS) [10] and Dynamic Largest Individual Sum (DLIS).

Variable State Independent Decaying Sum (VSIDS) Because of the overhead in the solver for maintaining the literals occurrences in State-dependent strategy, new more efficient branching rules are needed. Chaff [20] proposed VSIDS which imposes an order on literals, where each literal has a score related to its occurrences. The activity of the

variables is captured by the literals scores which do not depend on assignments. For branching, VSIDS will select the (unassigned) variable corresponding to the literal possessing the highest score. Periodically the scores of all literals are multiplied by a factor less than one, hence, decaying .

“ More recently, BerkMin [11] proposed a new strategy “ pushes the idea of VSIDS further ”, where the activity is captured by conflicts. More clearly, after each conflict, a learned clause is generated through a process called “ resolution process ” and as a result, the scores of all literals in the clauses involved in this resolution have to be increased. BerkMin also decayed the scores periodically.

2.4 Sat Solvers

The possibility of encoding many of practical problems in diverse fields (like software verification [12], circuit testing [23], AI planning [17]) as SAT instances was the motivation for the research in developing new SAT solvers with significant efficiencies. Some examples are multi-SAT [21], Glucose and Syrup in the SAT’17 [4], Nigma [14](and its improved versions Nigma 1.1 [15], Nigma 1.2 [16]) and Glulu [24]. The following is a comparison between the two solvers MINISAT, VANSAT(VAN der Waerden numbers SAT).

MINISAT, MINImal SAT-solver, is a CDCL (CHAFF [20] - based) SAT solver written by Eén and Sörensson. MINISAT makes no distinction between the two phases of a variable. It attaches each variable with an activity which is increased whenever the variable appears in a conflict clause(bumping). MINISAT bumps variables with larger and larger numbers until reaching a limit (predefined number), at that point all variables activities are scaled down. MINISAT uses a heap to sort the variables by the activity at all times and for branching, it selects the variable with the highest activity [9].

On another side, VANSAT is a modification of MINISAT1.14 where the activity of a variable is captured by its occurrences in the not satisfied clauses. Hence, activities are changed dynamically(increased and decreased) by adding and removing clauses.

More accurately, the strategy of VANSAT is:

I- increasing the activity of all variables that appear in each new added clause (learnt or problem clause).

II- decreasing the activity of all variables that were appearing in each deleted clause (where deletion of clauses occur in many situations).

2.5 SAT and van der Waerden numbers

The problem of computing van der Waerden numbers is one of the most interesting examples in combinatorics which can be encoded as formulas. These problems can be represented by parameterized formulas in such a way that decisions concerning the satisfiability of these formulas determine the van der Waerden numbers in question. The next two sub-subsections define the van der Waerden numbers and discuss the SAT Encoding of them.

2.5.1 van der Waerden numbers

The van der Waerden number $w(r;t_1,t_2,\dots,t_r)$ is the least integer m such that for every partition $C_1 \cup C_2 \cup \dots \cup C_r$ of the set $\{1, 2, \dots, m\}$, there is an index j in $\{1, 2, \dots, r\}$ such that C_j contains an arithmetic progression(AP) of t_j terms [2].

where, $r > 0$ is the number of the blocks of the partition, t_j 's are the lengths of the AP's, C_j 's are the blocks of the partition and arithmetic progression(AP) is a sequence of numbers such that the difference between the consecutive terms, d , is constant. For example, the sequence 5, 7, 9, 11, 13, 15, ... is an arithmetic progression with common difference(d) of 2.

Finding the value of van der Waerden numbers presents a challenging problem, since the underlying principle behind their computation is still unknown. The interest of many researchers in van der Waerden numbers was the reason behind the computing of many of them, Table 1 lists some of the known van der Waerden numbers:

2.5.2 SAT Encoding of van der Waerden numbers

Given positive integers r,t_1,t_2,\dots,t_r and n , we construct a SAT formula, F , which is satisfiable if and only if $n < w(r;t_1,t_2,\dots,t_r)$.

To compute the van der Waerden numbers we use the following algorithm:
For consecutive integers $m = r+1, r+2, \dots$ we test whether the formula F is satisfiable. If so, we continue. If not, we return m and terminate the algorithm. (note: $t_1, t_2, \dots, t_r \geq 2$ and then $m > r$)

Consider the following two cases :

- (I) $r = 2$, we have n variables, x_1, x_2, \dots, x_n and two types of clauses:
 (1) $\{\bar{x}_a, \bar{x}_{a+d}, \dots, \bar{x}_{a+d(t_1-1)}\}$ with $a \geq 1, d \geq 1, a + d(t_1 - 1) \leq n$
 (2) $\{x_a, x_{a+d}, \dots, x_{a+d(t_2-1)}\}$ with $a \geq 1, d \geq 1, a + d(t_2 - 1) \leq n$

Table 1: SOME OF THE KNOWN VAN DER WAERDEN NUMBERS

$w(r; t_1, t_2, \dots, t_r)$		REFERENCE
$w(2; 3, 11)$	114	LANDMAN, ROBERTSON AND CULVER [18]
$w(2; 3, 13)$	160	LANDMAN, ROBERTSON AND CULVER [18]
$w(3; 2, 3, 7)$	55	LANDMAN, ROBERTSON AND CULVER [18]
$w(3; 3, 4, 4)$	89	LANDMAN, ROBERTSON AND CULVER [18]
$w(4; 2, 2, 3, 6)$	48	LANDMAN, ROBERTSON AND CULVER [18]
$w(4; 2, 2, 3, 7)$	65	LANDMAN, ROBERTSON AND CULVER [18]
$w(5; 2, 2, 2, 3, 3)$	20	LANDMAN, ROBERTSON AND CULVER [18]
$w(5; 2, 2, 3, 3, 3)$	41	LANDMAN, ROBERTSON AND CULVER [18]
$w(6; 2, 2, 2, 2, 4, 4)$	56	Tanbir Ahmed [1]
$w(6; 2, 2, 2, 3, 3, 3)$	42	Tanbir Ahmed [1]
$w(7; 2, 2, 2, 2, 2, 3, 3)$	24	Tanbir Ahmed [1]
$w(7; 2, 2, 2, 2, 2, 3, 4)$	36	Tanbir Ahmed [1]
$w(8; 2, 2, 2, 2, 2, 2, 3, 3)$	25	Tanbir Ahmed [1]
$w(9; 2, 2, 2, 2, 2, 2, 2, 3, 3)$	28	Tanbir Ahmed [1]

Where $x_i = \text{TRUE}$ encodes $i \in C_1$ and $x_i = \text{FALSE}$ encodes $i \in C_2$.

Clauses (1) prevent the existence of an arithmetic progression of length t_1 in C_1 and Clauses (2) prevent the existence of an arithmetic progression of length t_2 in C_2 .

(II) $r > 2$, we have nr variables, $x_{i,j}$'s , with $i=1,2,\dots,n$ and $j=1,2,\dots,r$ Where the variable $x_{i,j}$ takes the value TRUE if and only if the integer i belongs to a block C_j of a partition. Here , we have three types of clauses:

(1) $\{x_{i,1}, x_{i,2}, \dots, x_{i,r}\}$, for each integer i , to ensure that integer i belongs to at least one block of the partition.

(2) $\{\bar{x}_{a,j}, \bar{x}_{a+d,j}, \dots, \bar{x}_{a+d(t_j-1),j}\}$, for $1 \leq j \leq r$, $1 \leq a \leq n - t_j + 1$ and $1 \leq d \leq \lfloor (n - a)/(t_j - 1) \rfloor$, to ensure that no arithmetic progression of length t_j in block C_j .

(3) $\{\bar{x}_{i,s}, \bar{x}_{i,t}\}$, for $1 \leq i \leq n$, $1 \leq s < t \leq r$ to ensure that integer i belongs to at most one block of the partition.

3 Experimental work and Results

The experimental work was as follows, first, since most of the known branching rules take into account the occurrences of the variables in the clauses that have not been satisfied we computed the occurrences of the variables in the formulas . Codes are written to find the variables with the maximum number of occurrences based on the above SAT encoding for the van der Waerden numbers. The output of the programs indicated a *symmetry* for the number of occurrences of the variables in the clauses. Second, we wrote a program to generate the instances in DIMACS cnf format since the majority of SAT solvers, including MINISAT, accept the input in this format. Finally, to implement a new sat solver based on MINISAT, it is required first to install its source-code from the MINISAT web-page and then editing it.

The following tables(except the last) show a comparison between results of both of VANSAT and MINISAT Solvers for a number of examples of van der Waerden numbers in terms of the number of Restarts, Conflicts, Decisions, Propagations, Conflict literals, Memory used and Cpu Time where VANSAT works better. The last table shows a comparison between the two solvers for a Random-3-SAT instance.

Table 2 shows a comparison between VANSAT and MINISAT in finding the van der Waerden number $w(3;2,3,3)$ for the integer $n=13$. It is clear that the VANSAT gives much better results in the number of Conflicts, Decisions, Propagations and Conflict literals. It also used less Memory.

Table 2: van der Waerden number : $w(3;2,3,3)$ for $n=13$

EXAMPLE1: $w(3;2,3,3)$, $n=13$	MINISAT	VANSAT
Restarts	1	1
Conflicts	18	1
Decisions	25	6
Propagations	274	51
Conflict literals	132	7
Memory used	5.81 MB	4.81 MB
Cpu Time	0 s	0 s

Table 3 shows a comparison between VANSAT and MINISAT in finding the van der Waerden number $w(3;2,3,3)$ for the integer $n=14$, where VANSAT works better than MINISAT and gives less number of Conflicts, Decisions, and much better in Propagations, and Conflict literals. It also used less Memory.

Table 3: van der Waerden number : $w(3;2,3,3)$ for $n=14$

EXAMPLE2: $w(3;2,3,3)$, $n=14$	MINISAT	VANSAT
Restarts	1	1
Conflicts	67	61
Decisions	76	70
Propagations	1101	985
Conflict literals	387	229
Memory used	5.81 MB	4.88 MB
Cpu Time	0 s	0 s

Table 4 shows a comparison between VANSAT and MINISAT in finding the van der Waerden number $w(3;2,3,3)$ for the integer $n=8$, though the VANSAT works as good as the MINISAT, it used less Memory.

Table 4: van der Waerden number : $w(3;2,3,3)$ for $n=8$

EXAMPLE3: $w(3;2,3,3)$, $n=8$	MINISAT	VANSAT
Restarts	1	1
Conflicts	0	0
Decisions	7	7
Propagations	24	24
Conflict literals	0	0
Memory used	5.81 MB	4.81 MB
Cpu Time	0 s	0 s

Table 5 shows a comparison between VANSAT and MINISAT in finding the van der Waerden number $w(3;2,3,5)$ for the integer $n=20$, where VANSAT works better than MINISAT and gives less number of Conflicts, Propagations and Conflict literals. It also used less Memory.

Table 5: van der Waerden number : $w(3;2,3,5)$ for $n=20$

EXAMPLE4: $w(3;2,3,5)$, $n=20$	MINISAT	VANSAT
Restarts	1	1
Conflicts	4	0
Decisions	14	19
Propagations	141	60
Conflict literals	64	0
Memory used	5.81 MB	4.81 MB
Cpu Time	0 s	0 s

Table 6 shows a comparison between VANSAT and MINISAT in finding the van der Waerden number $w(4;2,2,3,3)$ for the integer $n=16$, where VANSAT works much better than MINISAT and gives less number of Restarts, Conflicts, Decisions, Propagations and Conflict literals. It also used less Memory and was better in CPU usage.

Table 6: van der Waerden number : $w(4;2,2,3,3)$ for $n=16$

EXAMPLE5: $w(4;2,2,3,3)$, $n=16$	MINISAT	VANSAT
Restarts	3	1
Conflicts	255	58
Decisions	367	263
Propagations	5149	1369
Conflict literals	2810	319
Memory used	5.81 MB	4.81 MB
Cpu Time	0.015 s	0 s

Table 7 shows a comparison between VANSAT and MINISAT on Random-3-SAT instance for clause length=3, 50vars and 218clauses, where VANSAT works better than MINISAT and gives less number of Conflicts, Decisions, Propagations and Conflict literals. It also used less Memory and was better in CPU usage.

Table 7: Random-3-SAT : clause length=3, 50vars, 218clauses

EXAMPLE6:Random-3-SAT	MINISAT	VANSAT
Restarts	1	1
Conflicts	44	35
Decisions	58	44
Propagations	594	501
Conflict literals	130	107
Memory used	5.81 MB	4.81 MB
Cpu Time	0.015 s	0 s

4 Conclusion and Future Work

This paper concerned the Satisfiability Problem (SAT) as a central problem in theoretical computer science and its well known DPLL algorithm. It introduced the solver VANSAT as a modification of the MINISAT solver(which based on DPLL). In other words, the paper introduced an implementation of DPLL different from that of the MINISAT where a new measurement of the activity of a variable is considered. Experimental results proved that VANSAT worked better in finding some of van der Waerden numbers in terms of the number of Restarts, the number of Conflicts, the number of Decisions, the number of Propagations, the number of Conflict literals, Memory usage and CPU Time.

We intend to study the symmetry that has been found in the occurrences of the variables during implementation. We also plan to find new van der Waerden numbers.

References

- [1] Tanbir Ahmed, *Some new van der Waerden numbers and some van der Waerden-type numbers*, Integers 9, A06, 65–76. MR2506138, 2009.
- [2] Tanbir Ahmed, *Two new van der Waerden numbers: $w(2; 3, 17)$ and $w(2; 3, 18)$* , Integers, 10, A32, 369–377, 2010.
- [3] Gilles Audemard and Laurent Simon. Glucose <http://www.labri.fr/perso/lsimon/glucose>
- [4] Gilles Audemard and Laurent Simon. *Glucose and Syrup in the SAT'17*. In Proceedings of SAT Competition 2017 , Solver and Benchmark Descriptions, pp.16–17, 2017.
- [5] M. Buro, H. Kleine Büning. *Report on a SAT competition*. Technical report, University of Paderborn, November 1992.
- [6] Martin Davis, George Logemann and Donald Loveland. *A machine program for theorem-proving*, Communications of the ACM,5(7):394–397, MR0149690, 1962.
- [7] Michael R. Dransfield, Lengning Liu. *Satisfiability and computing van der Waerden numbers*. The ELECTRONIC JOURNAL OF COMBINATORICS 11, 2004.
- [8] Niklas Eén, Niklas Sörensson. MINISAT <http://minisat.se>
- [9] Niklas Eén, Niklas Sörensson. *An Extensible SAT-solver*. In Theory and Applications of Satisfiability Testing, pp. 333–336. Springer, 2004.
- [10] Jon William Freeman. *Improvements to Propositional Satisfiability Search Algorithms*, PhD thesis, Departement of computer and Information science, University of Pennsylvania, Philadelphia, 1995.
- [11] Eugene Goldberg, Yakov Novikov. *BerkMin: A fast and robust SAT-solver*, In Proceedings of the Conference on Design Automation and Test in Europe pp. 142-149 Paris France Mar. 2002.
- [12] Daniel Jackson and Mandana Vaziri. *Finding Bugs with a Constraint Solver*, Proc. International Symposium on Software Testing and Analysis (ISSTA'00), Portland, Oregon, August 2000.
- [13] Robert G. Jeroslow and Jinchang Wang *Solving propositional satisfiability problems*. Annals of Mathematics and Artificial Intelligence, Volume 1 pp.167–187, 1990.

- [14] Chuan Jiang and Ting Zhang. *Nigma: A Partial Backtracking SAT Solver*. In Proceedings of SAT Competition 2013 , Solver and Benchmark Descriptions, pp. 62–63 , 2013.
- [15] Chuan Jiang and Gianfranco Ciardo *Nigma 1.1*. In Proceedings of SAT Competition 2014 , Solver and Benchmark Descriptions, p. 53 ,2014.
- [16] Chuan Jiang and Gianfranco Ciardo *Nigma 1.2*. at SAT Race 2015, 2015.
- [17] Henry Kautz and Bart Selman. *Planning as Satisfiability*. In European Conference on Artificial Intelligence, volume 92, pp. 359–363, 1992.
- [18] Bruce Landman, Aaron Robertson, Clay Culver. Some new exact van der Waerden numbers, *Integers: ELECTRONIC JOURNAL OF COMBINATORIAL NUMBER THEORY*, 5(2) (2005), A10, MR2192088.
- [19] João P. Marques-Silva, Karem A. Sakallah. *GRASP: A New Search Algorithm for Satisfiability*, *IEEE Transactions on Computers*, 48: 506–521, 1999.
- [20] Matthew W. Moskewicz, Conor F. Madigan, Ying Zhao, Lintao Zhang, and Sharad Malik. *Chaff: Engineering an Efficient SAT Solver*. In Proceedings of the 38th conference on Design Automation, pp. 530–535, New York,USA, 2001.
- [21] Sajjad Siddiqi and Jinbo Huang. *multi-SAT: An Adaptive SAT Solver* . In Proceedings of SAT Competition 2016 , Solver and Benchmark Descriptions, p.54 , 2016.
- [22] Mate Soos. Cryptominisat <http://www.msoos.org/cryptominisat2>
- [23] Paul R. Stephan, Robert K. Brayton, Alberto L.Sangiovanni-Vincentelli. *Combinational Test Generation Using Satisfiability*, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, pp. 1167-1176, 1996.
- [24] Aolong Zha. *Glulu*. In Proceedings of SAT Competition 2017 , Solver and Benchmark Descriptions, p. 18 , 2017.
- [25] Lintao Zhang, Sharad Malik. *The Quest for Efficient Boolean Satisfiability Solvers*, In Proceedings of the 14th International Conference on Computer Aided Verification . Springer-Verlag, July 2002.

A proposed Architecture for Shopping Association Mining

Essam Eldin Mosallam¹, Reda Abd Elwahab, Alkhoribi², Mohamed Ragaie Sayed Osman³

Abstract

This paper presents Mobil application for association mining of shopping based on Apriori algorithm. The proposed system shows and displays some offers and deals from various branches. The system provides the analytics for the seller; as the demand for some varieties and their weakness in other varieties and the whole application is organized on the Cloud. The architecture includes three levels; the front-end, middle and back-end. The front-end level of the site-based Mobile shopping application is made up of Android Mobile devices, to buy miscellaneous products from various nearby branches. The front-end level also displays the link between items purchased. The middle repository level provides a Web service to generate returned from relational database. The Exchanged information and data between application and servers is stored in the Cloud. The background level offers a Web server and a MySQL database. In this paper, we propose a architecture that reduce the communication overhead in existing Mobile Agent-based Distributed Association Rule Mining (MAD-ARM).

Keywords: Cloud, MCC, SaaS, Market Basket Analysis (MBA) and (MAD-ARM).

1. INTRODUCTION

We know that mobile devices are constrained by their processing power, battery life and storage. However, cloud computing provides an illusion of infinite computing resources. Mobile cloud computing is a new platform combining the mobile devices and cloud computing to create a new infrastructure, whereas cloud performs the heavy lifting of computing-intensive tasks and storing massive amounts of data. In this new architecture, data processing and data storage happen outside of mobile devices. Cloud computing combined with data mining can provide powerful

¹ Master student in Arabic Academy for Science, Technology and Maritime Transport.

² Vice Dean for Educational and Student Affairs, Faculty of Computers and Information, Cairo University.

³ Assistant Professor, Arabic Academy for Science, Technology and Maritime Transport, College of Computing and Information Technology

capacities of storage and computing and an excellent resource management. Data mining in the cloud computing environment can be considered as the future of data mining because of the advantages of cloud computing paradigm. Mobile cloud computing (MCC) is a new emerging research field. Today's mobile devices have many advanced features such as mobility, communication, and sensing capabilities, and can serve as the personal information gateway for mobile users. However, when running complex data mining and storing operations, the computation, energy, and storage limitations of mobile devices demand an integrated solution relying on cloud-based computation and storage support. The Mobile shopping application consists of Mobile devices which limited memory and processing speed. The Cloud, on the other hand, provides a large storage and speed for data stores. The Mobile Cloud Computing (MCC) is an area where three things are involved: Mobile devices, networking, and the Cloud. Data is stored and processed on the Mobile devices on the cloud using a giant computer, and the results then are displayed through output devices such as the monitor. Google Drive, Gmail, Google Drive and Google Maps are already used as examples of Mobile cloud computing. Thus, Mobile cloud computing covers the disadvantages and weaknesses of Mobile devices such as short storage area and processing power. [1] The research is implemented in the Cloud for supermarket shopping products. Graphical user interface (GUI) was designed by using location-based services and association rule mining. This research covers these drawbacks and implements association rule mining on the data gathered from the Mobile application. [2]

- **Objective**

This paper is specially designed for product purchasing in the city for malls and a local market nearby the customer location. It uses Wi-Fi Global Positioning System (GPS) and Mobile network to get the current position of the customer and displays any registered supermarket's branch on the server from customer's location. Association rule mining as a technique of data mining is used to find the offers associated with products. Technically, data mining is the process of extraction of interesting information or patterns from data in the large database. Association rule mining is widely used in market basket analysis. This method benefits retailers in numerous ways for marketing or planning shelf space.

The paper organization

The paper is organized as follows:

2. Background: review the Mobile Cloud Computing (MCC), Market Basket Analysis (MBA), Association Rule Mining and Location-based services
3. Proposed Architecture: review the proposed, system architecture, software and data Set.
4. Related Work: reviews the work previously done by researchers in this field of interest.
5. Conclusions and Future works: final results of this research and future work suggestions.

2. Background

In this section we review the Mobile Cloud Computing (MCC), Market Basket Analysis (MBA), Association Rule Mining and Location-based services

2.1 Mobile Cloud Computing (MCC)

Mobile cloud computing consists of three modules: Mobile devices, communication network and cloud as a server. The application which is large can be decomposed to smaller ones to process concurrently. This method is called as application partition. Offloading is the process of transferring Mobile application to cloud. This saves the device memory, processing power, and ultimately battery consumption. The classic facilities needed by a Mobile cloud client are, synchronization, push, i.e., updates the notifications sent by the cloud server, offline App automatically handles synchronization and notification, network, database, inter App Bus; provides low-level coordination between applications.[15]

2.2 Market Basket Analysis (MBA)

Market Basket Analysis is a forming technique based on the theory that if you buy a certain group of items, you are more likely to buy another group of items. MBA uses this information to: (1) Understand why they make certain purchases, (2) Products which are purchased together, and (3) Products which might benefit from the promotion. This system has used the Market Basket Analysis method for analyzing the data. The following techniques are used in the analyzing process.[14]

2.3 Association Rule Mining

MAD-ARM is the Mobile Agent-based distributed data mining architecture. It contains knowledge server which works on the generation of association rule and data coming are from different remote sites. The item sets are always upgrading on remote sites at the immobile agent. [3]

The association rules have been recently recognized as an important tool for knowledge discovery in databases. The problem of discovering association rules was first investigated in pioneering work in [4]. Here we examine a database of records which consist of both customer profile (such as salary and age) and behavior (such as buying decision) information). The association rule problem was originally proposed for the case of binary itemset data.[9] The intuitive implication of the association rule $X \sim Y$ is that a presence of the set of items X in a transaction also indicates a strong possibility of the presence of the itemset Y . The measures used to evaluate the strength of the rule are support and confidence. The *support* of a rule $X \sim Y$ is the fraction of transactions which contain both X and Y . The *confidence* of a rule $X \sim Y$ is the fraction of transactions containing X which also contain Y . A considerable amount of research effort [5, 6], has been devoted to the problem of speeding up the itemset method for finding association rules from very large sets of transaction data. Several variations of this problem have been discussed in [7, 8]. The quantitative association rule problem is a generalization of the problem of discovering association rules on sales transaction data, in which both categorical and quantitative attributes are allowed.[10]

2.4 Location-based services

The Mobile location-based application for any shopping product was designed and developed to find nearby branch and stores in local markets, the association among the products purchased, display association to customer side screen, post order, and it is deployed on the Cloud (See Figure1). The grouping of web map service API and association rule mining using Mobile in the cloud, it is potential to gather large-scale shopping habit of people, with lower data collection cost. The developed application model represents an environment for data analysis, and the proposed Algorithm is chosen dynamically at each handset. This is based on the environment where data stream mining process runs on user's android handset. As the data streams in continuously, possible concept drift is updated. There is a specific central Mobile decision agent

which switches several others stream mining agents. Stream mining agents working on local Mobile phones decides the best possible algorithm to run on the local data.

3. Proposed Architecture

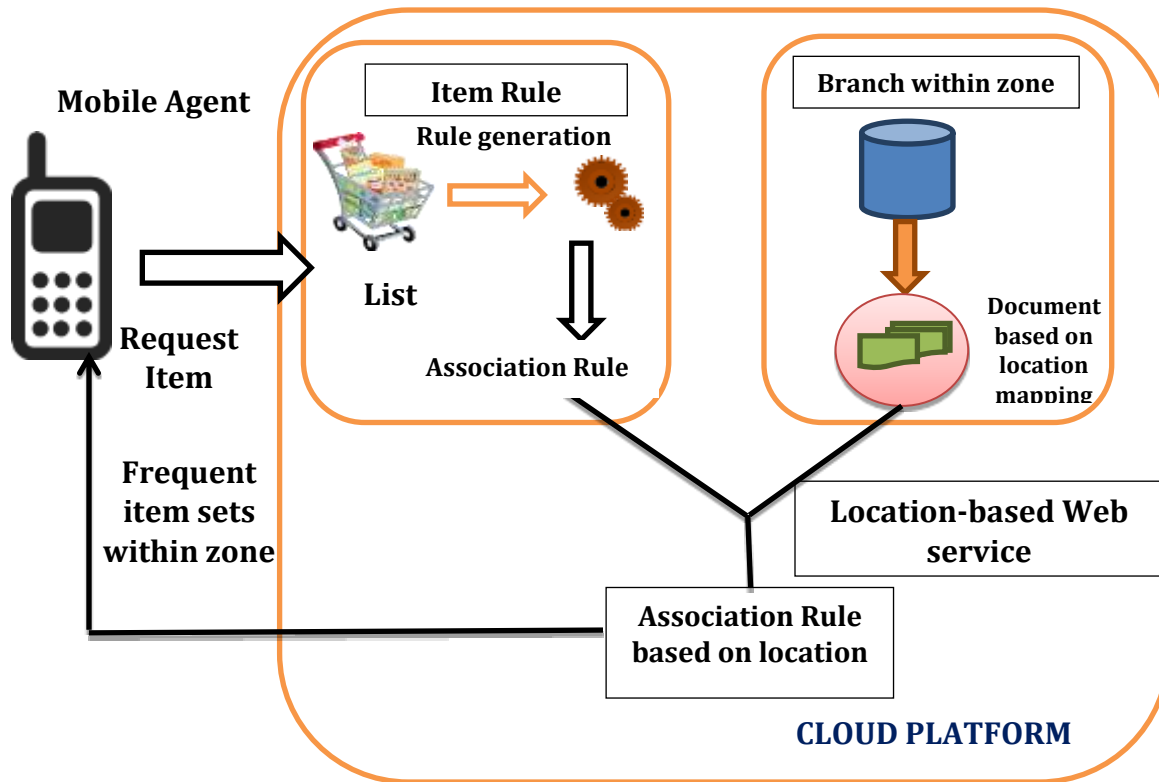


Figure 1 - Proposed Architecture Mobil Agent Association Mining System for Shopping

Any product or application that uses the location data of Mobile subscriber is called as location-based service. Location-based services like GPS uses the latitude and longitude data. Defines a location-based text mining approach to categorize texts into numerous categories based on their geospatial features, with the goals of discovering relationships between documents and area. There are three main modules in this architecture, including geographic data group and reprocessing, mapping forms into corresponding regions, and framing maximizes zones. Data mining and processing are takes place based on zones. The tourism industry has also taken the benefit of location-based services. This application is designed and established using the cloud based platform. It discovers out the location of tourists, where they are directed or looking.

3.1 System Architecture

The architecture of shopping application is divided into three parts as Front end, Middle ware level, and backend.

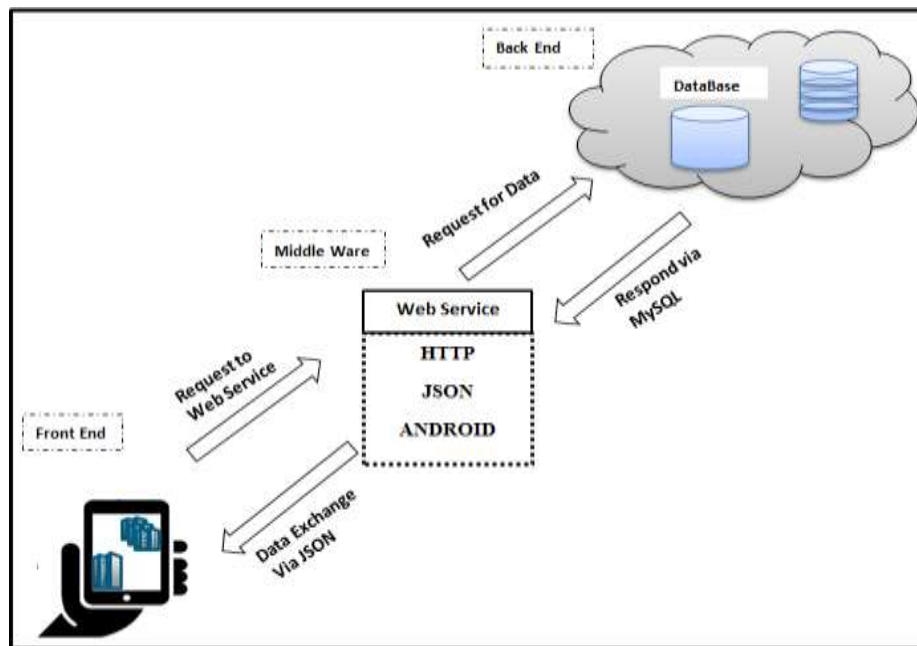


Figure 2: System Architecture. [11]

The Android open source platform is used to design and develop the shopping application. For the end user, in front-end user can able to select particular supermarket branch in the city and see the offers available in the specific showroom at the mall. It also provides offers of local market which are available to users nearby location. The registered retailers can upload and remove their offers and advertisement from this application. When the user searches for specific offers of the showroom in a mall, then the request is sent to middleware level that is to the Web Service. Web service acts like interface for front-end and back-end. The data exchange between front-end and back-end of shopping application happens via middleware level. Android shopping application

sends HTTP Request, and Web Service will send Query to fetch requested data from MySQL database located on the cloud (See Figure 2).

3.2 Software and data Set

In this paper we used data set form supermarket that contains 4700 Records, every record has 23 Items for purchasing case registration. By using and prepare the rules for Apriori diagram and test this data by using Weka application. The Weka is machine learning algorithm for data mining. Weka is efficient and has a user-friendly user interface. It is fully implemented by Java language there for it runs on almost any computing platform. However, it can only run in the local environment.

Apriori Algorithm

- **Product set:** a set of items
- **k-product_set:** a product set which consists of k items
- **Frequent product_set (i.e. large product_set):** an product_set with sufficient support
- **Lk or Fk:** a set of large (frequent) k-product_sets
- **ck:** a set of candidate k-product_sets
- **Apriori property:** if an item A is joined with item B, $\text{Support}(A \cup B) = \min(\text{Support}(A), \text{Support}(B))$

In the process of Apriori, the following explanations are needed:

Definition 1: Suppose $T = \{T_1, T_2, \dots, T_m\}$, (m_1) is a set of transactions, $T_i = \{I_1, I_2, \dots, I_n\}, (n_1)$

is the set of items, and $k\text{-product_set} = \{i_1, i_2, \dots, i_k\}, (k_1)$ is also the set of k items, and $k\text{-product_set} \subseteq I$.

Definition 2: $\text{Support}_t(\text{product_set})$, is the support count of product_set or the frequency of occurrence of a product_set in transactions.

Definition 3: Suppose C_k is the candidate product_set of size k, and L_k is the frequent product_set of size k.

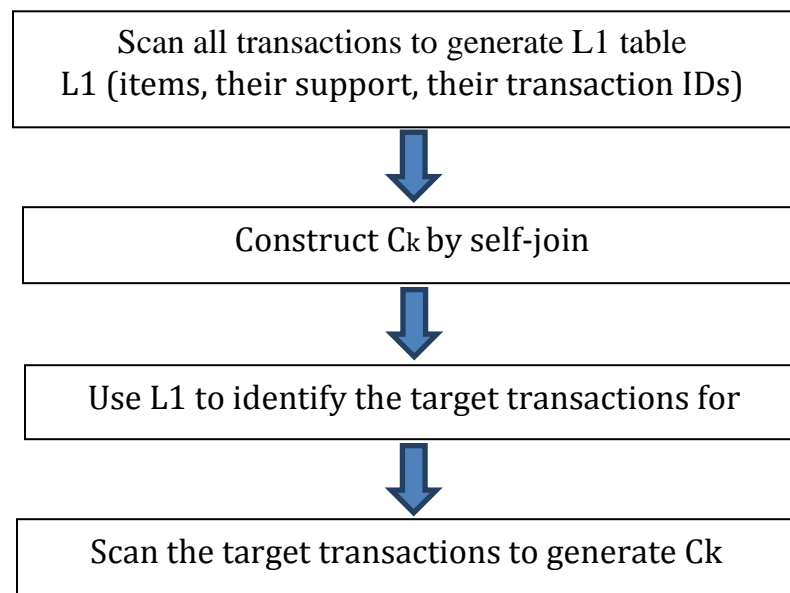


Figure 3: Steps for Ck generation

In the proposed approach, scan all transactions to create L1 which contains the products, their support count and Transaction ID where the products are found. And then use L1 later as a helper to generate L2, L3 ... Lk. When to create C2, make a self-join $L1 * L1$ to construct two product_set C (x, y), where x and y are the products of C2. Before scanning each transaction records to count the support count of every candidate, use L1 to get the transaction IDs of the least support count between x and y, and thus scan for C2 only in these specific transactions. The same thing for C3, construct three product_set C (x, y, z), where x, y, and z are the products of C3 and use L1 to get the transaction IDs of the least support count between x, y, and z, then scan for C3 only in these specific transactions and repeat these steps until no new frequent product_sets are identified. The whole process is shown in (Figure 3).

4. Related Work

- In [11] this paper designed and established a location-based mobile shopping application for malls and local markets for android platform. This application shows nearby local markets and mall's stores that are registered to the application. The main objective of marketing is achieved at a very low cost in comparison of advertisements, announcements, ground level marketing etc.
- In [12] this study has represented a recommendation engine by using association rules. The system had been tested by existing data in terms of the accuracy and the coverage. Best results have determined for 25 days data with 87.74% coverage and 16.43% accuracy. Processing time was 318 minutes for those data.
- In [13] this paper, we discussed the problem of online mining of profile association rules. Such rules may be very useful in developing relationships between consumer profiles and behavioral information. We discussed how to use multidimensional indexing to generate profile association rules in online fashion.

5. Conclusions and Future works

We designed this architecture after survey of the research on the use of mobile in the purchase process using an application on the cloud computing systems and tried to improve this service, also to monitor the influence of the factor of space and distance between the consumer and supermarkets and calculated the distances between the place of contact of the consumer and the nearest branches of supermarkets and the lowest prices in the same Time and therefore there are more opportunities for the consumer. We designed a mobile shopping architecture based on the site for the very large supermarket has many Branches of the Android platform. The main objective of marketing is achieved at a very low cost compared to advertising, and marketing at the regional level and so on. Data is exchanged between different levels of structural design using a web service station that generates a JSON format for data transfer. With Mobile Computing Cloud Computing Mobile processing and storage move to the cloud as a server, helping to save battery consumption and extend performance or speed of implementation. In this paper we designed architecture for the use of mobile in the purchase process of the cloud computing system and monitored the influence of the distance between the mobile and the nearest supermarket branch to the consumer. In future works, we can focus on small organization more precisely. The architecture can feature more options for live stations for small household business. A broad

market survey is needed to be done to provide marketing and selling for such developers and products.

REFERENCES

- [1] Emil Almrot, Sebastian Andersson. A study of the advantages & disadvantages of mobile cloud computing versus native environment. Bachelor Thesis in Software Engineering May 2013.
- [2] Raubal, M. and Rinner, C., 2004. Multi-Criteria Decision Analysis for Location Based Services. In S.A. Brandt (ed.) Proceedings of the 12th International Conference on Geoinformatics, pp. 47-53.
- [3] A.Saleem Raja, E.George Dharma Prakash Raj, “MAD-ARM: Mobile Agent based Distributed Association Rule Mining”, International Conference on Computer Communication and Informatics (ICCCI -2013), Jan. 09 – 11, 2013, Coimbatore, INDIA, DOI 10.1109/ICCCI.2013.6466112.
- [4] Agrawal R., Imielinski T., and Swami A. 1993. Mining association rules between sets of items in very large databases. Proceedings of the A CM SIGMOD Conference on Management of data, pages 207-216, Washington D. C.
- [5] Agrawal R., and Srikant R. 1994. Fast Algorithms for Mining Association Rules in Large Databases. Proceedings of the 20th International Conference on Very Large Data Bases, pages 478-499.
- [6] Savasere A., Omiecinski E., and Navathe S. 1995. An Efficient Algorithm for Mining Association Rules in Large Data Bases. Proceedings of the 21st International Conference on Very Large Data Bases. Zurich, Switzerland, pages 432-444.
- [7] Hun J. and Fu Y. 1995. Discovery of Multiple-Level Association Rules from Large Databases. Proceedings of the 21st International Conference on Very Large Data Bases. Zurich, Switzerland, pages 420-431.
- [8] Agrawal, R. and Srikant, R. Mining Generalized Association Rules. Proceedings of the 21st VLDB Conference. 1995
- [9] Agrawal R., Imielinski T., and Swami A. 1993. Mining association rules between sets of items in very large databases. Proceedings of the A CM SIGMOD Conference on Management of data, pages 207-216, Washington D. C.
- [10] Srikant R., and Agrawal R. Mining quantitative association rules in large relational tables. Proceedings of the 1996 ACM SIGMOD Conference on Management of Data. Montreal, Canada.
- [11] Shivani Rajput*, Rajeshree Pagar, Laxman Erande, Yogesh Alai. One Click Android Application for Shopping Based on Cloud. Proceedings of the 2015 IJSRSET | Volume 1 | Issue 5 | Print ISSN: 2395-1990 | Online ISSN: 2394-4099.
- [12] Ozgur Cakir, Murat Efe Aras. A recommendation engine by using association rules Procedia - Social and Behavioral Sciences 62 (2012) 452 – 456.

[13] Charu C. Aggarwal, Zheng Sun, Philip S. Yu. Online generation of profile association rules - T.J. Watson Research Center Yorktown Heights, NY 10598 charu@watson.ibm.com, Duke University, Durham, NC-27706 sunz@cs.duke.edu, T.J. Watson Research Center, Yorktown Heights, NY 10598 psyu@watson.ibm.com.

[14] Manpreet Kaur, Shivani Kang. Market Basket Analysis: Identify the changing trends of market data using association rule mining, Procedia Computer Science 85 (2016) 78 – 85.

[15] S M Shamim, Angona Sarker, Ali Newaz Bahar, Md. Atiqur Rahman. A Review on Mobile Cloud Computing. International Journal of Computer Applications (0975 – 8887), Volume 113 – No. 16, March 2015.

Mining IoT Data Streams

Zaki. A. M.¹

Abstract

Internet of Things “IoT” is the natural evolution of wireless sensor networks, mobile computing, and cloud computing. IoT is the result of trying to benefit from networking advances, hardware advances by connecting every possible object to the internet. The objective of IoT is to make every connected object intelligent enough to take decision to help people in their live. In this paper the concept of mining IoT data streams is reviewed. The mentioned data streams are the streams of data that are captured by or from IoT’s objects. This research gives a short background on IoT and state a set of research questions to answer. Then, a discussion to answer research question is conducted.

Keywords

IoT, Data Streams, Edge Analytics, Edge Mining

1. Introduction

In the last decade and due to advances in networking and virtualization a new term called cloud computing has arises. In cloud computing, processing and storage moved to the cloud and I/O, lite processing and lite storage remains on the client. Cloud computing goal was to provide a powerful and scalable services, platforms and infrastructures to satisfy business needs on demand [Mell, 2011].

On the other hand, there were advances in mobile computing. Mobile devices – including sensors, mobile phones, tablets, and laptops – have become more powerful in processing, memory, storage, networking and battery. Additionally, huge number of these mobile devices are produced and connected to Internet. Not only mobile devices which is connected to Internet. There are other products that are now contain an embedded computer and connected to internet such as cars, refrigerators, watches, TVs, air conditioner, satellite receivers, traffic lights, surveillance

¹ a.m.zaki@sadatacademy.edu.eg

Information Systems and Technology Department, Institute of Statistical Studies and research, Cairo University

cameras and weather sensors. All mentioned devices generate streams of data with different types. While these objects are connected to internet a new term coined which is Internet of Things “IoT”. Traditionally generated data is transferred to backend systems to be cleansed, stored, and analyzed.

Machine learning algorithms are used to make these things “Objects” intelligent. A new approach to handle these data is to bring intelligence on the edge of the network. The new approach called edge analytics. In edge analytics approach, data streams are processed during the journey to the cloud [Gaura, et al., 2013] [Simoens, 2015] [HP, 2017]

The objective of this research is to provide a complete picture of current literature in mining IoT data streams at the edge of networks. This research tries to answer a set of questions stated in section 2. In section 3, IoT concept is discussed. Section 4 shows big data analytics approach and its drawbacks. Section 5 introduces the concepts of edge computing, edge analytics. Section 5 also discusses edge computing motivations, challenges, opportunities and applications. Section 6 reviewed data stream mining concepts and shows the algorithms and models of data mining used with IoT. Finally, in section 7, Open issues of IoT is discussed.

2. Research Questions

Research questions to be addressed are:

Table 1 Research Questions

RQ#	Description
RQ1:	What are the existed approaches to mine data streams from IoT devices?
RQ2:	What are the reasons leading to the need to make analysis on the edge device?
RQ3:	Is there any algorithm that is already adapted to work on edge devices?
RQ4:	What are the benefits that returns by mining data streams on edge devices?

These questions could be expanded to be as follow:

RQ[1]. What are the existed approaches to mine data streams from IoT devices?

RQ[1.1] What are IoT Devices?

RQ[1.2] What are IoT devices characteristics?

RQ[1.3] How much Data IoT Devices collect?

RQ[1.4] How to get benefit from this data?

RQ[1.5] What are the different approaches to handle these data?

RQ[2]. What are the reasons leading to the need to make analysis on the edge device?

RQ[2.1] How traditional approach for handling IoT data Streams works?

RQ[2.2] What are the drawbacks of the traditional approach?

RQ[3]. Is there any algorithm that is already adapted to work on edge devices?

RQ[3.1] What are the existed machine learning algorithms for mining data streams?

RQ[3.2] Which algorithm could be adapted to work on edge devices?

RQ[3.3] What are algorithms that are already adapted to work on edge devices?

RQ[4]. What are the benefits that returns by mining data streams on edge devices?

RQ[4.1] What are the different applications of Edge Analytics?

RQ[4.2] What are the benefits from these kinds of applications?

RQ[4.3] What are the side effects of these applications on society?

3. Internet of Things

In this section, researcher tries to identify what Internet of Things “IoT” is and what IoT devices are. A lot of researches try to answer these questions. IoT could be defined as “The worldwide network of interconnected objects uniquely addressable based on standard communication protocols” which is RFID group definition according to [Gubbi, et al., 2013]. This definition is a network based definition which focuses on communication protocols and networking. It ignores the nature of connected object. Another view of IoT is an intersection between Internet, Thing and semantic oriented visions [Atzori, et al., 2010]. Figure 1 represents the previous definition as it considers the nature of object and semantics of environment and communication standards [Atzori, et al., 2010].

Gubbi give more general definition for IoT. They define IoT as “Interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified framework, developing a common operating picture for enabling innovative applications. This is achieved by seamless ubiquitous sensing, data analytics and information representation with Cloud computing as the unifying framework” [Gubbi, et al., 2013] . From the previous three definitions, IoT is to connect objects “Things” together with a cloud based backend systems over the internet, these things can sense, generates and capture data and then transmit that data to the cloud to be analyzed, visualized and stored. Sometimes these objects have an actuator to do tasks. IEEE defines IoT as “Internet of Things envisions a self-configuring, adaptive, complex network that interconnects ‘things’ to the Internet through the use of standard communication protocols. The

interconnected things have physical or virtual representation in the digital world, sensing/actuation capability, a programmability feature and are uniquely identifiable. The representation contains information including the thing's identity, status, location or any other business, social or privately relevant information. The things offer services, with or without human intervention, through the exploitation of unique identification, data capture and communication, and actuation capability. The service is exploited through the use of intelligent interfaces and is made available anywhere, anytime, and for anything taking security into consideration.” [Initiative, 2015] .In the previous definition IEEE was trying to define IoT with covering all perspectives and points of view.

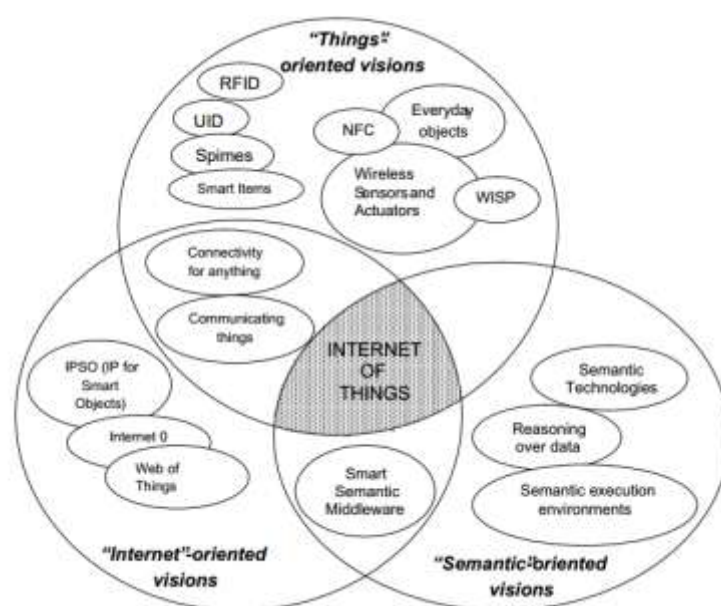


Figure 1 ‘Internet of Things’ paradigm as a result of the convergence of different visions, Source: [Atzori, et al., 2010]

The next question is what type of data that could be captured are and how much this data is and how to deal with this data. Types of data that could be captured are text in all formats, images and videos. Traditionally, all data captured and send to cloud over network. When the cloud receives those data, backend systems starts to clean and store it in form that facilitates analysis, visualization process.

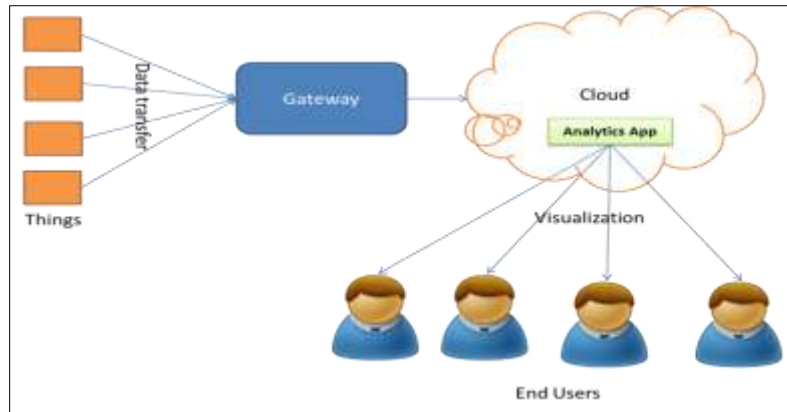


Figure 2 Simple illustration of Big Data Analytics

Actually, data type is not the problem. The major problem is data size, there is a forecasting tells that by the year 2020, there will be 26 billion connected devices and data size will reach to 40 zettabytes as illustrated in Figure 3 [Siemens, 2017].

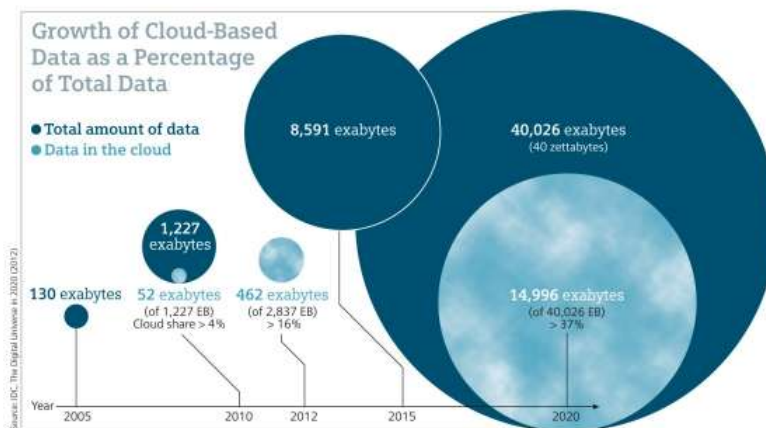


Figure 3 Growth of Cloud Based data as a Percentage of total data source [Siemens, 2017]

With this massive amount of data, the question arises, how to handle this amount of data. The next question, dose it worth to store these huge amounts of data. These problems will be discussed in the next section.

4. Big Data Analytics approach and drawbacks

As discussed earlier, the existed approach of big data assume that data is collected from sensors and other things, and then transferred to backend system on the cloud. These systems are responsible for preprocess, store, analyze and visualize analysis results to users.

This approach has its own drawbacks. These drawbacks are seven according to HP [HP, 2017]: latency, bandwidth, cost, threats, duplication, corruption and compliance. These are the drawbacks of existing approach and reasons for moving towards edge analytics. Table 2 summarizes these seven drawbacks. From researcher point of view these drawbacks may be solved by moving analysis into edge. But other problems may occur.

Table 2 : drawbacks of Existing approach summary

Drawback	Description
Latency	Network latency may cause a disaster in critical real time application. Where in existing approach it may take minutes before response.
Bandwidth	With huge amount of data, it consumes bandwidth to upload data on to the cloud which will also affect network latency.
Cost	To send huge amount of data to cloud, it needs high bandwidth which costs a lot of money. In some cases, it is not needed to move all data from sensors into backend.
Threats	Security Threats like while sending data over network (security attacks)
Duplication	Data could be duplicated while sending to the cloud
Corruption	During data transmission, data may be get corrupted due to any network issue.
Compliance	In some critical real time application, it is not allowed to move data outside country boundary. If the cloud is outside country this will cause a problem.

5. Edge Analytics

Edge analytics refers to the processing on the edge device at the edge of the network. In simple words, processing and analysis is done before sending anything to the cloud. It could be done on the edge “IoT” device or on the network gateway. According to [Varghese, et al., 2009], there are two layers of edge computing devices. The first is edge device which includes “Things”. The second layer is edge nodes which include networking devices such as routers, switches and base stations.

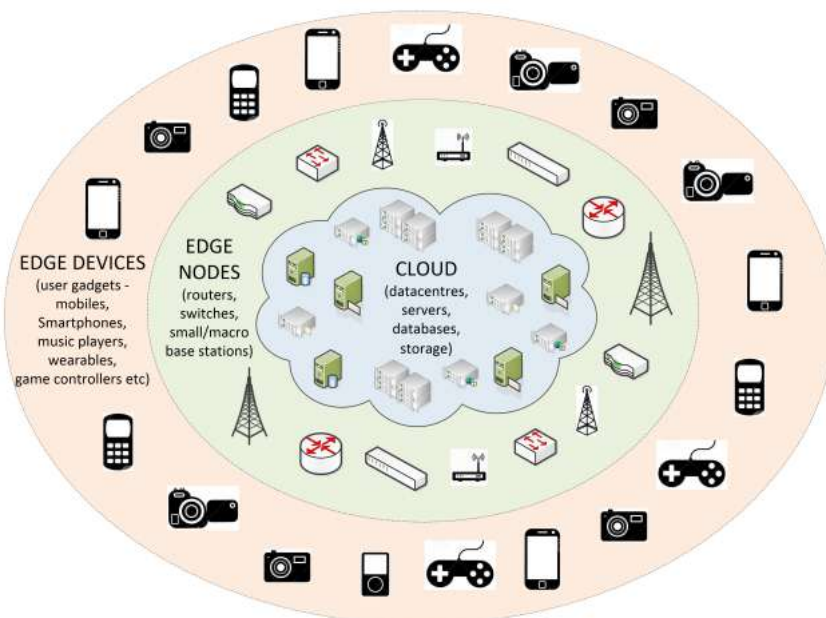


Figure 4 Edge Devices, Edge nodes and Cloud
source [Varghese, et al., 2009]

Edge analytics provide a solution for the previous drawbacks of big data model. But this also faces challenges. One of these challenges is that IoT devices are limited or constrained in resources compared with resources on the cloud. IoT devices may face a problem in power consumption and battery size. If analysis happened on network gateway, this will solve the power problem. But lead to another problem which is network gateways is a specific purpose computer system. Which means it needs to be modified or replaced with another system capable to do analysis. Another challenge is to adapt algorithms to run on such devices. Challenges and opportunities are discussed in more details in next subsection.

5.1. Edge computing motivations, challenges and Opportunities

Varghese proposes summary of motivation, challenges and opportunities in edge computing in Figure 5 [Varghese, et al., 2009]. According to [Varghese, et al., 2009] , motivations of edge computing are decentralizing cloud, enhancement in computational resource of client devices, huge energy consumption at the cloud, data explosion and smart computation techniques. Each of these motivations is discussed in

Table 3.

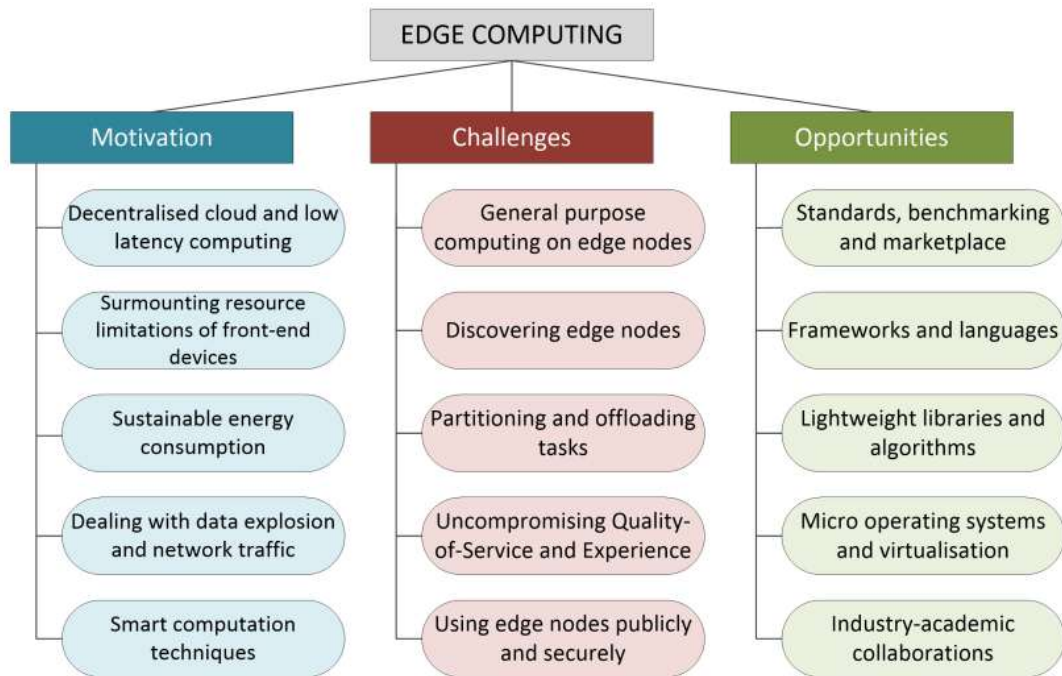


Figure 5 Edge Computing Motivation, Challenge and Opportunities

Source: [Varghese, et al., 2009]

Table 3: Edge Computing Motivation

Motivation	Description
Decentralized cloud	In real time applications, using central cloud to process data stream dose not fulfill time requirements. So, moving the processing to edge node or edge device satisfy time constrains. Decentralization is useful in such cases.
Enhanced Resources in Client Devices	Recently, edge devices could be considered resource limited if it compared to servers or workstations. But it is not resource limited to do filtrations and analytics.
Sustainable energy consumption	Data centers at Cloud consume huge amounts of energy and this amount is increasing. By using edge devices and nodes in some tasks, this leads to minimizing energy consumption at the cloud.
Data Explosion and network traffic	As mentioned earlier, data sizes are increasing dramatically due to the increase in number of connected devices. This affects network bandwidth while transferring data to the cloud. By moving processing on the edge, this help in minimizing data to transferred to the cloud.

Motivation	Description
Smart computation techniques	Using computations techniques to distribute the application using the idea of pipelining the processing in horizontal fashion in which the processing is starting at edge device, then edge node and finally at the cloud. Another technique is to use computational offloading in which the cloud offloads computational tasks to edge devices and node to be processed.

On the other hand of these motivations there are challenges that face edge computing. These challenges according to [Varghese, et al., 2009] are general purposes computing on edge nodes, edge node discovering, task partitioning and computational offloading, quality of services issues and edge nodes security concerns. Table 4 describes a summary of these challenges.

Table 4: Edge Computing Challenges

Challenge	Description
General Purpose computing on edge node	Edge nodes like routers, switches and base stations are not general purposes computers. According to that it is not designed to do any processing except the one it designed for.
Discovering edge node	Automatic resources discovery services for edge nodes are not existed. All applied methods in cloud environment do not apply on edge nodes or devices.
Task Partitioning and Computational offloading	Task partitioning or task offloading refers to the idea of distributing application between edge node or device and cloud, the major challenge here is how to determine which portion of task will be executed and where .
Quality Of Services “QoS”	This challenge is to use edge node or device efficiently to meet user expectations in service level without overloading edge node or device.

Finally, Varghese addresses a set of opportunities that could be done in academic researches. These opportunities are first creating benchmarks, standards, frameworks and toolkits for edge computing. Another opportunity is to create lightweight libraries and algorithms that fit edge

nodes and devices. Additional research area is micro operating systems, mobile containers for edge nodes and devices. [Varghese, et al., 2009]

5.2. Edge Analytics Applications

There are a lot of edge analytics on IoT Applications. A sample application of edge analytics on IoT is Surveillance cameras. According to [Simoens, 2015], surveillance cameras are common in police cars and also available for commercial uses. In [Simoens, 2015], they propose an architecture called GigaSight for video streaming analytics purposes. Another application of edge analytics is health care. Rahmani et al, propose architecture for an IoT-based health monitoring that could be deployed inside hospitals or homes. It depends on health monitoring sensors and smart e-health gateway. These edge devices and node do the analysis and call for help on time [Rahmani, et al., 2015] . Another example is smart city which introduced in [Bélissent, 2010] . According to Bélissent, IoT could be used in areas such as transportation, healthcare, education, public safety and security, building management, city administration and waste management. By adopting IoT in those areas, smart city will be existed. Bélissent also provide a comparison between new cities, existed cities and non-cities from the point of obstacles to apply smart city.

6. Data Stream mining

Data stream mining is concerned with mining streams of real time data generated from sensors in Wireless Sensor Networks “WSN”. Data Stream mining is one of two categories of data streams processing. According to [Gaber, et al., 2010] data streams processing has two categories which are data stream management and data stream processing. Data stream management is preparing data stream to other processing by querying and summarizing it. Whereas data stream mining is to apply data mining techniques on this stream of data. According to [Ullman, et al., 2014] , data stream management system has a main component for dealing with streams which is stream processor. Stream processor handles ad-hoc queries on stream input and store data which is needed and other data dose not stored. Figure 6 shows the idea of data streams management system.

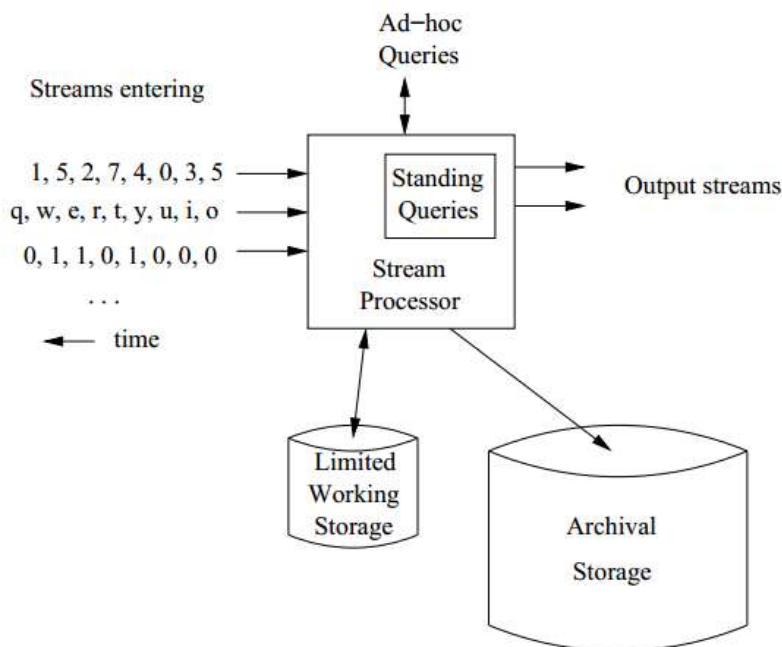


Figure 6 - Data Stream Management System
Source [Ullman, et al., 2014]

6.1. Data Mining for IoT

According to [Shen, et al., 2010] , there are four proposed data mining models for IoT. These four models are multi-layer model , distributed model , grid based model and multi-technology based model. In multi-layer model there four layers: data collection layer, data management layer, event processing layer and data mining service layer. Data collection layer is responsible for collecting objects data using RFID readers, GPS and any other adopted sensor. The next layer is data management layer which responsible for managing collected data using centralized or distributed database or warehouse. Then collected data is cleaned and processed to produce analytics. The third layer is event processing layer which is responsible for events filtering and events detection. This layer filters all unrequired events and keeps events that users are interested in. Finally, the last layer is data mining service layer. Data mining service layer is responsible for build learning models using data mining techniques to extract knowledge from objects data and events data [Shen, et al., 2010].

On the other hand, distributed data mining model tries to cope with IoT data characteristics, hardware limitations and security requirements

[Shen, et al., 2010]. According to [Shen, et al., 2010], distributed data mining model preprocess data locally at node and send to central backend system only required and necessary data for processing. The third model according to [Shen, et al., 2010] is grid based data mining model. In this model, IoT devices are considered as computing resources for a grid. This grid is used later for data mining tasks. The last model that was proposed by [Shen, et al., 2010] was based on mixing networking technologies, sensing technologies and mining algorithms then provide “Intelligence” to applications.

Additionally, [Shen, et al., 2010] discusses challenges of mining IoT data. These challenges were about data collection issues such as data sizes, data transmission and energy utilization. Another challenge is about what model to adopt in data centralized processing or distributed processing. Also [Shen, et al., 2010] refers to the challenge of studying data mining algorithms that are suitable for IoT.

In 2013, a paper entitled “Edge Mining the Internet of Things” define a term “Edge Mining” as “Processing of sensory data near or at the point which it is sensed, in order to convert it from a raw signal to contextually relevant information” [Gaura, et al., 2013] . This paper introduces four mining algorithms that used to reduce network traffic and reduce energy consumption. These algorithms are “General Spanish Inquisition Protocol” G-SIP [Gaura, et al., 2013] , “Linear Spanish Inquisition Protocol” L-SIP [Gaura, et al., 2013] , “ClassAct” [Gaura, et al., 2013] and “Bare Necessities” [Gaura, et al., 2013] BN. Figure 7 summarize edge mining process according to [Gaura, et al., 2013]

G-SIP and L-SIP are based on SIP. The basic idea of SIP is to send unexpected information to control center or to base station [Gaura, et al., 2013]. ClassAct is a classifier for human posture [Gaura, et al., 2013] . BN is an algorithm that calculates time spent in specific state [Gaura, et al., 2013].

In 2014, a survey paper about data mining for internet of things. This paper reviews motivation and problems then reviewed solution and results. Finally, it gives future directions and trends in this area. This paper was written by [Tsai, et al., 2014]

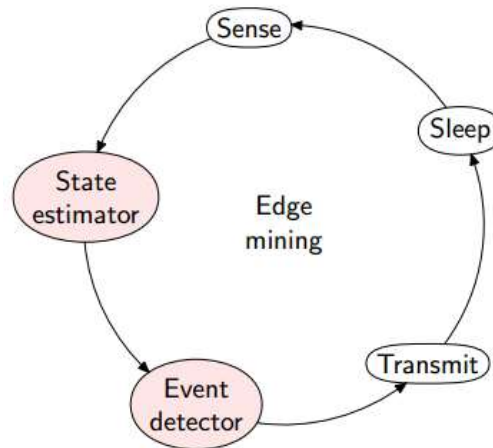


Figure 7: Edge Mining Process at node

Source: [Gaura, et al., 2013]

In [Tsai, et al., 2014] , researchers start with describing data captured from IoT and what this data about and data sizes. Tsai and the others refer to possible solutions of handling big data generated from IoT. Additionally, they refer to the benefits gained from analyzing IoT data. They also provide architecture for knowledge discovery and IoT. Figure 8 illustrates Tsai architecture [Tsai, et al., 2014]. This architecture starts with collecting data and preprocesses it then extracts knowledge then update the IoT device with that knowledge.

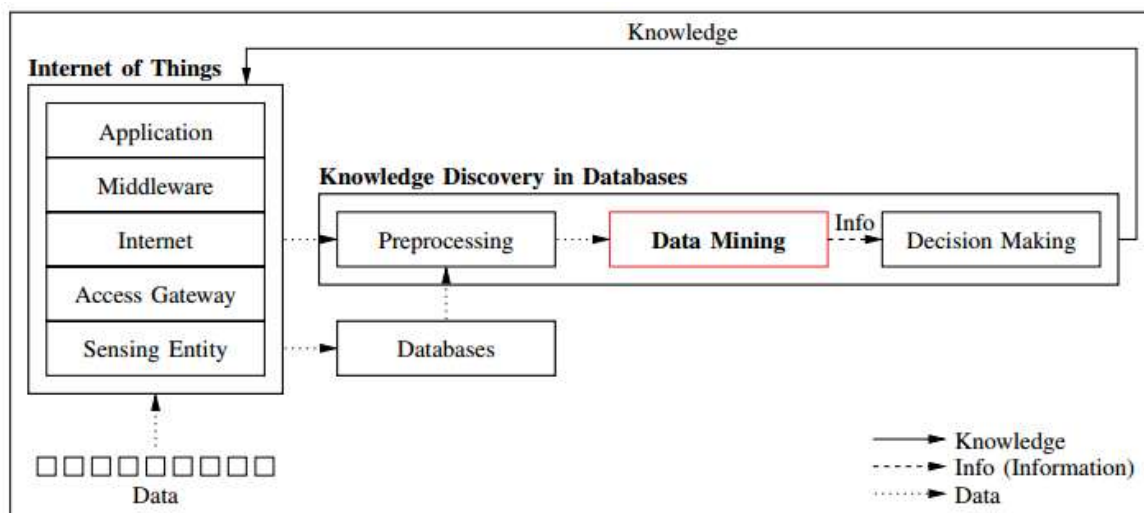


Figure 8: Architecture of IoT with KDD
source [Tsai, et al., 2014]

In [Tsai, et al., 2014], there are three considerations for choosing suitable data mining technique for IoT. These Considerations are objective, data and mining algorithm. Objective is about problem assumptions, limitations and measurement way. Data consideration is about data size, distribution and type. The last consideration is algorithm. The challenge is how to determine data mining algorithm. Additional challenge about algorithm, will the problem require to develop new algorithm or just adopt old one [Tsai, et al., 2014].

Tsai present a framework for data mining, called unified data mining framework. This framework is used to explain all mining algorithm that covered in their research. This framework is based on scan, construct and update operation. Scan input data set, then construct rules as output then update these rules. The previous operations continue until met a termination criterion [Tsai, et al., 2014]. Figure 9 shows the unified data mining framework presented by Tsai.

Algorithm 1 Unified Data Mining Framework		
1	Input data D	
2	Initialize candidate solutions r	
3	While the termination criterion is not met	
4	$d = \text{Scan}(D)$ [Optional]	S
5	$v = \text{Construct}(d, r, \theta)$	C
6	$r = \text{Update}(v)$	U
7	End	
8	Output rules r	

Figure 9 Unified data mining framework source [Tsai, et al., 2014]

According to Tsai, data mining techniques could be used for IoT infrastructure or services [Tsai, et al., 2014]. They discuss data mining techniques in terms of the unified data mining framework. Their study investigates each data mining technique for IoT from two perspectives. First perspective is infrastructure perspective (i.e. the design consideration of algorithm to fit in IoT). The second prospective is which is the service algorithm will be used in (i.e. used in which application). The researcher illustrates this taxonomy in Figure 10.

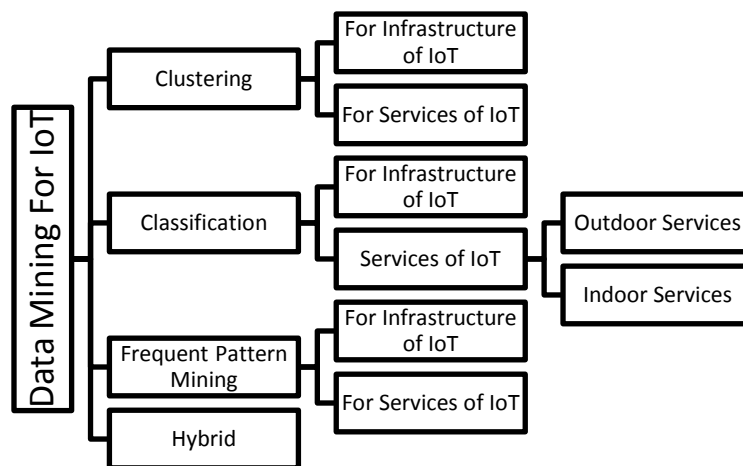


Figure 10 data mining techniques for IoT

Moreover, [Tsai, et al., 2014] provide a comparison of mining technologies in terms of goals and data sources and in reference that study each mining technique. Figure 11 is adapted from [Tsai, et al., 2014] to

show only goal and data source for each mining technique.

Mining Algorithm	Goal	Data Source
Clustering	Network performance enhancement	wireless sensor
	Inhabitant action prediction Provisioning of the needed services Housekeeping Managing the plant zones Relationships in a social network	X10 lamp and home appliances Raw location tracking data Vacuum sensor GPS and sensor for agriculture RFID, smart phone, PDA, and so on
Classification	Device recognition Traffic event detection Parking lot management Inhabitant action prediction	RFID GPS, smart phone, and vehicle sensor Passive infrared sensor RFID, sensor, video camera, microphone, wearable kinematic sensor, and so on
	Inhabitant action prediction Inhabitant action prediction physiology signal analysis	Video camera microphone wireless ECG sensor
Frequent Pattern	RFID tag management	RFID
	Spatial colocation pattern analysis	GPS and sensor
	Purchase behavior analysis Inhabitant action prediction	RFID and sensor RFID and sensor
Hybrid	Inhabitant action prediction	RFID and sensor

Figure 11 Mining Techniques Comparison

Source: [Tsai, et al., 2014]

7. IoT Open Issues

Finally, Tsai et al refer to open issues of IoT and data mining. These open issues are from four perspectives: Infrastructure, Data, and algorithm and security perspectives [Tsai, et al., 2014] . Table 5 tries to summaries these open issues from the different perspectives.

Table 5 open issues in IoT

Perspective	Open Issues
Infrastructure	<ul style="list-style-type: none"> • Decentralization • Heterogeneity • Low power • Small memory, limited computation power
Data	<ul style="list-style-type: none"> • Different data standards, types and resources • Sizes of Data and preserving of data
Algorithm	<ul style="list-style-type: none"> • Dynamic nature of sensor status affects algorithms • Redesign of algorithm to fit in application. • Computation load and services balance.

Perspective	Open Issues
Security	<ul style="list-style-type: none"> • Privacy concern while using video cameras • Sensitive information of IoT end User (e.g. health info)

8. Conclusion

In this paper, a systematic literature review is conducted on mining IoT data Streams. The paper discusses the term IoT, edge computing, data stream mining and data mining for IoT. The paper starts with reviewing drawbacks of big data approach for handling IoT data. Then it reviews challenges and opportunities of edge computing. Additionally, Applications of edge analytics is mentioned. Later, concept of data stream mining is reviewed. After that data mining for IoT is discussed. At the end, it could be notices that a lot of challenges still need research effort. This Research area is still open and there are a lot of opportunities. It is a hot topic for research.

References

- [1] Atzori Luigi, Iera Antonio and Morabito Giacomo The Internet of Things: A survey [Journal] // Computer Networks. - 2010. - Vol. 54. - pp. 2787-2805. - ISSN: 1389-1286.
- [2] Bélissent Jennifer Getting clever about smart cities: new opportunities require new business models [Journal] // Forrester Research, inc. - 2010. - p. 33.
- [3] Gaber Mohamed Medhat, Zaslavsky Arkady and Krishnaswamy Shonali Data Stream Mining [Book Section] // Data Mining and Knowledge Discovery Handbook / book auth. Ratanamahatana Chotirat AnnLin, Jessica Gunopulos, Dimitrios Keogh, Eamonn Vlachos, Michail Das, Gautam. - 2010.
- [4] Gaura Elena I [et al.] Edge mining the Internet of Things [Journal] // IEEE Sensors Journal. - 2013. - pp. 3816-3825.
- [5] Gubbi Jayavardhana [et al.] Internet of Things (IoT): A vision, architectural elements, and future directions [Journal] // Future Generation Computer Systems. - 2013. - Vol. 29. - pp. 1645-1660. - ISSN: 0167-739X.
- [6] HP IoT analytics at The Edge [Report]. - 2017.
- [7] Initiative I. E. E. E. Internet Towards a definition of the Internet of Things (IoT) [Journal]. - 2015. - p. 27. - ISBN: 9781509019410 ISSN: 09758887.

- [8] Mell Peter M. and Grance, Timothy SP 800-145. The NIST Definition of Cloud Computing [Report]. - Gaithersburg, MD, United States : National Institute of Standards & Technology, 2011.
- [9] Rahmani Amir Mohammad [et al.] Smart e-Health Gateway: Bringing intelligence to Internet-of-Things based ubiquitous healthcare systems [Journal] // 2015 12th Annual IEEE Consumer Communications and Networking Conference, CCNC 2015. - 2015. - pp. 826-834. - ISBN: 9781479963904 ISSN: 2331-9860.
- [10] Shen Bin, Liu Yuan and Wang Xiaoyi Research on Data Mining Models for the Internet of Things [Journal]. - [s.l.] : IEEE, 2010. - 978-1-4244-5555.
- [11] Siemens Internet of Things: Facts and Forecasts [Online] // Siemens.com. - Siemens, 2017. - April 2017. - <https://www.siemens.com/innovation/en/home/pictures-of-the-future/digitalization-and-software/internet-of-things-facts-and-forecasts.html>.
- [12] Simoens Pieter Edge Analytics in the Internet of Things [Journal]. - 2015. - pp. 24-31.
- [13] Tsai Chun-Wei [et al.] Data Mining for Internet of Things: A Survey [Journal]. - [s.l.] : IEEE COMMUNICATIONS SURVEYS & TUTORIALS, 2014. - 1 : Vol. 16.
- [14] Ullman J. D., Leskovec J and Rajaraman A Mining Data Streams [Book Section] // Mining of massive datasets. - 2014.
- [15] Varghese Blesson [et al.] Challenges and Opportunities in Edge Computing [Journal]. - 2009. - ISBN: 978-1-5090-5263-9.

Arabic Named Entity Recognition

Mariam Muhammad*

Abstract

Named Entity Recognition (NER) is an important task that is used in recognizing the proper names in the text such as “Person”, “Location”, and “Organization”. This task is very important in the most fields of the natural language processing (NLP) such as “Question Answering”, “Machine Learning”, and “Information retrieval”. There are more research studied the NER for the foreign language e.g. English but little of them that studied it for Arabic language because Arabic language has complicated morphology and that make the NER task difficult. In this research, we tried to know more about the Named Entity Recognition task and its importance in the NLP fields especially when using it with the Arabic language text. The challenges that faced the Arabic NER were studied and some the solutions for each challenge were introduced. The linguistic resources and tools that support the Arabic NER were presented and a link for each was illustrated. The main approaches of the NER were explained and an example for each approach was given. The evaluation metrics that are used for evaluating the Arabic NER also were presented.

Keywords: Arabic NER, NER Challenges, Linguistic Resources, NER Tools, NER approaches.

1. Introduction

In the 1990s, NER was introduced as a subtask of information extraction (IE) task and became important in more of the research studies [Shaalán, 2014].

The Named Entity Recognition (NER) task aims at detecting and classifying the proper names in the text. The proper names can be person, location or organization names [Nadeau & Sekine 2007]. For example, in the sentence “Mohammed and Ahmed work in IBM Company in Egypt”; “**Mohamed**” and “**Ahmed**” will be identified as Person NE, “**IBM**” will be identified as organization NE, and “**Egypt**” will be identified as location NE.

* Institute of Statistical Studies and Research, Cairo University, Egypt
eng_maryamadel@yahoo.com

The NER task helps the Natural Language Processing (NLP) applications such as Information Retrieval (IR), Question Answering (QA) and Machine Learning (ML) to enhance their performance [Shaalan & Raza 2009].

There are some challenges that face the NER task when it's used with the Arabic language because of the characteristics of Arabic language such as the rich morphology and syntax. Absence the capitalization and the short vowels, the ambiguity, and the lack of the resources are some of these challenges.

The annotated corpora and the lexical are the main linguistic resources for the Arabic NER. Also, there are some the tools that support the Arabic NER such as: "GATE" and "MADAMRA".

There are three main approaches are used in the NER are: (1) Rule Based NER that depends on the linguistic rules of the language, (2) Machine Learning Based NER that depends on the features of the NE classes in a large training corpus, and (3) Hybrid Based NER that combine between the two previous approaches.

It is necessary to evaluate the performance of the NER system. There are some the evaluation metrics that used to evaluate the performance of the NER system. These metrics are part of NLP conferences such as "MUC", "CONLL", and "ACE".

This research aims to give a good background on the Arabic NER for the interested researchers.

The remainder of the research is organized as follows. Section 2 discusses the importance of the Named Entity Recognition Task. Section 3 discusses and analyzes the challenges that face the NER when it is used with Arabic language and the solution for each. Section 4 presents the linguistic resources and the tools that support the Arabic NER. Section 5 explains the main approaches of the Arabic NER and gives an example for each approach. Section 6 reviews the evaluation metrics that are used to evaluate the performance of the Arabic NER. Finally, the conclusion remarks and the future work are presented.

2. The importance of the NER in NLP applications

"Names" play important role in any text for detecting, identifying and extracting them [Mohit, 2014]. So, recognizing the names can improve many applications in NLP. The NLP applications includes: Information Retrieval (IR) systems, Information Extraction (IE), Machine Translation (MT), Question Answering (QA) systems and others.

2.1. Integrated the NER with the information retrieval

The information retrieval (IR) is “the task of identifying and retrieving relevant documents from a set of data according to an input query” [Shaalán, 2014]. According to the studies on the information retrieval, there is a strong relation between the named entity recognition task and the retrieval systems.

NER can be useful for the IR through two possible ways: recognizing the “Named Entities” within the input query and determining the relevant document according to the existence of the recognized “Named Entities” within these documents. For example, the word (“الجزيرة”- Aljazeera) can be recognized as an *organization* Named Entity or a *Location* Named Entity corresponding to the word island. So, the correct recognition leads to retrieving and extracting of the relevant documents.

2.2. Integrated the NER with the Machine Translation

The Machine Translation (MT) is the task of translating a text in a language into another language. Improving the Named Entities translation can help in improving the performance of the MT system [Kaddoura, 2010]. The translation of some Arabic person names to Latin languages faces the ambiguity problem; because the Arabic person name can be found as regular words in the language that isn't a named entity type. For example, the word “سعيد” (Saeed) can be used in Arabic text as a noun (non-NEs) that means “Happy”, and also as a Person name (Person NEs). So, the translation of the following phrase “محمد سعيد” can be “Mohamed Saeed” or “Mohamed Happy”.

2.3. Integrated the NER with the Question Answering

The Question Answering (QA) task is related to Information Retrieval field where the questions are taken as input and the QA system returns brief answers. For improving the retrieved data in the QA system, the NER task is used; that is through identifying the relevant documents and then extracting the correct answers from candidate passages. For example, the words “الشرق الأوسط” may be classified as *Organization* NE or *Location* NE according to the context.

3. Challenges of the NER in Arabic Language

There are several challenges/problems face the NER task when using it in Arabic Language; this section describes these challenges and studies the possible solution for each.

3.1. Absence of Capitalization Problem

Most Named Entities in the Latin languages like English begin with capital letters such as proper names e.g. (“Ahmed”, “Mohammed”) and abbreviations e.g. (“ACM”, “IBM”). But this feature doesn’t exist in the Arabic language because the Arabic language can’t support the capitalization. Absence of this feature effects on the Named Entity recognition task [Shaalan, 2014].

The solution of this problem

The dictionary lookup method wouldn’t be a suitable solution to face this problem, because some of the words can be used as proper nouns and they also can be used as non-proper noun in the text. For example, the Arabic proper name “عين” can be used in the sentence with different meanings as in Table 1.


Table 1 Example of the absence of the capitalization problem

Word meaning	Word Category	Sentence
Ain	Proper-Noun	عين جالوت
wellspring	Noun	عين الماء
eye	Noun	عين الانسان
delimitate/be delimitate	Verb/passive Verb	عين وزيراً للخارجية

[The source (Saad & Ashour 2010)]

So, the solution of this problem is analyzing the context surrounding the Named Entity.

3.2. Absence of Short Vowels

Arabic text contains the diacritics  that “affect the phonetic representation and give different meaning to the same lexical form” [Shaalan, 2014]. But nowadays, Modern Standard Arabic (MSA) is written literally, without considering the addition of diacritics [Alotaibi, 2015]. Absence of the diacritics in the Arabic words leads to the ambiguity because different diacritics represent different meanings. For example, the word “قطر” without any diacritics can be recognized as *country* “location NEs” or as *radius* “Measure NEs”.

The solution of this problem

Consider the context of the word in order to predict the correct meaning.

3.3. Complicated morphology

Finding many different patterns for the one Arabic word is from the characteristics of the Arabic language where each word can consist of one or more prefixes, a stem or root, and one or more suffixes in different combinations, that leads to complicated morphology.

Table 2 Arabic Patterns and Roots

Affixes in Arabic	Examples
Prefixes of length 3	ولل ، وال ، كال ، بال
Length 2 prefixes	ال ، لل
Length 1 prefixes	ل ، ب ، ف ، س ، و ، ي ، ت ، ن ، ا
Length 3 suffixes	تمل ، همل ، تان ، تين ، كمل
Length 2 suffixes	ون ، ات ، ان ، ين ، تن ، كم ، هن ، نا ، يا ، ها ، تم ، كن ، ني ، وا ، ما ، هم
Length 1 suffixes	ة ، ه ، ي ، ك ، ت ، ا ، ن

[The source (Saad & Ashour 2010)]

The solution of this problem

There are two possible ways to resolve this problem

1. By deleting all the affixes and keeping only the root of the Arabic word. For example, the analysis of "وبالقاهرة" word will result "القاهرة" as a location name after deleting the "و" and the "ب". Although this way is faster, it leads to a loss of valuable information from the affixes [Benajiba et al., 2007].
2. By separating the suffixes with spaces. For example, the word "وعاصمتها" can be segmented into three parts "و" "عاصمة" "ها". This way is more accurate, because it keeps all affixes and then keeps the meaning.

3.4. Ambiguity Problem

The problem of ambiguity can be between two or more Named Entities because the Named Entity phrases can be formed by different POS such as 'common nouns', 'adjectives' or more complex phrases of more than one token. For example, In the sentence "أحمد أباد فاز بالجائزة": the phrase "أحمد أباد" can be recognized as "Person NE" or as a "location NE¹". Table 3 presented more examples for this problem.

¹ "Ahmed Abad" is the largest city and former capital of Gujarat, which is a state in India.

Table 3 Ambiguous Examples

Ambiguous example	English translation	Incorrect	Correct
1.6985 فرنك سويسري	1.6985 Swiss francs	Person	Price
15 رمضان الكريم 2005	15th of Ramadan Al karim 2005	Person	Date
جاسم المتحدة للعقارات والصيانة العامة	Jassim united for real estate and general maintenance	Person	Company
1.5 بليون دولار سنغافورة	1.5 billion Singapore dollars	Location	Price
شركة أرامكو السعودية	Saudi Aramco	Location	Company
راشيل فيكتوريا كيون	Racheal Victoria Queen	Location	Person
اليزابيث الثانية في مساء	In the evening Elizabeth II	Time	Person
نقطة تحول في سبتمبر سنة 1954 تقدم مارتن	... a turning point in September 1954 Martin presented ...	Measurement	Date

[The source (Shaalán & Raza, 2009)]

The solution of this problem

Rule based approach can be used for resolving this problem. (Shaalán & Raza, 2009) used the heuristic rules for resolving this problem by “preferring one Named Entity type over the other”.

3.5. Transliteration Problem

An NE can be transliterated in many ways. The lack of standardization leads to many shapes of the same word that are spelled differently but with the same meaning. Another reason for this is that “Arabic has more speech sounds than Western European languages, which can ambiguously lead to an NE having more variants” [Shaalán, 2014]. For example, the city of “Washington” could be expressed using four forms such (وشنطن، واشنطن، واشنغطن، واشنجطن)

The solution of this problem

Make a standard form/ a canonical form and normalize each occurrence of the variant to this form; this requires a mechanism (such as string distance calculation) for matching between a name variant and its normalized representation.

3.6. Lack of Resources

There are large annotated corpora as well as Arabic lexicons that can be used for implementing and testing the performance of an Arabic NER system; but most of these available Arabic NER resources are expensive and have limited capacity.

The solution of this problem

Researchers depend on their own corpora, which require human annotation and verification. Few of these corpora have been made freely

and publicly available for research purposes as in [Benajiba et al., 2007] and [Mohit, 2014]; whereas others are available but under license agreements.

4. Arabic Linguistic Resources and Tools supporting Arabic NER task

As it is shown in the previous section, the lack of digital linguistic resources and the tools that support the NER represents a challenge for the NER task especially when it is used with the Arabic language.

This section presents the available linguistic resources and the tools that support the Arabic NER. It detects each of them is open access.

4.1. Arabic Linguistic Resources

There are two types of linguistic resources that are commonly used in NER: (1) Corpora and (2) Lexical resources.

4.1.1. Corpora

The datasets or corpora are used to evaluate and compare the systems. For the NER task, we need “large annotated corpus” where every NE has a type assigned to it.

Some NER corpora are available under paid license agreements, for example, “ACE²”. And others are freely available such as “ANERcorp³” that is a Corpus of more than 150,000 words annotated for the NER task [Benajiba et al., 2007].

Fig. 1 shows a sample of annotated corpora where the words of corpora are collected from different resources and each word are classified into its type: “Loc” means Location NE, “PER” means person NE, “ORG” means organization NE, and “O” means other NE. The letters “B, I,L” means the site of the word in the sentence “Begin”, “In the middle” and “Last”.

² <https://www ldc.upenn.edu/collaborations/past-projects/ace/annotation-tasks-and-specifications>

³ <http://www1.ccls.columbia.edu/~ybenajiba/downloads.html>

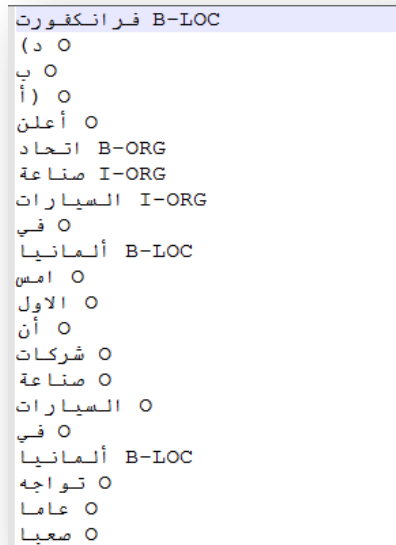


Fig. 1 Sample of annotated Corpora

4.1.2. Lexical Resources (gazetteer)

Gazetteer is another lexical resource which is a collection of predefined lists of typed entities. It can be called as dictionary or wordlist. The contents of a gazetteer should be consistent and belong to only one type of NE. For example, a location gazetteer consists of names of countries, cities, states, and so on [Shalan & Raza, 2009].

When the researchers found these resources aren't available freely; they built their own gazetteers from different resources such as the Web and from organizations. For example, (Benajiba et al., 2007) built "ANERGazet" that is a collection of 3 Gazetteers:

- a) **Locations:** a Gazetteer containing names of countries, cities, states, etc.
- b) **People:** a Gazetteer containing names of people recollected manually from different Arabic websites
- c) **Organizations:** a Gazetteer containing names of Organizations like companies, football teams, etc.

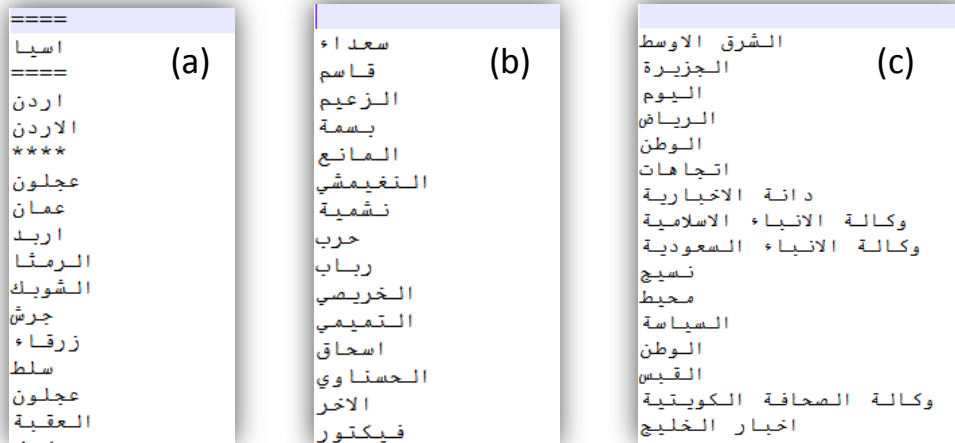


Fig. 2 Sample of the “ANERGazet”
 (a) Loc_Gazetteer, (b) Person_Gazetteer and (c) ORG_Gazetteer

Table 4 presents a summary of the important Arabic NER resources that are available to be used.

Table 4 Summary of Arabic NER Resources

Resource name	Resource Type	Availability	Link
“ACE”	Corpora	under paid license agreements	https://www ldc.upenn.edu/collaborations/past-projects/ace/annotation-tasks-and-specifications
“ANERcorp”	Corpora	Free	http://www1.ccls.columbia.edu/~ybenajiba/downloads.html
“AQMAR”	Corpora	Free	http://www.cs.cmu.edu/~ark/ArabicNER/
“Fine-grained”	Corpora	Free	https://sourceforge.net/projects/arabic-named-entity-corpora/
“WIKIFANE_Gazetteer”	Lexical	Free	https://sourceforge.net/projects/arabic-named-entity-gazetteer/?source=directory
“CJK”	Lexical	under paid license agreements	www.cjk.org/cjk/arabic/arabsam.htm
“ANERGazet”	Lexical	Free	http://www1.ccls.columbia.edu/~ybenajiba/downloads.html

4.2. Tools supporting the Arabic NER task

In Arabic language, there is a lack in the NER tools which have more importance for the NLP systems [Kaddoura, 2010].

In this section, some NER tools that have been used in the Arabic NER literature are presented. The tools can be classified into two categories according to their functions: (1) Integrated Development Environments tools and (2) Basic Preprocessing Tools for Arabic.

Table 5 presents a summary of these tools with link for the official site for each.

4.2.1. Integrated Development Environments tools

- **GATE (The General Architecture for Text Engineering):**
 - It is a freely available.
 - It supports these languages (English, French, Italian, German, Arabic, Chinese, Hindi, Romanian, and Cebuano).
 - It can be used to extract basic Arabic entities, such as (date, name, location, organization, and so on).

- **NooJ:**
 - It is a freely available.
 - It supports constructing, testing, and maintaining large lexical resources, and applying morphological analysis for Arabic processing.
 - It can recognize all Unicode encodings.

- **LingPipe:**
 - The free version has limited production capabilities and in order to obtain full production abilities, it must to be upgraded.
 - It is a toolkit for text engineering and processing.
 - It supports POS tagging, spelling correction and NE recognition.

4.2.2. Basic Preprocessing Tools for Arabic

In this section, the Arabic morphological pre-processing tools are presented that are used in the Arabic NER literature, including BAMA, MADA, AMIRA and MADAMIRA toolkit.

- **Buckwalter Arabic Morphological Analyser (BAMA)**
 - It contains over 80,000 words, 38,600 lemmas.
 - It contains three dictionaries (Prefix, Stem, Suffix) and three compatibility tables (Prefix-Stem, StemSuffix, Prefix-Suffix).

- **Morphology Analysis and Disambiguation for Arabic (MADA)**
 - It is a development of BAMA.
 - The TOKAN component allows the user to specify any tokenization scheme that can be generated from disambiguated analyses.
 - The MADA+TOKAN package provides one solution to all of the basic problems in Arabic NLP, including *tokenization*, *diacritization*, stemming, and lemmatization.

- **AMIRA**
 - This is a set of tools including a tokenizer, POS tagger and Base Phrase Chunker.
 - It is used for different NLP applications because of its speed and high performance.

- **MADAMIRA**
 - It combines aspects of two previously used systems for NLP; MADA and AMIRA.
 - It includes several tasks that are useful for NLP processes such as POS tagging, tokenized forms of words, diacritization, lemma stemming, base phrases, and NER.

Table 6 Summary of Arabic NER Tools

Tool	Availability	Link
“GATE”	 Free	https://gate.ac.uk/
“NooJ”	 Free	http://www.nooj4nlp.net/
“LingPipe”	Free for limited capabilities	http://alias-i.com/lingpipe/
“BAMA”	 under paid license agreements	https://catalog ldc.upenn.edu/LDC2004L02
“MADA + TOKAN”	 Free	https://lists.cs.columbia.edu/pipermail/mada-users/
“AMIRA”	A demo of the system is available	http://nlp.ldeo.columbia.edu/amira/
“MADAMIRA”	 Free	http://www1.cs.columbia.edu/~rambow/software-downloads/MADA_Distribution.html

5. Main Approaches for Arabic NER

According to the research works on the Arabic NER, it is found many approaches for recognizing named entities from text. These approaches have been divided into three categories [Mohit, 2014] as following: (1) Rule Based NER, (2) Machine Learning based NER, and (3) Hybrid NER.

5.1. Rule-Based NER

The Rule-Based is from the early approaches to the NER. It depends on linguistic rules (grammars). It is also called “Linguistic-based” approach [Shaalán, 2014]. There are three main components are used in the Rule-Based systems, as following:

- (1) A set of Rules for the named entity extraction task (grammatical rules).
- (2) Gazetteers/ Dictionaries that contains different types of named entity classes
- (3) Extraction engine that applies the rules to the text.

5.1.1. The advantage and disadvantage of this approach

The advantage of the Rule-based approach is that it needs to high linguistic knowledge to build it.

The problem of this approach is when needing to make any updates or maintenance on the system but the linguists with the required knowledge and background are not available. It requires expensive manual effort and it is time-consuming [Shaalán, 2014].

5.1.2. Example on using this approach

(Shaalán & Raza, 2009) used the Rule-Base approach for developing NER system for Arabic (called NERA system, see Fig. 3). They depended on their own resources, corpora and gazetteers, to train and test the system.

For example, when applying the proposed system in [Shaalán & Raza, 2009] to recognize the NEs in a text, a phrase such “أحمد آباد” can be recognized as a person name or a location. So, it is necessary to create a filter rule to return one correct result. The authors put the following filter rule:

“If a possible match M1 for a location entity reported by the location extractor intersects with a match M2 of a person entity that is also reported by the person extractor, then the match as a location name will be discarded”

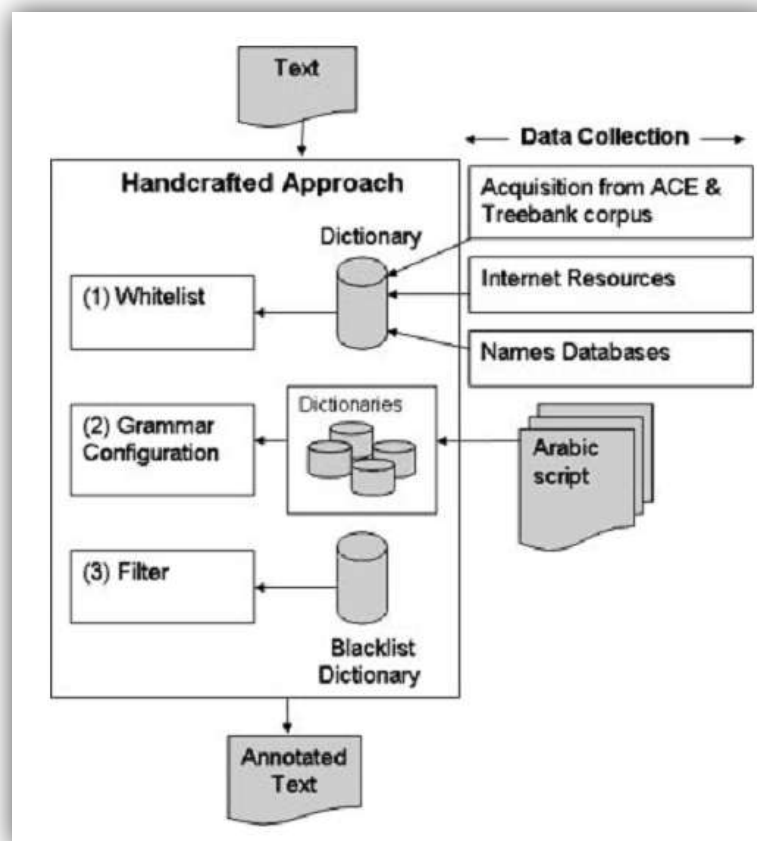


Fig. 3 Architecture of NERA system
(Source: [Shaalán & Raza, 2009])

So, in case of an intersection, “أحمد آباد” will be recognized as person name. In my opinion, there are some the drawbacks in “NERA” system when using this rule; where the phrase “أحمد آباد” can be a location but in this case the system can’t retrieve the correct answer!. So, we suggest using the probabilities and the context surround the NEs for improving this linguistic rule and the similar rules.

Also, there are some the ways that can improve the performance and enhance the accuracy such as using large corpora and large dictionaries, and using Arabic text with error-free spelling.

5.2. Machine Learning (ML)-Based NER

The idea in ML-based approach is to study the features of the Named Entity classes in a large training corpus. There are two main components are used in the ML-Based systems, as following: (1) Large annotated corpora and (2) a probabilistic representation of the training data (A statistical model). Decision Trees, Support Vector Machines (SVM), Maximum Entropy (ME), Conditional Random Fields (CRF) and Hidden Markov Models (HMM) are examples of the statistical models.

5.2.1. The advantage and disadvantage of this approach

The advantage of this approach is that the use of ML approaches reduces the human effort needed for building a set of rules and gazeteers. In some cases, the ML-based approach is more flexible than the rule based approach.

The disadvantage of the machine learning approach is the needing to large corpora of annotated text. And this problem appears highly in Arabic NER because of the lack of linguistic resources.

5.2.2. Example on using this approach

(Benajiba, et al., 2007) built a NER system called “NERsys” for Arabic texts based-on n- grams and maximum entropy. The authors depend on their own testing and training corpora and gazeteers (ANERcorp & ANERgazet)⁴.

First, we would know how to use maximum entropy

Maximum entropy (ME) is “a general technique for estimating probability distributions from data” [Benajiba, et al., 2007]. The Maximum entropy calculates the best probability distribution according to the defined information. The following example will explain “how the ME classifier performs”:

*“Sudan’s **Darfur** region remains the most pressing humanitarian problem in the world”*

If we want to classify the word “**Darfur**” as one of these classes (person, location, organization, or other)

- If we assumed that we don’t have any information on the word “**Darfur**” then the distribution of the probability on the four classes

⁴ Available at: <http://www1.ccls.columbia.edu/~ybenajiba/downloads.html>

will be the same. $P(\text{person})=P(\text{Location})=P(\text{Organization})=P(\text{other}) = \frac{1}{4}= 0.25$.

- But, if we assumed that we have information for “Darfur” that it a word that begins with “Capital letter” and isn’t “in the start of the sentence”.
 - By using this information we will guess that “Darfur” is a proper name (one of those: Person, Location, or Organization) and it isn’t any other NE.
 - So, distribution of the probability will be different. The most probability will be assigned to the ‘Person’, ‘Location’, and ‘Organization’. $P(\text{person})=P(\text{Location})=P(\text{Organization})= \frac{1}{3} = 0.3$.

In (Benajiba, et al., 2007), the results showed that using the maximum entropy can enhance the performance of the Arabic NER task without using any POS-tag information or text segmentation.

Table 7 shows a Comparison between Rule-Based and ML-Based

Table 7 Comparison between Rule-Based and ML-Based

	Rule-Based	ML-Based
Language Specialists	It depends on linguistic rules “hand-constructed rules” that require language specialists	It doesn’t need language specialists
The Required Training Data	Small amount of training data	Large amount of annotated training data (very large corpora)
Time Consuming	Very time consuming	Automated
Changes	Some changes may be hard	Some changes require re-annotation of the entire training corpus
Quality	High quality	Less quality

5.3. Hybrid NER

The idea in the hybrid NER systems is combining the rule-based approach with the ML-based approach to overcome on the problems in both methods. Using the hybrid Arabic NER can improve the Arabic NER task.



5.3.1. The advantage and disadvantage of this approach

In some cases, the depending only on rule-based features doesn't improve the performance; and the depending only on machine learning based features doesn't improve performance. But when integrating the features of rule-based with Machine learning classifiers, in this case the performance can be improved.

5.3.2. Example on using this approach

(Abdallah, et al. 2012) proposed a simple method for integrating the Machine learning-based approach with rule-based approach for Arabic NER as it is shown in Fig. 4.

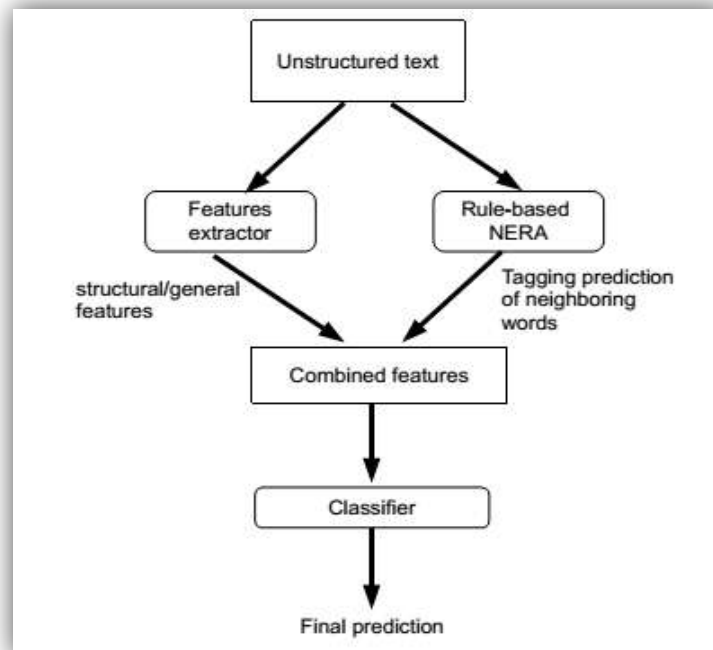


Fig. 4 Block diagram illustrating the integration between the Rule-Based and ML-Based (Source: [Abdallah, et al. 2012])

The authors focused on only three named entities (person name, location, and organization).

For training and testing the proposed method, they used two annotated corpora:

- (1) The ACE 2003 Multilingual Training Set⁵.
- (2) ANERcorp Corpus⁶.

Steps to build their rule-based system:

1. Performing the recognition based on a dictionary lookup that containing lists of known named entities.
2. Using a parser, based on a set of grammar rules (represented as regular expressions).

Steps to build the proposed integrated approach:

1. Using the Stanford POS Tagger⁷ to compute some of the general features such as word category and affixation that are defined as machine learning features.
2. Complementing the rule-based features with the other extracted features
3. Feeding all combining features to a decision tree classifier.

In (Abdallah, et al. 2012), the results of proved that the proposed hybrid approach is better than the pure rule-based system or the pure machine-learning classifier.

6. Evaluating the performance of the NER systems

The aim of the evaluation is to ensure if the NER system can enhance the performance or no and with which degree according to the used datasets. NER systems are evaluated by “running them on human-labeled data and comparing their results against this gold-standard” [Mohit, 2014]. If there are standard evaluation corpora, it will be easy to compare between the existed NER systems. Some researchers used the annotated datasets where every NE has a type assigned to it such as (“ANERcorp” and “ANERGazet”), and the evaluation measurements (precision, recall,

⁵ Available to BUID under License from <https://www ldc.upenn.edu/collaborations/past-projects/ace/annotation-tasks-and-specifications>

⁶ Available for download from <http://users.dsic.upv.es/ybenajiba/>

⁷ available at <http://nlp.stanford.edu/software/stanford-postagger-2010-05-26.tgz>

f-measure and accuracy) for evaluating their NER systems [Benajiba et al., 2007].

In this section, we presented the evaluation measurements and the evaluation metrics that are used for evaluating the performance of the Arabic NER systems.

6.1. Evaluation Measurements

The evaluation measurements such as “Precision, Recall, F-measure, and Accuracy” are the most used of the evaluation techniques [Al-Jumaily et al., 2012]. The equations of each measure will be defined in terms as in the following table:

Table 8 Terms in the evaluation measures equations

Term	Stands for	Definition
TP	((True-Positives))	“It counts the tokens correctly assigned to this category”
FP	((False-Positives))	“It counts the tokens incorrectly tagged to this category”
FN	((False-Negatives))	“It counts the tokens incorrectly rejected from this category”
TN	((True-Negatives))	“It counts the tokens correctly rejected from this category”

Precision

The precision is “the ratio of the retrieved tokens which are relevant in the corpus, i.e., it evaluates the exactness of the system” [Jumaily et al., 2012].

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

Recall

Recall is “the ratio of the retrieved relevant tokens. It measures the ability of the system to retrieve a complete set of the relevant tokens from a corpus” [Jumaily et al., 2012].

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

F-measure

F-measure evaluates the effectiveness of the system.

$$F_{measure} = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \quad (3)$$

6.2. Evaluation Metrics

There are three main NER scoring metrics used as part of NLP conferences: (1) Message Understanding Conference “MUC”, (2) Computational Natural Language Learning Conference “CoNLL”, and (3) Automatic Content Extraction “ACE”.

6.2.1. MUC Evaluations⁸

“MUC” is an evaluation metrics where a system is scored on two axes:

1. Its ability to find the correct type of the NE (TYPE).
2. Its ability to find the boundaries text surrounds the NE (TEXT).

The advantage of this method is taking into account all possible types of errors. But the ambiguity in the boundaries can lead to a problem.

6.2.2. CoNLL Evaluations⁹

“CoNLL” provides an exact match evaluation, where the entity is considered correct if it exactly matches the same type and text. This method is simple in calculating and analyzing results [Shalan, 2014]. In this method, the Precision, recall and F-measures are used to calculate performance.

6.2.3. ACE Evaluations¹⁰

“ACE” evaluation deals with several kinds of errors into an integrated scoring mechanism where each type of error and each type of entity has different weight. Compared by “MUC” and “CoNLL” methods, “ACE” is more complex. Because of the complexity of this evaluation method, most studies used the “CoNLL” method for the evaluation [Alotabi, 2015].

According to the literature, we concluded that the CoNLL method can be a standard method for evaluating the Arabic NER.

⁸ http://www.itl.nist.gov/iaui/894.02/related_projects/muc/muc_sw/muc_sw_manual.html

⁹ <http://universaldependencies.org/conll17/evaluation.html>

¹⁰ <https://www ldc.upenn.edu/collaborations/past-projects/ace/annotation-tasks-and-specifications>

7. Conclusion and Future Work

The Named Entity Recognition task has an importance in each field of NLP fields. In this research, we studied the importance of the NER task in some of these fields. Also the challenges that faced the NER task when using it with the Arabic language and the solutions that are used in the literature for each challenge were studied and analyzed. We founded that “analyzing the context surrounding the NEs” is the solution for the most of these challenges.

The studies on the Arabic NER used some the linguistic resources such as Corpora and Gazetteers and they also used some tools that support the Arabic language such as BAMA and MADAMIRA. Some of these resources and tools are freely available and the others are available under paid license agreements.

There are three approaches of the NER: (1) Rule-Based approach, (2) Machine Learning approach and (3) Hybrid approach. According to the literature the Hybrid-based can be the best approach in some cases.

NER systems are compared based on standard evaluations such as “MUC”, “CoNLL”, and “ACE”; and standard evaluation measurements such as “Precision”, “Recall”, and “F-measure”. According to the literature, we concluded that the “CoNLL” Matrix was the most used because of its simplicity in the calculating and the analyzing the results. To the best of our knowledge, there isn’t research that evaluates and compares the performance of the available Arabic NER systems.

In the future, we would to make a comparative study between using many statistical models (SVM, HMM, Maximum Entropy, CRF, etc.); and to detect the role for each for improving the Arabic NER task.

References

- Abdallah, S., Shaalan, K., & Shoaib, M. (2012). Integrating rule-based system with classification for Arabic named entity recognition. In Proceedings of the 13th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), Springer- Verlag, Berlin Heidelberg, pages 311-322.
- Al-Jumaily, H., Martínez, P., Martínez-Fernández, J. L., & Van der Goot, E. (2012). A real time Named Entity Recognition system for Arabic text mining. Language Resources and Evaluation, 46(4), 543-563.
- Alotaibi, F. (2015). Fine-grained Arabic named entity recognition (Doctoral dissertation, University of Birmingham).
- Benajiba, Y., Rosso, P., & Benedíruiz, J. M. (2007, February). Anersys: An arabic named entity recognition system based on maximum entropy. In International

Conference on Intelligent Text Processing and Computational Linguistics (pp. 143-153). Springer Berlin Heidelberg.

Dandashi, A., Al Jaam, J., & Foufou, S. (2016). Arabic Named Entity Recognition— A Survey and Analysis. In *Intelligent Interactive Multimedia Systems and Services 2016* (pp. 83-96). Springer International Publishing.

Kaddoura, H. (2010). ALNER: Arabic location named entities (Doctoral dissertation, The British University in Dubai (BUiD)).

Mohit, B. (2014). Named entity recognition. In *Natural Language Processing of Semitic Languages* (pp. 221-245). Springer Berlin Heidelberg

Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Lingvisticae Investigationes*, 30(1), 3-26.

Saad, M. K., & Ashour, W. (2010). Arabic morphological tools for text mining. In: *EEECS10 the 6th International Symposium on Electrical and Electronics Engineering and Computer Science*. European University of Lefke, Cyprus, pp. 112–117.

Shalan, K. & Raza, H. (2009). NERA: Named entity recognition for Arabic. *Journal of the American Society for Information Science and Technology*, 60(8), 1652-1663.

Shalan, K. (2014). A survey of Arabic named entity recognition and classification. *Computational Linguistics*, 40(2), 469-510.

Image Encryption Scheme with Hashed Biometric Key

Ali M. Meligy¹, Hossam A. Diab² and Marwa S. Eldanaf³

ABSTRACT

Communities have always interested for the security of information over the eras. This has produce to the improvement of several information security methods, prominent among them is cryptography. In this paper we suggest a technique for image encryption that has fast execution speed and high level safety. The design tools of our algorithm are depend on a chaos-based feedback cryptographic scheme using the tent map and an external biometric secret key of 160-bit using hash function (SHA-1). Accordingly, the created external key is utilized to extract the initial seeds of the used chaotic mapping. We will see soon, the experiences show the efficiency of the suggested scheme in addition to its vulnerability to secret key changes and its resistance to various kinds of threats.

Keywords

Stream Cipher, Chaos Cryptography, Hash Function, SHA-1 Hash Algorithm, Tent Map, Logistic Map, Arnold Cat Map, Statistical Tests, Statistical analysis, Security analysis.

1. INTRODUCTION

Historically, the security of data was accomplished by a collection of physical security methods and trust. Stamps and subscriptions were confirmed of substantiation of legitimacy, this is appeared in the royal seals and signets by last emperors. Confidentially was accomplished by closing up documents and lockers. Integrity of information was fundamentally based on trust. Devoted manpower such the royal postal service was entrusted with the task of transferring information. The degree of information security was based on the trouble faced in compromising these security methods and breaking confidence. In the information age, though the demand for information security yet stays the same. Many of these physical security techniques are not only inadequate, but many are impracticable. The networked age has raised the danger of information safety probability of accumulation assets. Vulnerable data establishing on unlock and un-trusted networks (e.g. the Internet) can be simply accessed, duplicate or altered [3]. In impose to counter the security threat in the new age; a distinguished technique that has improved is encryption. The subject of encryption is called cryptography.

A hash function is a utility that takes a comparatively arbitrary quantity of input and makes an output of fixed size [6]. This concept feature makes them important in data structure, checksum algorithms for error detection, digital signature in information security etc. Cryptographic hash functions have another merit that is beyond hash functions - it is very rough to discover two distinct inputs that make the same output. This property gives a high level of certitude though not ultimate that several input values would product different output signature in most of the cases. This make cryptographic hash functions the hash functions that are used in information security related applications. SHA-1 (Secure Hash Algorithm) is one of hashing algorithms; it is used to generate the hashing value. It creates the hash value of 160 bits (20 bytes). It has the 80 number of rounds. The user which has the hash value can adjust the data. The hashing algorithm supply authenticity and integrity. If any user alters the data then the hash value will be varied [1].

¹Prof.Dr of Computer science Faculty of science–menoufiya University–Mathematics department

meligyali@hotmail.com

²Assistant Professor of Computer Science Faculty of science– Menoufiya University– Mathematics department

hossamdiab_86@yahoo.com

³Researcher in Academy of Scientific Research and Technology

marwa277.saeed@gmail.com

According to [7], Cryptographic Hash functions are one of the most significant tools in the field of cryptography and are used to realize a number of security goals. In [8] the essential characteristics of the Hash function that allow them to withstand to a satisfactory level all known cryptanalytic attack was highlighted.

In [2] mentioned to collision resistance property as collision freeness or strong collision resistance, second pre-image resistance as weak collision resistance and preimage resistance as one-wayness.

In [4] categorized collision resistance as the strongest property of all three, hardest to satisfy and easiest to breach, and breaking it is the goal of most attacks on hash functions.

In [5] the notion of the hash function security was expanded. In this extension they realized seven several security notions, three construct on pre-image resistance, three based on second pre-image resistance and one on collision resistance. Their work is based on public concept of hash function family that is a limited set of hash functions with common domain and range.

Similarly some chaos – based cryptosystems are used to solve the privacy and security problems of biometric templates. The secret keys are randomly generated and each session has different secret keys. Thus biometric templates are encrypted by means of chaotic cryptographic scheme which makes them more difficult to decipher under attacks [11]. Most properties are related to some requirements such as mixing and diffusion in the sense of cryptography. Therefore, chaotic cryptosystems have more useful and practical applications.

One of the simplest chaos functions that have been studied recently for cryptography applications is the logistic map. The logistic map function is expressed as:

$$x_{n+1} = r x_n (1 - x_n) \quad (1)$$

Where x_n takes values in the interval $[0, 1]$ and $r \in [0, 4]$. It is one of the simplest models that present chaotic behavior [12].

Also, the proposed cipher utilizes the chaotic Tent map which can be depicted as follows:

$$f_{\mu}(x_{n+1}) = \begin{cases} \mu x_n & \text{for } x_n < 0.5 \\ \mu (1 - x_n) & \text{for } 0.5 \leq x_n \end{cases} \quad (2)$$

For the chaotic Tent map [12], the control parameter $\mu = 2$ yields a chaotic sequence $x_n \in [0, 1]$ with a random behavior.

Also, popularized Arnold map in 1960, Russian mathematician Vladimir Arnold used the most general two-dimensional chaotic map for an image [9]; the name was Arnold Cat Map. If a matrix $N * N$, pixel with coordinates (x, y) , we wrote to Arnold will be:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & p \\ q & pq + 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \pmod{N} \quad p, q=1 \quad (3)$$

Where q and p are real numbers, the (x', y') is the new position of original pixel (x, y) . Since there only exists a linear transformation and mod function, it is very efficient to mix the pixel positions using the Arnold cat map. After several iterations, the correlation among the adjacent pixels can be muddled totally.

In this paper, present a technique for image encryption that has fast execution speed and high level safety. The design tools of our scheme are depend on a chaos-based feedback cryptographic scheme using the tent map and an external biometric secret key of 160-bit using hash function SHA-1. Accordingly, the generated external key is employed to derive the initial seeds of the applied chaotic mapping. Moreover, the

pixels are masked based on an iterative module which exploits a data-dependent feedback mechanism to mix the current cipher conditions with the previously masked pixels to get the encryption results.

-2-

The paper is organized as follows: Section 2 presents the structure of the proposed image encryption scheme. Experimental tests and numerical computations to emphasize the encryption quality of the presented scheme are suggested in Section 3. Section 4 debates the diverse security analyses of the presented scheme including statistical analysis and sensitivity analysis related to key and plaintext changes. Finally, Section 5 drafts the conclusions.

2. PROPOSED IMAGE ENCRYPTION SCHEME

The suggested cipher is a symmetric key stream cryptography algorithm in which three essential functions (the key expansion, encryption and decryption modules) are utilized by the dispatcher and recipient to acquire the encrypted and decrypted image, severally. First, a secret biometric image is utilized by the dispatcher/recipient to create the secret key by applying hash function (SHA-1). The encryption operations are utilized to the plain image to obtain the cipher image. The structure of the suggested cipher depends on a feedback mechanism in which the encryption of each pixel is made dependent on the encryption merits of the preceding cipher pixel, which in turn, makes the cryptosystem powerful against any kind of threats.

The next subsections demonstrate the three phases of the suggested scheme.

2.1 Key Expansion by Using SHA-1

In view of the fundamental wants of cryptology, the cipher text should be strongly related to the secret key and the security of the encryption algorithm only relies on obscuring this key. Moreover, the cipher should be strictly sensitive to little changes of the secret key. Thus, the strategy of randomly generating the key ensures these requirements. The proposed mechanism for key scheduling utilizes from a selected biometric image to obtain the desired key by using SHA-1.

The steps for key generation can be represented as follows:

Step 1: Input a biometric image Q .

Step 2: Transform biometric image Q to be square matrix $m \times m$ then Exchanged biometric image Q by using Arnold cat map Eq. 3, the existing matrix is titled by A .

Step 3: Produce matrix B with same length of A by utilizing Logistic map which mentioned in Eq. 1, where $X_0 = \left(\frac{\text{mean}(A)}{255} \right) \text{ mod } 1$.

Step 4: However, the periodicity of Arnold cat map should degrade the security, because the possible threats may iterate the Arnold cat map incessantly to reemerge the original image. As a remedy, we modify the pixel values next to increase efficient key by BitXoring A and B .

$$C = A \oplus B \quad (4)$$

Step 5: Derive 160 bit secret key by utilizing SHA-1 which take C in step 3 as input.

$$\text{SecretKey} = \text{SHA-1}(C) \quad (5)$$

The test suite from NIST [10] was chosen to check the randomness of the sequence (secret key) produced by the suggested key expansion mechanism. This suite consists of a set of tests. Each test is independently applied to an n bits sequence (the same sequence in each test) to get a P-value. Particularly, the statistical

package consists of 16 tests [10]. These tests are performed on our proposed key generator and the obtained results are brief in Table 1. The estimations values confirm that the proposed generator can pass many of the underlying statistical tests and the basic requirements for the uniform distribution are met. Thus, the generated key stream is uniformly distributed and cannot be predicted by an enemy.

-3-

Table 1: Results of Statistical Tests NIST

No.	Statistical tests	P_Value	No.	Statistical tests	P_Value
1	Frequency (Monobit)	0.7518	9	Maurer's Universal, (L=7,Q=1280)	0
2	Block Frequency (M =20)	0.4224	10	Lempel-Ziv Compression	0
3	Runs Test	0.1522	11	Serial	0
4	Longest Runs of Ones(M=5)	0.0167	12	Approximate Entropy	0
5	Binary Matrix Rank(M = 7)	1.1872	13	Linear Complexity (M = 22)	0.1736
6	Spectral DFT	1.000	14	Cumulative Sums (Backward) Zero & One	0.7200& 1.000
7	Non-overlapping, M=12,B= [1 0 1]	0.0153	15	Random Excursions	0.6126
8	Overlapping, M =22,B =[1 1 0 1 1]	0.7856	16	Random Excursions Variant	0.4319

2.2 Proposed Encryption Scheme

The image encryption process utilizes an external biometric secret key of 160-bit long from SHA-1. Further, the secret key is divided into blocks of 8-bit. Now, the proposed scheme generates the initial seed of the employed chaotic tent map, y_0 and the initial cipher pixel C_0 from the extracted external biometric key K . Presume that the external secret key is represented as follows:

$$K = K_1 K_2 \dots K_{20} \quad (6)$$

Where K_i performs a block of 8-bit of the overall 160-bit biometric key K .

To compute y_0 and C_0 , the following two steps are carried out:

$$y_0 = \left(\sum_{i=1}^{length(K)} K_i + \frac{K_{20}}{256} \right) \bmod 1 \quad (7)$$

$$C_0 = \left(\left(\bigoplus_{i=1}^{length(K)} K_i + K_1/256 \right) * 10^4 \right) \bmod 256 \quad (8)$$

The suggested encryption handles the plain image as a stream of pixels, each pixel is represented by 8-bit, and encrypts the input image pixel by pixel according to the following steps:

Step 1: Convert the 2D plain image P into 1D vector by reading the pixels from top left to bottom right sides. The obtained plain image vector and its corresponding cipher vector are denoted by P and C , respectively.

$$P = P_1 P_2 \dots P_m \quad (9)$$

$$C = C_1 C_2 \dots C_m \quad (10)$$

Step 2: Encrypt the current pixel P_i to obtain its corresponding cipher pixel C_i according to

$$C_i = [P_i \oplus \{([K_i + T(y_i)] \bmod 256) \oplus C_{i-1}\}] \ggg T_i \quad (11)$$

Where y_i denotes the current input for the Tent map T depicted in Eq. 2 and can be calculated as follows:

$$T(y_i) = \sum_{i=1}^{\text{length of tent map}} y_{i-1} \quad (12)$$

-4-

If the next value obtained is within the subinterval between (0.2, 0.8), the iteration goes on until a desired number not be subinterval between (0.2, 0.8) then exit from tent map. After encryption of each pixel, we modify the K_i and initial value of the tent map y_0 as follows:

$$K_i = K(T(y_i) \bmod 19) + 1 \quad (13)$$

$$y_{i+1} = (((K(i-1) \bmod 20) + 1) + K((T(y_i) \bmod 19) + 1) / 255) \bmod 1 \quad (14)$$

Step 3: Set $i=i+1$ and apply the step 2 until all pixels are encrypted.

Step 4: Convert C to 2D array to obtain the final encrypted image.

2.3 Proposed Decryption Scheme

Decryption is very simple; the same pad is generated but this time un-merged with the ciphertext to retrieve the plaintext. The decryption module receives an encrypted image (cipherimage) and the same 160-bit biometric secret key is generated and returns the original image (plainimage). The decryption scheme applies the same steps with the replacement of the encryption mapping with the inverse mapping of Eq. 11.

3. EXPERIMENTAL RESULTS

To demonstrate the efficiency of the proposed cipher, several experiments are performed on a set of biometric images downloaded from CASIA (Chinese Academy of science and institute of Automation) database [15]. Also, for the numerical evaluation of the encryption quality, the correlation coefficient (C.C) between the plainimage and cipherimage is estimated. Mathematically, C.C can be expressed according to [13, 14] as follows:

$$C.C = \frac{N \sum_{j=1}^N (x_j \times y_j) - \sum_{j=1}^N x_j \times \sum_{j=1}^N y_j}{\sqrt{(N \sum_{j=1}^N x_j^2 - (\sum_{j=1}^N x_j)^2) \times (N \sum_{j=1}^N y_j^2 - (\sum_{j=1}^N y_j)^2)}} \quad (15)$$

Where x and y denote the grey values of the pixels for the plainimage and the corresponding encryption result.

The encrypted images which are illustrated in Fig. 1 emphasize the feasibility of the proposed scheme. Obviously, the proposed scheme effectively conceals all features of the plainimage which means that the encrypted image is visually indistinguishable. Also, the results are compared with the standard encryption algorithms (AES, RC5, and RC6) in Table 2, where the proposed scheme retains the smallest Correlation Coefficients (C.C).

Table 2: The evaluation of encryption quality

Images	Correlation Coefficient			
	RC5	RC6	AES	Proposed Cipher
Iris001_1_1	0.0118	0.0225	0.0152	- 0.0073
Iris001_2_1	- 0.0258	0.0091	- 0.0884	0.0064
Fingerprint100_L0_0	0.0400	0.0078	-0.0119	0.0052
Fingerprint100_R0_0	0.0386	0.0056	0.0183	0.0031
Palmprint0001_m_L_01	- 0.0284	0.0067	0.0040	- 0.0044
Palmprint0001_m_R_01	- 0.0061	0.0050	0.0031	0.0032

-5-

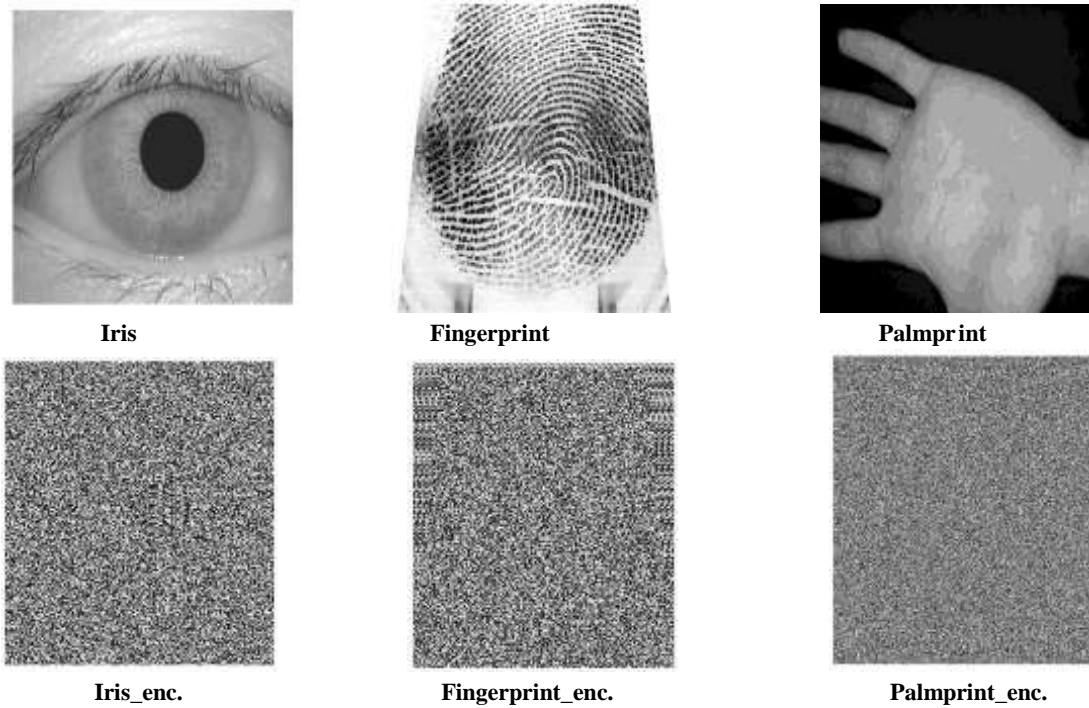


Fig 1: Encryption by the propose scheme

4. SECURITY ANALYSIS

An acceptable encryption algorithm must thwart all kinds of cryptanalytic threats such as statistical attacks and exhaustive search attacks, differential attacks and related key attacks [13, 14, 16]. In this section, several security tests are applied to the proposed cipher to demonstrate its satisfactory security level.

4.1 Statistical Analysis

From cryptanalysis point of view, statistical analysis may enable an attacker to crack the cipher and recover the plain image from its cipherimage. Indeed, several cryptography schemes have been successfully broken through the statistical analysis such as permutation based ciphers. Hence, to confirm the strength of the proposed cryptosystem, the statistical analysis based on histogram and adjacent pixel correlations analysis are performed. The obtained results demonstrate the ideality of the proposed cipher with respect to statistical attacks.

4.1.1 Correlation of Two Adjacent Pixels

The correlations between neighboring pixels are tested for horizontal, diagonal, and vertical adjacent pixels for the plainimage and the associated cipherimage. First, several pairs of adjacent pixels in different directions are randomly selected. Then, calculate the correlation coefficient between them according to Eq. 15. The results of the adjacent correlation analysis for horizontal pixels for iris image and its related cipherimage are illustrated in Fig. 2. The obtained values for the correlation coefficients in the plainimage and cipherimage are tabulated in Table 3 for different directions. Obviously, there is an extraneous correlation between adjacent pixels in the cipherimage. On the other hand, the plainimage appears well correlated adjacent pixels which prove that the success of the proposed scheme in decreasing such correlation.

Table 3: Obtained Values of C.C between Adjacent Pixels for Plainimage/Cipherimage

Direction	Plainimage	Cipherimage
Horizontal	0.9701	0.0679
Vertical	0.9752	0.0009
Diagonal	0.9539	-0.0459

4.1.2 Gray Histograms Analysis

Additionally, to prevent an opponent from exploiting the statistical features of the cipherimages to obtain valuable information about the plain image, the cipherimage must bear a high dissimilarity to the original image. The histogram of several encrypted biometric images and its related original biometric images are studied. One of these examples shown in Fig. 2 displays the histogram of a cipherimage and the histogram of corresponding original image denoted by iris image. It is clear that the encipher image histogram is uniform distributed and notably dissimilar to the relevant histogram of the corresponding original image and consequently does not afford any indication about the original plainimage. Thus, an opponent cannot apply any statistical analysis on the proposed cipher.

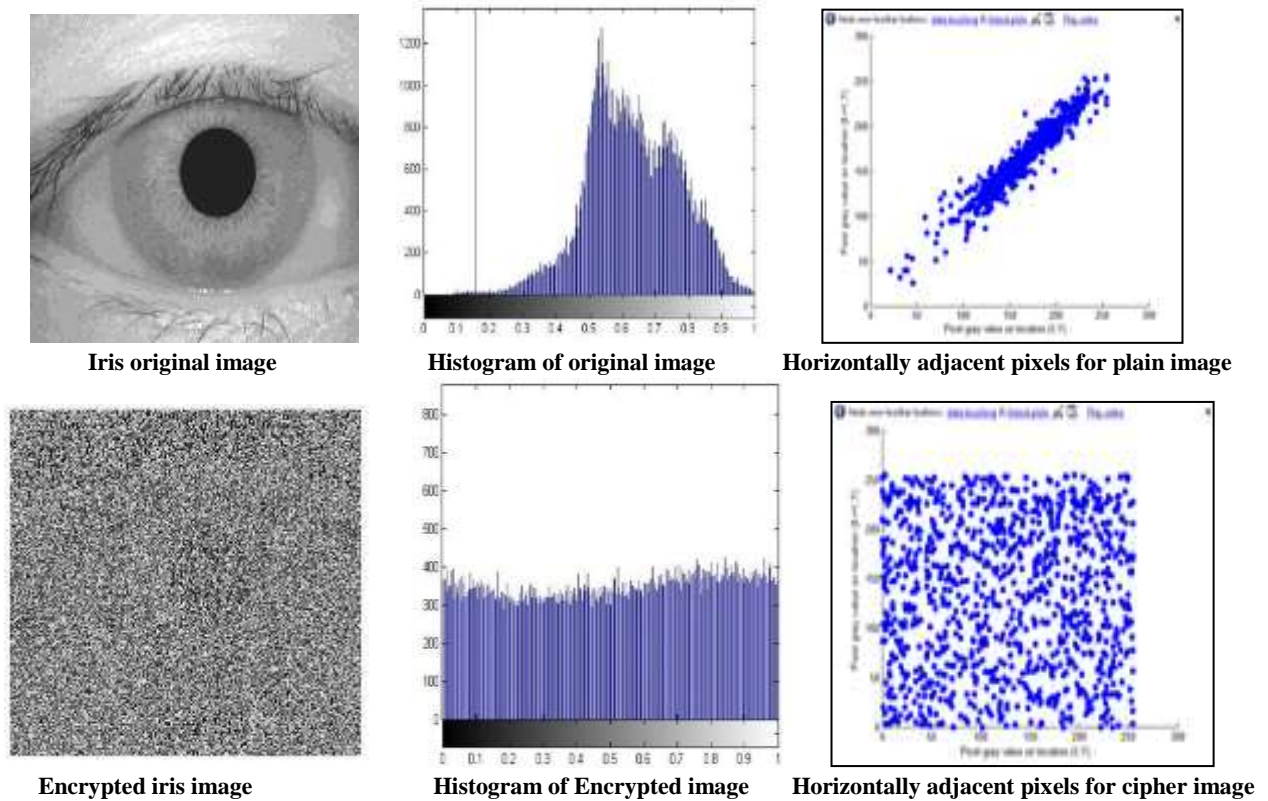


Fig. 2. Histogram and correlation coefficient of two horizontally adjacent pixels in plain image and its related cipher image

4.2 Information Entropy Analysis

Information entropy is considered as a significant indicator to randomness degree. According to Shannon’s theory [17], the entropy of an information source IS can be defined as follows:

$$H(IS) = - \sum_{i=0}^{2^L-1} P(IS_i) \log_2 P(IS_i) \quad (16)$$

Where $P(\mathbf{IS}_i)$ denotes the probability of symbol \mathbf{IS}_i , and L is the number of bits used in representation of symbols of the source IS. According to this definition, it is found that the idea value of entropy for a random image with 28 (256) gray levels equals 8. To test the safety of the suggested scheme against the entropy attack, the entropy values for several images encrypted by the proposed scheme are estimated and are displayed in Table 4. The obtained estimators are too close to the expected value of 8 of a perfect random image. Thus, the proposed image cipher can defy the entropy attacks.

Table 4: Results of Information Entropy Analysis.

Images	Entropy	Images	Entropy
Iris 001_1_1	7.9977	Fingerprint 100_R0_0	7.9982

Iris 001_2_1	7.9982	Palmprint 0001_m_L_01	7.9995
Fingerprint 100_L0_0	7.9984	Palmprint 0001_m_R_01	7.9994

-7-

4.3 SENSITIVITY ANALYSIS

A perfect encryption procedure must be wholly sensitive to small modification in the associated encryption key and the original plainimage. Namely, the trivial variation on a single bit in either the plainimage or the secret key should yield a significant change in the cipherimage (i.e. completely different enciphered image). To verify the robustness of the proposed algorithm, the following analysis is employed.

4.3.1 Key Sensitivity Analysis

One aspect of key sensitivity for a secure cipher is the failure of restoring the plainimage from cipher image if there is only a slight diversity between decryption and encryption keys. Really, this feature also promises the high resistance of the cryptosystem to brute-force attacks. On the other hand, the resulting cipher image from a small changing in the encryption key must result in extremely different enciphered image. To test the high sensitivity of the proposed cipher to the changes in the secret key, the following steps are carried out:

- 1) Use the secret key K_1 to encrypt the plainimage shown in Fig. 4(a) and the ciphered image is denoted as A as depicted in Fig. 4(b).
- 2) The same plainimage is enciphered again by the key K_2 where this key is different from the previous key K_1 in only the most significant bit. The resulting image B is displayed in Fig. 4 (c).
- 3) Finally, the same image is enciphered again by the key K_3 where this key is different from K_1 in only the least significant bit. The resulting image C is illustrated in Fig. 4 (d).
- 4) Compare the enciphered images A , B and C to find their differences.

Fig. 4 shows the plainimage and the three corresponding cipher images produced from the applications of the aforementioned steps. To computationally compare these encrypted images, the correlation between each pair of them is evaluated. Table 5 lists the obtained results for correlation. It is obvious that modifying only one bit of the secret key yields entirely distinct enciphered images with insignificant correlation between them.

Table 5: Obtained C.C Analysis for Cipher by Slightly Different Secret Key.

Image1	Image2	Correlation coefficient
Encrypted image A	Encrypted image B	- 0.0005
Encrypted image B	Encrypted image C	- 0.0017
Encrypted image C	Encrypted image A	- 0.0056

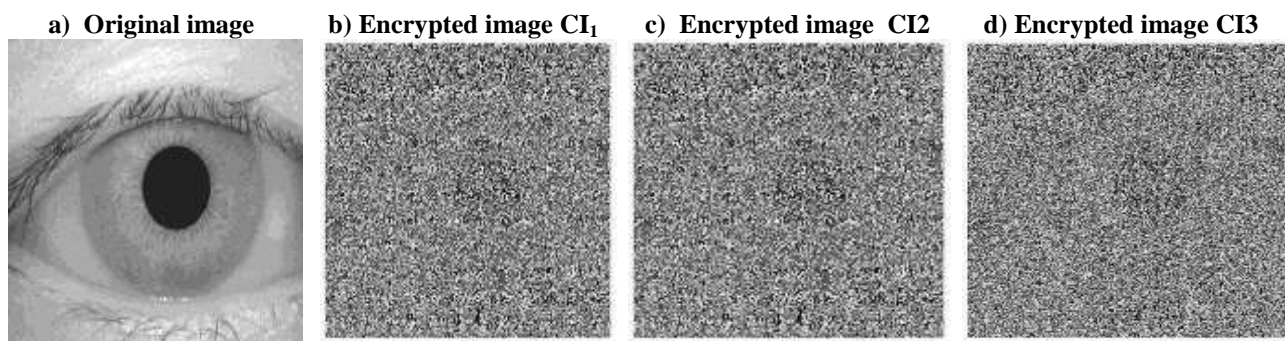


Fig. 4: Key sensitive

Moreover, the attempt to recover the original image from the enciphered one with slightly different key fails.

Particularly, Fig. 5 (a) and Fig. 5 (b) explain the original plainimage and the associated encrypted image produced using the secret key K1, Severally, whereas Fig. 5 (c) and Fig. 5 (d) draw the retrieved images from the decryption procedure with a correct key K1 and a slightly different key K2, respectively. Distinctly, the deciphering with a somewhat distinct key cannot succeed.

-8-

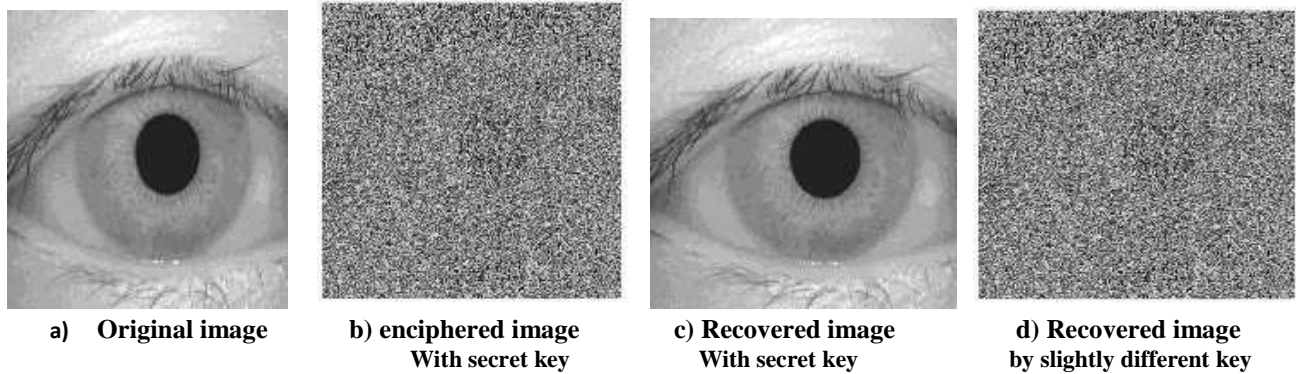


Fig. 5: Key sensitive test

4.3.2 Plainimage Sensitivity Analysis

Another demand for a perfect encryption technique is its high allergy to little changes of the original image. To evaluate the suggested algorithm in this trend, two standards can be applied. The first one is the Number of Pixels Change Rate (NPCR) and the second is the Unified Average Changing Intensity (UACI). Presume that P1 and P2 are two images with only one pixel various, and the secret key applied is indicated by K, the next proceedings are utilized:

- 1) Use the secret key K to encrypt the first plainimage P_1 displayed in Fig. 4(a) and the related enciphered image is denoted as C_1 .
- 2) The same secret key K is used to encrypt the second plainimage P_2 which is different from P_1 in only one pixel. The resulting image is named C_2 .
- 3) Finally, compute the values of UACI and NPCR according to the following equations[13,16]:

$$UACI = \frac{1}{W \times H} \left[\sum_{i,j} \frac{|C_1(i,j) - C_2(i,j)|}{255} \right] \times 100\% \quad (17)$$

$$NPCR = \sum_{i,j} \frac{D(i,j)}{W \times H} \times 100\% \quad (18)$$

Where W and H are the width and height of C_1 or C_2 and $D(i,j)$ is evaluated as follows:

$$D(i,j) = \begin{cases} 0 & \text{if } C_1(i,j) = C_2(i,j) \\ 1 & \text{otherwise} \end{cases} \quad (19)$$

The test is executed on iris biometric image of size 280 x 320, and the result of NPCR estimator is found to be over 99.76%. As well, the value of UACI estimation is calculated to be over 33.49%. The obtained values of NPCR and UACI demonstrate that the suggested cipher is robustly sensitive to modifications happened in the original image.

5. Conclusion

A prospective technique for encryption images is suggested which employs one of chaotic systems (tent map). The suggested is depended on a chaos- feedback mechanism. The major merit of the suggested scheme depends on the employment of biometric secret image extract an external secret key by utilizing

Arnold cat map for shuffling places of biometric pixels and then employs SHA-1 hash algorithm and logistic map to change the values of pixels, which enhance effective key. The encoding of each pixel depends on the secret key, the preceding enciphered pixel and the output of tent map. The performed empirical results explain that the encrypted image has tiny correlations between neighbor pixels, roughly uniform image histogram which can be seen as roughly random image. Furthermore, the plain image allergy parameters are near to their exemplary values of 99.76% and 33.49%. The suggested scheme has high plain-image sensibility. So the suggested scheme is resistive to statistical and differential threats. Thus, the suggested scheme can be powerful versus the joint assaults.

-9-

References

- [1] J. Menezes, S. A. Vanstone, and P. C. V. Oorschot, " *Handbook of Applied Cryptography*", CRC Press, Inc., 1996.
- [2] Gauravram, P. " *Cryptographic hash functions: Cryptanalysis, design and applications*", Ph.D. thesis, Faculty of Information Technology, Queensland University of Technology, Brisbane, Australia. 2003.
- [3] Kahn, D. " *The Codebreakers: The comprehensive history of secret communication from ancient times to the Internet*", (Revised edition). Scribner. 1996.
- [4] Lai, X., and Massey, J. L. " *Hash function based on block ciphers*", In EUROCRYPT, 1992, pp.55-70, 1992.
- [5] Rogaway, P., and Shrimpton, T. " *Cryptographic hash function basics: Definitions, implications and separations for preimage resistance, second preimage resistance, and collision resistance*", in FSE, pp.371-388, 2014.
- [6] Silva, J. E. " *An overview of cryptographic hash functions and their uses*", GIAC Security Essentials Practical Version 1.4b Option. 2003.
- [7] Sobti, R., and Geetha, G. " *Cryptographic hash functions: A review*", IJCSI International Journal of Computer Science Issues, 9(2), 461-479, 2012.
- [8] In Wikipedia (2014), " *Information Security*", Retrieved from, [online]. Available: http://en.wikipedia.org/wiki/Information_security
- [9] R Jun, P., J. Shangzhu, and L. Yongguo, " *Design and Analysis of an Image Encryption Scheme Based on Chaotic Maps*", International Conference ICICTA, pp. 1115-1118, 2010.
- [10] National Institute of Standards and Technology. " *A statistical test suite for random and pseudorandom number generators for cryptographic applications*", Special publication 800-22. Revision 1. 2008.
- [11] Z.Yun-peng, and Z. Zheng-jun. " *Digital Image Encryption Algorithm Based on Chaos and Improved DES*". Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics San Antonio. TX, USA. 2009.
- [12] Muhammad Khurram Khan, and Jiashu Zhang, " *Implementing Templates Security in Remote Biometric Authentication System*", IEEE Conf. Proceedings on CIS'06, China. pp. 1396-1400. vol.2, 2006.
- [13] G. Chen, Y. Mao, and C.K. Chui, " *A symmetric image encryption based on 3D chaotic maps*", Chaos Solitons Fractals. vol. 21. pp. 749-761, 2004.
- [14] Pareek ,N. K., Patidar , V., Sud, .K .K. , " *Image encryption using chaotic logistic map*", In Image and Vision Computing 24-926-934. Elsevier, 2006.
- [15] CASIA Iris Database . [Online March, 2009] <http://sinobiometric.com>.

- [16] I. A. Ismail, M. Amin, and H. Diab, "A digital image encryption algorithm based a composition of two chaotic logistic maps", International Journal of Network Security. vol. 11. no. 1. pp. 1-10, 2010.
- [17] Jiri Giesl, Ladislav Behal and Karel Vlcek, "*Improving Chaos Image Encryption Speed*". International Journal of Future Generation Communication and Networking Vol. 2. No. 3, 2009.
- [18] N.K. Pareek, Vinod Patidar, and K.K. Sud, "*Cryptography using multiple onedimensional chaotic maps*", Commun. Nonlinear Sci. Numer. Simul. 10 (7) 715–723, 2005.
- [19] Nitumoni Hazarika and Monjul Saiki, "*A Novel Partial Image Encryption using Chaotic Logistic Map*". International conference on Signal Processing and Integrated Networks (SPIN), 2014.

Cauchy Based fuzzy neural network with Mutual Subsethood Product Inference

Nelly S. Amer ¹, Hesham A. Hefny²

Abstract

This paper presents a mutual subsethood product fuzzy neural model based on Cauchy fuzzy sets. The proposed model has the ability to perform classification using both of numeric and linguistic inputs simultaneously. Fuzzy rule-based knowledge is translated into network architecture. Connections in the network are represented by Cauchy fuzzy sets. The firing degrees of the fuzzy IF-Then rule are based on fuzzy mutual subsethood similarity measure. Which is computed neither approximated nor numerically. It is computed by an exact formula. We focus on the classification ability of the model and demonstrate its performance on two benchmark classification problems: the IRIS data classification, Hepatitis medical diagnosis.

Keywords: fuzzy neural network, Fuzzy mutual subsethood.

1. Introduction

Fuzzy systems have been successfully used in a variety of applications, such as pattern recognition, automatic control and fuzzy inference systems [1]. It is very difficult to find a global function or analytical structure for a nonlinear system. In contrast, fuzzy logic provides an inference morphology that enables approximate human reasoning capability to be applied in a fuzzy inference system. Therefore, a fuzzy inference system employing fuzzy logical rules can model the quantitative aspects of human knowledge and reasoning processes without employing precise quantitative analysis.

In recent past, artificial neural network has also played an important role in solving many engineering problems. Neural network has advantages such as learning, adaptation, fault tolerance, parallelism, and generalization [2]. Fuzzy systems utilizing the learning capability of neural networks can successfully construct the nonlinear input output mapping for many applications. The term neuro fuzzy system [3] refers to an adaptive fuzzy systems, where neural network learning techniques are employed to adjust the shape of fuzzy sets, i.e. linguistic values, of the fuzzy rule base that constitute the model of the considered nonlinear phenomenon. Such a hybrid model has been applied efficiently in several applications including classification, function

¹ Dept. of computer sciences Institute of statistical studies and researches ,Cairo University.

² Dept. of computer sciences Institute of statistical studies and researches ,Cairo University.

approximation [4]. In the subsethood based fuzzy neural models, the adaptive fuzzy model can handle simultaneous admission of fuzzy or numeric inputs along with the integration of a fuzzy mutual subsethood measure for activity propagation

Fuzzy neural model based on mutual subsethood measure has a lot of applications. Several approaches have been proposed in literature to compute the firing degrees of the fuzzy IF-Then rule based on mutual subsethood similarity measure. Most research works made in this area adopt triangle or trapezoidal fuzzy sets due to the simplicity for obtaining an exact mathematical formula for the overlapping area. However, when dealing with continuous fuzzy sets, most of research works are based on numerical approximation, or approximate similarity formulas. Some researchers also used the possibility measure as a measure of similarity. In [5,6,7,8 and 9] the authors tried to calculate exact formula for similarity measure of Gaussian fuzzy sets, But these exact formulas were for four different possible cases of intersection, a general closed analytical formula could not be found, and hence high computational effort was done. Finally, in [10] , the generalized analytical formula of similarity measure of Gaussian fuzzy sets has been successfully obtained which can be used for distinguishability quantification. in [11], an exact analytical formula of mutual subsethood similarity measure of Cauchy fuzzy sets has been obtained which the work in this paper is inspired by it.

The organization of the paper is as follows: section two provides the architectural of the model. Section three describes the numeric and linguistic inputs. Section four presents the learning algorithm. Section five provides the set theoretic similarity measure for Cauchy fuzzy sets, and Section six demonstrates the application of the model pattern classification.

2. Architecture of the model

The proposed fuzzy neural model represents fuzzy rules of the form

If A_1 is SHORT and A_2 is TALL then B is MEDIUM.

Where A_1 , A_2 , B, SHORT, MEDIUM, and TALL fuzzy are fuzzy sets defined, respectively, on input or output universes of discourse (UODs). As seen in fig.1. The input nodes of the first layer in the fuzzy neural network represent domain variables or features, and output nodes represent target variables or classes. Each hidden node represents a rule, and input-hidden node connections represent fuzzy rule antecedents. Each hidden-output node connection represents a fuzzy-rule consequent. Fuzzy sets corresponding to linguistic labels of fuzzy if-then rules (such as SHORT, MEDIUM, and TALL), are defined on

input and output UODs and are represented by Cauchy membership functions specified by a center and spread. Fuzzy weights w_{ji} from rule nodes j to input nodes i are thus modeled by the center w_{ji}^a and spread w_{ji}^b of a Cauchy fuzzy set and denoted by $w_{ji} = (w_{ji}^a, w_{ji}^b)$. By the same way, consequent fuzzy weights from output nodes to rule nodes are denoted by $w_{kj} = (w_{kj}^a, w_{kj}^b)$.

Subsethood Product Fuzzy Neural Inference system (SuPFuNIS) can simultaneously admit numeric inputs as well as fuzzy inputs. Numeric inputs are first fuzzified hence all inputs to the network are fuzzy. Since antecedent weights are also fuzzy, then it requires a method to transmit a fuzzy signal along a fuzzy weight. In our SuPFuNIS model fuzzy mutual subsethood is employed to handle Signal transmission along the fuzzy weight.

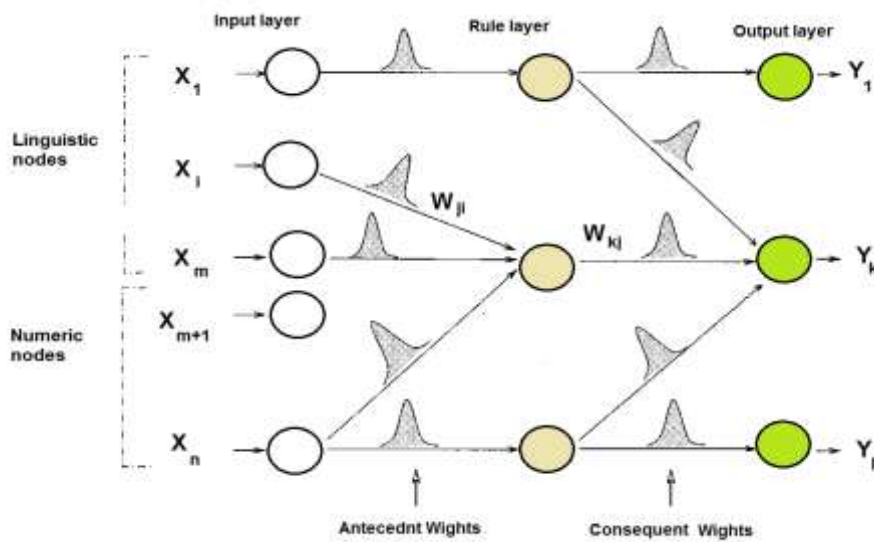


Fig. 1. Architecture of the fuzzy neural model.

3. Numeric and linguistic inputs

In this section we will show how the model can accept numeric inputs or linguistic inputs. It is known that The input features (x_1, \dots, \dots, x_n) can be either numeric or linguistic. Linguistic nodes accept a linguistic input represented by a fuzzy set with a Cauchy membership function and modeled by a center a_i and spread b_i . The linguistic input feature x_i is represented by the pair (a_i, b_i) . Numeric nodes accept numeric inputs and fuzzify them into Cauchy fuzzy sets. They are fuzzified by treating them as the center of a Cauchy membership function with a specified chosen spread. We chose a spread value of 0.5 for the applications presented in this paper. Therefore, the signal from a numeric node of the input layer also is represented by the pair (a_i, b_i) . Antecedent connections uniformly receive signals of the form (a_i, b_i) . Signals $S(x_i) = (a_i, b_i)$ are

transmitted to hidden rule nodes through fuzzy weights which is of the form (a_{ji}, b_{ji}) .

4. The learning algorithm

The SuPFuNIS model is trained by back propagation supervised learning [9]. The model is trained by repeating presentation of a set of input patterns drawn from the training set. The firing degree of the fuzzy IF-Then rule will be computed set theoretic which will be shown in the next section. The output of the network is compared with the desired value to obtain the error, and network weights are changed on the basis of an error minimization criterion. Once the network is trained to the desired level of error, it is tested by presenting a new set of input patterns drawn from the testing set.

5. The Fuzzy mutual subethood

The fuzzy mutual subethood is the set theoretic similarity measure which is used to measure the similar degree between two fuzzy sets [12].

The set-theoretic similarity measure usually used in interpretability analysis is:

$$S(A, B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}, \text{ where } A \text{ and } B \text{ are fuzzy sets.} \quad (1)$$

Where $|\cdot|$ denotes the cardinality of the set [13].

For the Cauchy fuzzy sets, first, the general form of Cauchy fuzzy set is

$$\mu(x) = \frac{1}{1 + \left(\frac{x-a}{b}\right)^2} \quad (2)$$

where a is the center, b is the width and x belong to the domain

Let $\mu_A(x)$ and $\mu_B(x)$ be two Cauchy fuzzy sets of the following forms

$$\mu_A(x) = \frac{1}{1 + \left(\frac{x-a_1}{b_1}\right)^2}, \quad (3)$$

$$\mu_B(x) = \frac{1}{1 + \left(\frac{x-a_2}{b_2}\right)^2} \quad (4)$$

The cardinality of Cauchy fuzzy sets calculation becomes integration as follows [11]:

$$|A| = \int_{-\infty}^{+\infty} \mu_A(x) = \int_{-\infty}^{+\infty} \frac{1}{1 + \left(\frac{x-a_1}{b_1}\right)^2} = b_1\pi \quad (5)$$

$$|B| = \int_{-\infty}^{+\infty} \mu_B(x) = \int_{-\infty}^{+\infty} \frac{1}{1 + \left(\frac{x-a_2}{b_2}\right)^2} = b_2\pi \quad (6)$$

$$|A \cap B| = b_{min}\pi + \Omega \quad (7)$$

Where

$$\Omega = (b_{max} - b_{min}) \tan^{-1} \left(\frac{a_{max} - a_{min}}{b_{max} - b_{min}} \right) - (b_{max} + b_{min}) \tan^{-1} \left(\frac{a_{max} - a_{min}}{b_{max} + b_{min}} \right)$$

$$b_{min} = \min(b_1, b_2)$$

$$b_{max} = \max(b_1, b_2)$$

$$a_{min} = \min(a_1, a_2)$$

$$a_{max} = \max(a_1, a_2)$$

(8)

The generalized formula for the similarity measure between two Cauchy fuzzy sets A and B with centers a_1 and a_2 and widths b_1 and b_2 can be obtained directly as follows:

$$S(A, B) = \frac{b_{min}\pi + \Omega}{b_{max}\pi + \Omega} \quad (9)$$

6. Applications in pattern

The pattern classification problem deals with categorization of an unseen pattern to 1-of- classes. These classes are either predefined or are learned based on the similarity of patterns. Common approaches for pattern recognition include neural networks, statistical methods, discriminant analysis, fuzzy systems, neural networks, and hybrid methods [5]. Attempts have also been made to design hybrid systems involving multiple models.

Our model in this paper has been tested on two different pattern classification benchmark data sets: IRIS data and hepatitis diagnosis [6]. For the IRIS classification problem, we report the resubstitution error which is uses the same data for training and testing and it is computed as the number of misclassification of patterns of all classes. For the hepatitis diagnosis

classification problems, we report the test errors where the network is trained on a training data set that is distinct from the test set. The test error is the number of misclassification of patterns of all classes.

6.1 IRIS Data Classification

IRIS data involves classification of three classes of the IRIS flower namely, IRIS setosa, IRIS versicolor, and IRIS virginica on four feature of the IRIS flower:

- 1) Sepal length;
- 2) Sepal width;
- 3) Petal length;
- 4) Petal width.

There are 50 patterns (of four features) for each of the three subspecies of IRIS flower. The input pattern set thus comprises 150 four-dimensional patterns. This data is available from UCI repository of machine learning databases from <http://www.ics.uci.edu/~mlearn/MLRepository.html>.

The input layer consists of four numeric nodes and output layer comprises three class nodes. The number of nodes in the rule layer depends on the number of rules for which the network is trained. To train the network, initially, the centers of antecedent weight fuzzy sets were randomized in the range of the minimum and maximum values of respective input features of IRIS data. These ranges are (4.3000, 7.9000) which represent the minimum and the maximum of sepal length of 150 IRIS patterns, (2.0000, 4.4000) which represent the minimum and the maximum of sepal width of 150 IRIS patterns, (1.0000, 6.9000) which represent the minimum and the maximum of petal length of 150 IRIS patterns, and (0.1000, 2.5000) which represent the minimum and the maximum of petal width of 150 IRIS patterns. The centers of hidden-output weight fuzzy sets were randomized in the range (0, 1), and the spreads of all fuzzy weights were randomized in the range (0.2, 0.9). The feature spreads are taken as 0.5. All 150 patterns of the IRIS data were presented sequentially to the input layer of the network for training. The learning rate and momentum were both taken as 0.0007 and kept constant during the training period.

Once the network was trained, the test patterns (which again comprised all 150 patterns of IRIS data) were presented to the trained network and the resubstitution error computed.

TABLE I
BEST RESUBSTITUTION ACCURACY FOR IRIS DATA FOR DIFFERENT
SOFT COPMTING ALGORITHMS

Method	Rules	Resubstitution accuracy (%)
FuGeNeSys	5	100
NEFCLASS	7	96.7
ReFuNN	9	95.3
EFuNN	17	95.3
FuNe-I	7	96.0
SUPFUNIS	5	100
Our Model	3	100

Table I, [8] compares between our model and other techniques according to resubstitution accuracy which indicates that our model can strongly classify IRIS data with least number of rules.

6.2 Medical Diagnosis

This problem deals with hepatitis diagnosis which requires classifying patients into two classes Die or Live on the basis of features which are both numeric and linguistic. The data can be obtained from <http://www.ics.uci.edu/~mlearn/MLRepository.html>.

The hepatitis data set has 155 patterns of 19 input features with a number of missing values. There are six numeric features namely Age, Bilirubin, Alk Phosphate, SGOT, Albumin, and Protime, and the remaining 13 features are linguistic in nature.

As there are a number of missing data, preprocessing of data is required. The data set consists of 75 patterns that have one or more features unspecified. A new set of data was formed by fitting some of the missing numeric values. Twenty patterns which had either a missing symbolic feature value, or more

than two missing numeric feature values were first discarded. The missing numeric values in the remaining 55 incomplete cases were filled with the average value of the missing feature calculated on a class-wise basis from the 80 original complete data [5 and 6]. By this method we were able to reconstruct a data set of 135 patterns. The numeric features of these 135 patterns were normalized feature-wise in the range [0, 1]. Symbolic features(yes/no or male/female) were represented by constructing two fuzzy sets: the symbolic value “no” represented by a fuzzy set with Cauchy membership function having center as zero and spread as 0.5, and “yes” represented by a Cauchy membership function centered at one and spread 0.5. The spreads were assumed to be trainable during the learning procedure.

Experiments were conducted using two data sets: Data Set 1 comprising of only 80 of 155 patterns that were originally complete in all respects; and Data Set 2 comprising 135 patterns (80 originally complete and 55 reconstructed). For training, 70% patterns were randomly chosen and the remaining 30% were used for testing. Five combinations of such 70 %(train) and 30 %(test) were randomly generated separately for Data Set 1 and Data Set 2. Experiments were then conducted on each of these individual data set combinations using SuPFuNIS architecture which had 19 input nodes, 3 nodes in the hidden layer and 2 nodes in the output layer. During the training, both learning rate and momentum were kept constant as 0.0007.

Table II shows the average classification accuracy obtained using SuPFuNIS for both data sets, compared with Kumar [8] to solve the same problem and it refers to the testing accuracy is raised in our model where the average of testing accuracy of data set 1 in our model is 94 % where in Kumar was 91.67. In data set 2 our testing accuracy average is 97 where in Kumar model was 96.5.

TABLE II
TESTING ACCURSCY IN % USING THREE RULES FOR HEPATITIS
DATA

Method	Experiment	1	2	3	4	5	Average %
Kumar Model	Data Set1	91.67	100	87.5	87.5	91.67	91.67
	Data Set2	97.5	97.5	97.5	92.5	97.50	96.50
Our Model	Data Set1	95.83	95.83	95.83	91.67	91.67	94
	Data Set2	97.5	95	97.5	97.5	97.5	97

Conclusion

In this paper we proposed a SuPFuNIS model which can handle both numeric and linguistic inputs and we showed that it was different of other SuPFuNIS models where first it used Cauchy function membership and a lot of other model used Gaussian membership. Second it computed the mutual subsethood by exact formula not approximated, not numerically and not exact with different cases for intersections as Kumar. And hence the model was low computationally and its efficiency appeared in the two patterns classification cases.

References

- [1] Song Hengjie et al, “a fuzzy neural network with fuzzy impact grade”, *Neurocomputing* 72, pp 3098–3122, April 2009.
- [2] Hengjie Song et al, “Implementation of Fuzzy Cognitive Maps Based on Fuzzy neural Network and Application in Prediction of Time Series”, *IEEE Transactions on fuzzy systems*, Vol. 18, No. 2, pp 233-250, April 2010.
- [3] Amit Mishra and Zaheeruddin, “design of hybrid fuzzy neural network for function approximation”, *J.Intelligent learning systems and applications*, 2, pp97-109, May 2010.
- [4] Antonios D. Niros and George E. Tsekouras, “A novel training algorithm for RBF neural network using a hybrid fuzzy clustering approach”, *fuzzy sets and systems*, 163, pp 62-84, 2012.
- [5] Sandeep Paul and Statish Kumar, “Subsethood product fuzzy neural inference system”, *IEEE Transactions on neural networks*, Vol.13, No.3, pp578-599, May 2002.

[6] Sandeep Paul and Statish Kumar, “Subsethood based adaptive linguistic networks for pattern classification”, IEEE Transactions on systems, man and cybernetics- part c: applications and reviews Vol.33, No.2, pp248-258, May 2003.

[7] Sandeep Paul and Statish Kumar, “fuzzy neural inference system using mutual subsethood products with applications in medical diagnosis and control”, IEEE international fuzzy systems conference, pp 728-731, 2001.

[8] C.Shunmuga Velayutham and Statish Kumar, “Some applications of an Asymmetric Subsethood Product Fuzzy Neural Inference System”, IEEE international conference on fuzzy systems, pp 202- 207, 2003.

[9] C.Shunmuga Velayutham and Statish Kumar, “Asymmetric Subsethood-Product Fuzzy Neural Inference System (ASuPFuNIS)”, IEEE Transactions on neural networks, Vol. 16, No. 1, pp 160 – 174, January 2005.

[10] Hesham A.Hefny, “Comment on: Distinguishability quantification of fuzzy sets”, Information Sciences (177), pp 4832-4839, February 2007.

[11] Nelly S. Amer and Hesham A.Hefny, “Analytical formulas for similarity, possibility and distinguishability measures of Cauchy type fuzzy sets with comparison to Gaussian fuzzy sets”, Proceedings of the 7th IEEE International Conference on Intelligent Computing and Information Systems (ICICIS15), vol. 3, pp. 22-27, 2015.

[12] Shifei Ding and Fengxing Jin, “A Novel Fuzzy Likelihood Measure Algorithm”, IEEE International conference on computer Sciences and software Engineering, pp 945-948, 2008.

[13] Shang- Ming Zhou and John Q.Gan, “Low –level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling”, Fuzzy Sets and Systems 159, pp 3091-3131, June 2008.

Survey on Land Change Modeling

Khalid A. Eldrandaly^a, Mamdouh M. Abdeen^b, Safa A. Abdelkareem^c

Abstract

The presented paper covers the subject of land change modeling, including the current approaches, applications, and the future prospects. A lot of classifications schemes do exist for the land change modeling approaches, many of which don't cover the massive variety of the land change models. Land change models requires to be organized and grouped to ease the procedure of selecting the most suitable one given a certain application or to satisfy a specific purpose. Applications of land change models and the preferred model for each application form a critical step included within the modeling land change process. The more precise the model fits a certain application including all its requirements and specifications the higher the accuracy of the simulated results, and more related to real world they will turn out to be. Although the last decade showed a great development not only in the field of land change modeling, but also in the land change science as a whole, still a number of problems do exists and needs to be addressed in order to move the land change modeling to the next level. Despite of the existing problems, chances for developments in the future are present. Developments such as the integration between the different approaches of land change modeling must be further studied to benefit from the advantages of different approaches and to limit drawbacks that accompany each approach. Some other future development ideas described in the future prospects section.

Key words: Land use/cover change, Land change modeling, LULCC applications.

1. Introduction

Land use land cover change(LULCC) presents the way human being change and modify the surface of Earth[1]. Land change modeling approaches form the primary foundation for creating a model, a model that will not only be concerned with the study of changes in land use and land cover, but also will identify the causes and consequences of such a change [2]. Prediction of future changes and simulation of different scenarios based on different policies can also be estimated using land change models. Such capabilities form a great interest in different disciplines [3], [4]. This paper aims to, discuss and review the current practices of land change modeling approaches; to identify a set of models and some examples of its applications, and to propose ideas for improving land change modeling. This paper is organized to discuss some of the different classifications of land change modeling approaches, pick out the most appropriate classification that does handle most, if not all models of land change, discuss the conceptual basis that convey the idea of each approach and the models built upon it, also to talk about the cons and pros of each approach, and to demonstrate each approach separately and in comparison with each other. Also to demonstrate different applications and models that may be used to study those applications. Based on the discussion of different approaches, models, and applications opportunities to improve the future of land change models will be shown in the last sections of the paper.

^{a, c} Information Systems Department, Faculty of Computers and Informatics, Zagazig University, Al-Sharqiyah, Egypt

^b Remote Sensing and Space Sciences, 23 Joseph Broz Tito St., El-Nozha El-Gedida, Cairo, Egypt

Section 2 will discuss the current state of land change modeling approaches. Although many approaches do exist it would be helpful if each approach can be classified under a particular category. The current practices section discusses a classification of different approaches that can include almost all land change modeling approaches under one of its categories. Section 3 discusses the applications that use pattern based models to model past changes and predict future ones. . Section 4 of this survey paper focus on the future prospects proposing ideas for possible advancements that can help improve land change modeling. Finally section 5 concludes this paper.

2. Current practices:

Land change modeling approaches can be sorted according to different features. Although plenty approaches do exist [1], [2], [5]–[7]. Some researchers focused on a specific set of approaches such as spatially explicit modeling approaches [8], [9] or approaches for descriptive models (All, 2016). Others described a much wider range of approaches [1], [4], but still a much general classification that all or at least most of the approaches can be categorized under one of its classes. The classification proposed in [10], [11] suggested different approaches for land change modeling (Figure. 1). The classification covered a wide range if not all of modeling approaches and maintained a common ground to distinguish between modeling approaches while keeping in mind the theoretical and practical structure of each approach. Anyone who is interested in the theoretical and empirical foundation and much more detail of any given modeling approach can refer to previous reviews i.e.[10], [12]

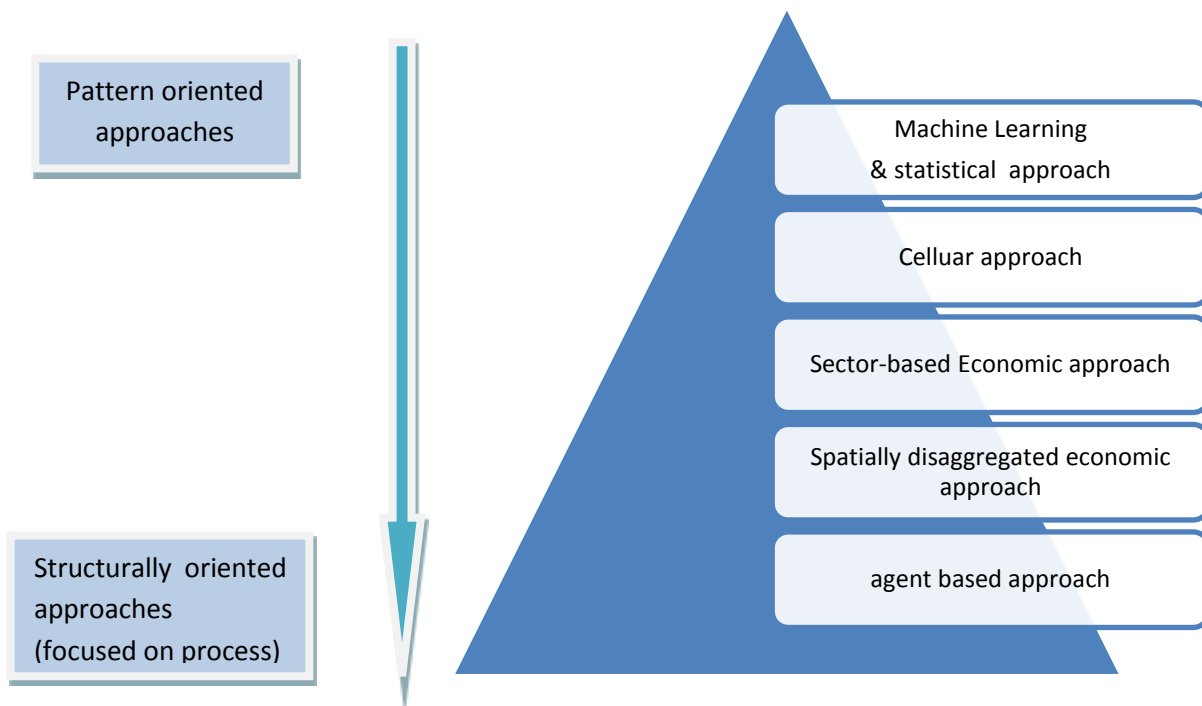


Figure.1. LCM Approaches: current practices.

2.1. Machine learning and statistical approach

This approach is considered to be a pattern based approach, and it has been using some kind of statistical model along with algorithms (i.e., Classification and regression trees, logistic regression, and artificial neural networks[7], [13]) to demonstrate relationships between changes of land use or cover and leading variables of the locations where they are likely to take place[4], [10], [12]. Machine learning is more appropriate whenever data related to patterns can be obtained while theory required for process can't. Whenever short-term views and predictions of the near future are in need machine learning approaches proven to be useful. In event of huge data or larger quantity of data sets the ability of the algorithms contained within this approach have proven to perform highly. This particularly means that the algorithm can run with limited computational requirements[4]. The algorithms are particularly used to project and generalize relationships found between inputs and outputs[10].

2.2 Cellular approach

The concept behind cellular automata is that each cell can carry a specific current state and can be moved to a next state which can be calculated using the current state, some transition rules and the neighbor cells also affect the process[4]. The bases of this method are shaped merging three components. The three components are suitability maps which indicate the degree in which a specific land can be used for a particular purpose[14], the second component focus on the impress of neighborhood on the land units used to mimic the study area, and the third component include information about the quantity of expected change of land[10].

The latter component enhances the power of the approach to predict and study the changes that are more likely to happen in the future[15]". The similarity between data structure in cellular models and land cover data abstracted from remote sensing makes it simple to start processing. Despite the simplicity of the three basic components forming cellular method seem to be, the format of models starts to develop a form of complicity when replacing the regularly formed pixels with spatial units. The selection of cellular models is not necessary because of the way it fits a specific case, sometimes one can choose to formulate the case study using cellular models due to their simplicity and well defined structure [10].

2.3 Econometric approaches

The third approach depends on the economic sector. Economic models connect outcomes that are generated from microeconomics actions that define the interaction between demand and supply to produce land use patterns[4]. The economic models determine the relationships between available and desired states which produces a link between prices and land use patterns. One of the features that characterize economic models its capability to explicitly determine how human being decisions affects land use, this quality is not contained within different approaches like the statistical or machine learning approaches[4]. this is particularly not an easy task taking in consideration the complexity of human

decisions and thinking mechanism[4]. The following two subsections and figure 2 define two different types of economic models. The sector based models and the spatially disaggregated models [4], [10]. The sector based models which can work on aggregated scale or spatially disaggregated models which operated at disaggregated scale.

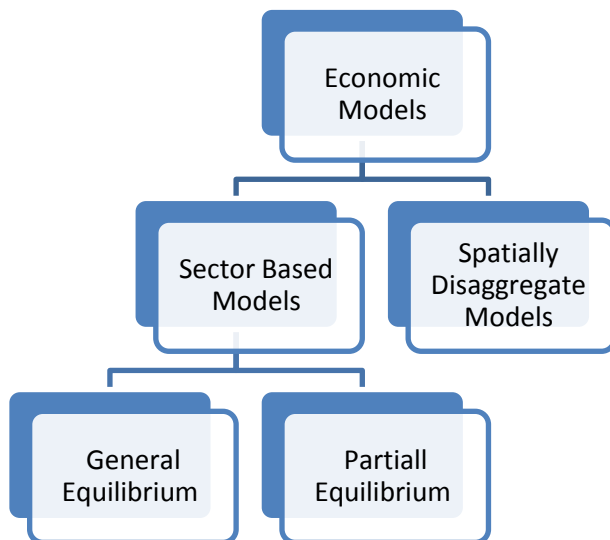


Figure.2: Economic models.

2.3.1 Sector-Based Models

The sector based economic approach makes a great use of equilibrium structural models, whether general or partial, this is for the representation of demand and supply processes included in land system and defined by economic sectors which are based on trade and economic activity. Models of this approach basically concentrated on trade activities which can be detected in regions or sectors along with inputs and outputs, this is mainly to define demand and need for a specific land type. As mentioned before, those models are mostly structural which means they represent supply and demand explicitly as a contributing factor to market equilibrium. The scope of an economic system which is represented by sector-based model differentiates between models: partial equilibrium and general equilibrium models differ in their representation of the economic system. General equilibrium represents interactions and feedbacks between several sectors along with the global economy. While partial equilibrium models cares more about the details that describe a specific sector as a closed system, which means no linkages between the model and the rest of the outer economic.

Partial equilibrium models also determine prices, productions, and shares of land within a specific geographic area such as a country, afterwards uses the same land in different sectors. In addition, partial equilibrium models assume that the economic conditions in the rest of the world are fixed.

2.3.2 Spatially Disaggregated Economic Models

Spatially disaggregated economic models are thought to be structural models either fully structured or reduced structural. Reduced structural model mechanisms can be only implicitly illustrated through model specification and chosen variables. Models of this approach define the relationships that effect on the spatial equilibrium in land systems. One of the aims of this approach is to help investigators to fully interpret and acknowledge land-use and land-cover changes as an outcome of decisions made individually, this is possible because this approach models is designed based on the macroeconomic theory. The theory of microeconomics and the approaches built on its conceptual bases usually used to evaluate the implications of variables included in the model.

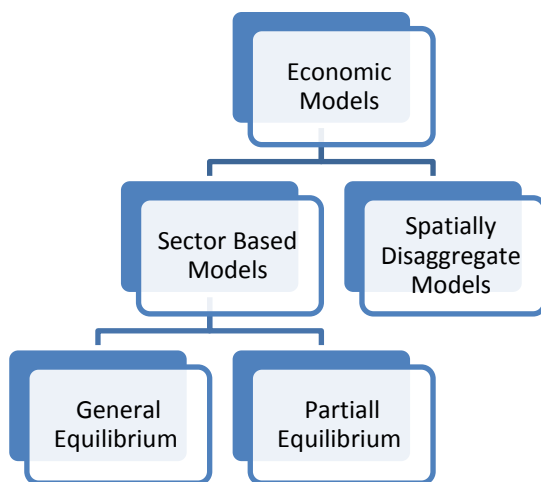


Figure.2: Economic models

2.4 Agent-Based Approach

Agent based models represent systems that are organized by a group of actors that interact with each other, these systems basically referred as multi-agent[4], [10]. When it comes to land change we can call land owners, farmers or any decision making figure as agents. Agents can be defined as any actor that contribute to the operation of decision making or the actor that perform a set of actions which has an effect on land use or land cover patterns[10]. Agents may be conceived as distinct objects that are accompanied by attributes and behaviors. Agents are associated with attributes that may be discretely or continuously measured. Interactions may occur between agents or between agents and surrounding environment this is particularly helpful in the process of collecting information or taking actions that affects their settings[2]. Although being spatially explicit is not always true for agent-based models, those involved in researching land change usually are. This means that agents and any associated actions are pointed at specific locations on Earth’s surface. Agent based models are also considered to be structural models that use explicit processes to represent land change[10].

2.5 Hybrid Approaches

This section is dedicated to illustrate how a mix of the previously stated approaches can be used to address land change in the form of changed patterns and processes. Hybrid approaches are responsible for integrating theories and data for different environments[2]. Different theories, including economics and geography and many other contribute in the process of interpreting land change. Most theories cannot fully explain efficiently the complexity that is involved in land use decision making. The main concept that the hybrid modeling approaches are based on is its capability to merge various conceptual frames and theories while representing a system. The strengths of hybrid modeling approaches are derived from the individual strengths of each approach, also it minimize the drawbacks that may happen to each approach alone when they are combined[16].

2.6 Comparison of land change modeling approaches

Although the comparison of different models and modeling approaches may be difficult because of the different fundamental bases[1]. The following sub-sections include some aspects to compare the main modeling approaches stated above.

2.6.1 Pattern based vs. Process based

When describing the arrangement of the approaches, whether the approach focused on patterns or based on processes, the most patterns based approaches are the Machine Learning and Statistical and cellular approaches, while the Spatially Disaggregated Economic Models, Sector-Based Economic Models, and the Agent Based Models are more process focused as shown in figure. 1.

2.6.2 Likely outcome type

Each approach tented to be used for a certain outcome type. Machine Learning and Statistical approach most likely used to produce land cover. While Spatially Disaggregated Economic Models and Sector-Based Economic Models are usually used for land use. Both the land use and land cover have been produced from cellular approaches and Agent Based Models.

2.6.3 Data requirements

Usual input data requirements differ from one approach to the other, for instance the Machine Learning and Statistical approach requires land cover maps that were taken across two or more different time slots and some maps for single or multiple variables the will be used in the prediction process as inputs. While cellular approaches may require the same data inputs as machine learning approach the only difference it can only require a single map of land cover at a certain time point. The third approach discussed above (i.e. Spatially Disaggregated Economic Models) requires only data not maps like the previous two approaches about land use or land cover at one or more time points. Also takes economic and biophysical variables that greatly affect the process of demand and supply for land. Sector-Based Economic Models need the following inputs, the economic variables that effects on the aggregate demand and supply which include commodity price and trade value at regional or county scale. The final approach mentioned above which is the agent-based requires data for describing the agent characteristics, whether quantity or quality data involved in decision processes. Also requires data on land use or land cover at one or more time points.

2.6.4 Preferred uses

As shown in figure (4) Machine Learning and Statistical approach commonly used for producing predictions by recruiting with past patterns, while Cellular approaches recommended to be used for predicting land cover patterns, also for evaluation spatial changes without using market feedbacks. Spatially Disaggregated Economic Models recommended for recognizing the effect of the key variables on land change outcomes, this is true if the model was a reduced form model. While structural models of this approach used to simulate effects of changes of policy on the outcomes of the land market. Sector-Based Economic Models used to predict the aggregate land change which happened under a various set of changes built upon market which may affect the demand and supply. Agent-based requires data explores processes of land change usually under certain set of conditions. Also explores the effect of out changes in a system where they haven't occurred. Exploring future scenarios where past patterns proven to be insufficient and cannot refer to future patterns

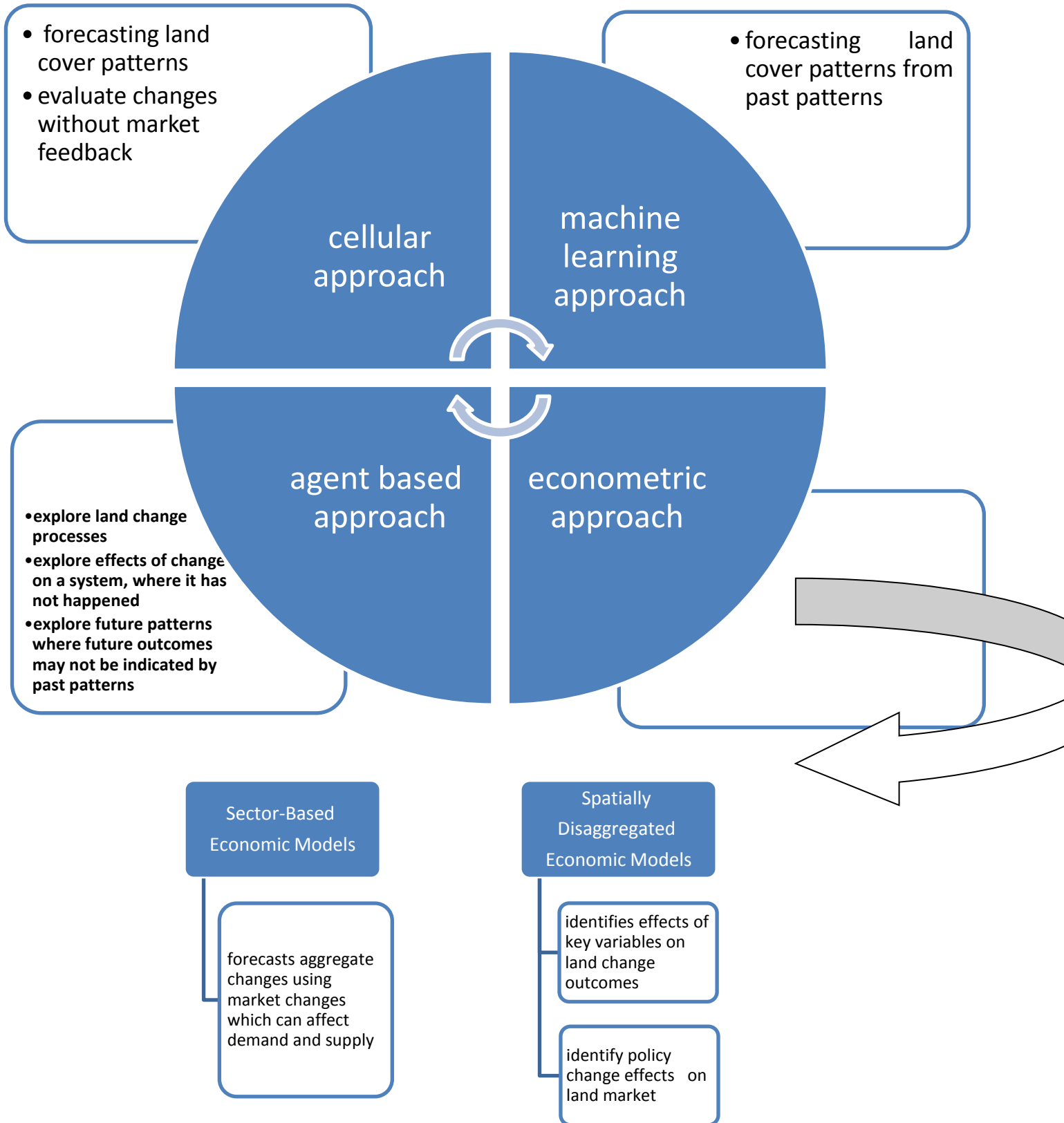


Figure.3. Recommended uses for each modeling approach

Figure (4) describes a hierarchy that the coming subsections will illustrate. Starting with the modeling approach or model methodology category, a subset of selected models was chosen based on how common they are or based on previous literature and has been reviewed by a number of technical experts in the field of land change science[18]. This is only a sample of models from each category farther models do exist, but they are not explained in this paper. This paper is more focused on the pattern based approaches which are the machine learning and statistical approach and the cellular approach this is why the models application's of those approaches are described in details while the other approaches are mentioned with some references for more information.

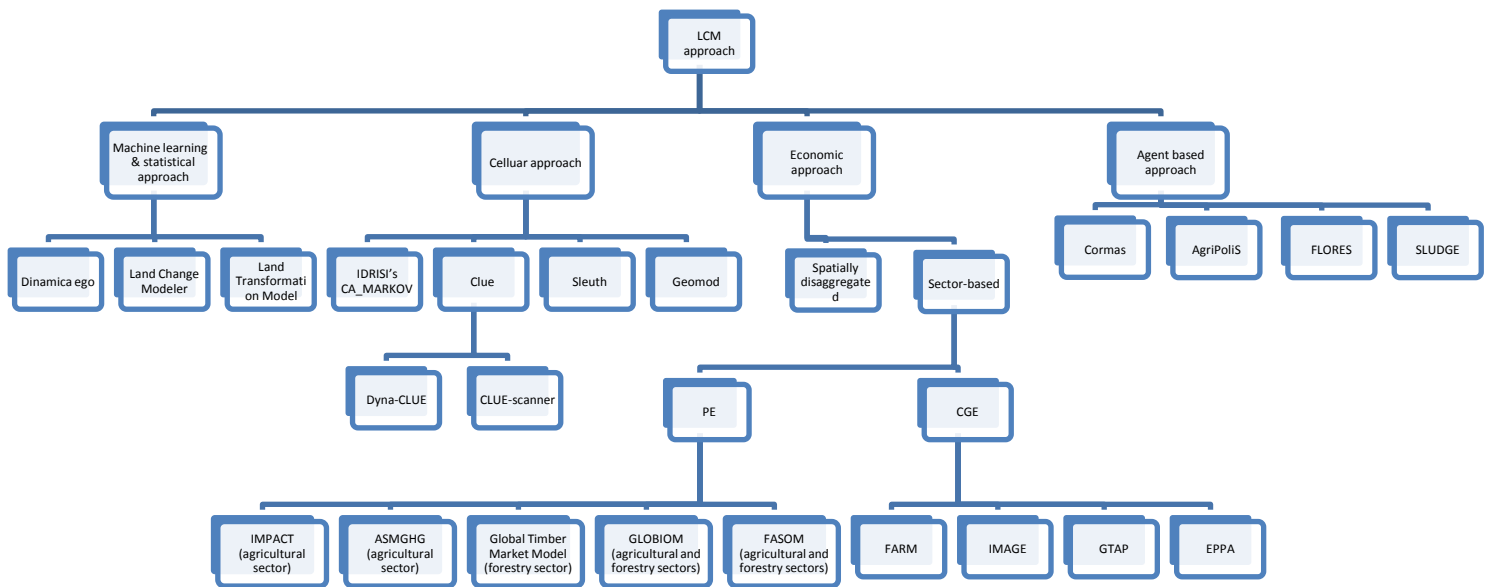


Figure.4. Hierarchical diagram to classify models used for the applications section

3. Applications of land change modeling approaches

The following figure describes a hierarchy that the coming subsections will illustrate. Starting with the modeling approach or model methodology category, a subset of selected models was chosen based on how common they are or based on previous literature. also a large number of the selected models were mentioned in the "Advancing Land Change Modeling: Opportunities and Research Requirements" book which was written by "Committee on Needs and Research Requirements for Land Change Modeling" and has been reviewed by a number of technical experts in the field of land change science. This is only a sample of models from each category farther models do exist, but they are not explained in this paper.

During the search conducted using both Google scholar scientific search engine and ISI's web of science/knowledge using the name of the model as a keyword, a large number of articles were found to be related to the topic. For the sake of simplification and organization farther search criterion was proposed so that no articles before the year 2010 were reviewed in this paper, also all articles that are not written in English were neglected and only the most rated or top cited or most recent articles when the search was conducted are discussed.

This paper is more focused on the pattern based approaches which are the machine learning and statistical approach and the cellular approach this is why the models application's of those approaches are described in details while the other approaches are mentioned with some references for more information.

3.1 Applications of Machine learning and statistical approach

Model name	Model description
Dinamica ego	is a software written in C++ and java and it includes a set of algorithms that help perform all the operations concerned with spatial analysis[19].
Land Change Modeler	Land change modeler is considered to be Terrset's (previously idrisi's) integrated set of tools designed for facilitating the process of assessment and prediction of land changes[20].
Land Transformation Model	The LTM considered being a neural network model. [21],the LTM relates both the potential (dependent) variables with the independent variables and develop numerical framework to include both inputs and output and trains on the relationship between them unit acceptable fit is reached to simulate land use cover change [21], [22].

Table.2 machine learning and statistical approach models example

Model name	Application	Integration with other models	References

Dinamica ego	deforestation, urban dynamics, agriculture expansion, forest fires	-	[23]
	to study the role of protected areas in reducing carbon fluxes caused by deforestation and the model helped predicting the carbon emission	carbon bookkeeping model	[24]
	to study changes in oil palm plantation expansion	-	[25]
	to study the effect of urbanization and climate changes on future flooding	climate change model	[26]
	to predict and model changes and growth of urban land	-	[27]
	deforestation modeling & projecting the future deforestation	-	[28]
	Modeling, calibrating, and validating fire regimes	-	[29]
	to optimize the ecosystem services	-	[30]
	helped studying the expected effects of agricultural expansion and climate change and how they both will affect on soil erosion	-	[31]
Land Change Modeler	to study urbanization, deforestation, habitat modeling and much more	-	[32]
	To monitor land cover changes of the dry land forest landscape.	-	[33]
	the impacts of land use changes and how it affects the erosion risk.	-	[34]
	maintain protected & to minimize biodiversity loss	climate change model	[35]
	to analyze urban sprawl & to model urban growth	-	[36]
	the spatial analysis of land abandonment	-	[37]
	To model the land use and climate patterns and how they affect surface water quality	-	[38]
	spatial planning and to produce different land-use zoning scenarios	-	[39]
	Model to assess and evaluate global biodiversity loss.	-	[40]

	to spatially asses farming systems and how they affect deforestation, forest re-growth, and agriculture change	-	[41]
	Timber plantation and forest expansion from other land cover and the analysis of causes of change	-	[42].
	Prediction of future land use land cover	-	[43]
	Urban growth analysis and prediction	-	[44],[45]
	The prediction of transformation to certain land category in the future	-	[46]
Land Transformation Model	to study urban growth	-	[47]
	Urban prediction	-	[48], [49]
	predict vacant land	-	[50]
	how land use and water quality affect each other	hydrologic models	[21]

Table.3 machine learning and statistical approach applications

3.2 Applications of Cellular Approach

The following subsections will discuss some of the most used cellular based models to simulate land changes and to project future changes.

Model name	Model description
IDRISI's CA-MARKOV	The CA-Markov model is a model that is included in the IDRISI Selva software, and it is essentially used to simulate land cover changes and to predict the future patterns[51].
Clue	The clue model is concerned with studying the conversion of land use and how such changes happen and what effects will arise from such changes. The clue model can operate at international, continental, and regional scales based on the required application and study area. New versions of the CLUE model has been developed two of the latest versions used are: Dyna-CLUE and CLUE-Scanner [52].
Sleuth	Sleuth model is considered to be one of known and strongly tested examples of cellular automata models[4], [53]..Sleuth contains two coupled CA models: the land change model and the urban growth model to asses historical changes and predict future land use patterns[54].
Geomod	The Geomod model had two versions Geomod1 and Geomod2, and it was integrated into IDRISI. Geomod can be applied on continental, country, and local scale and it has been used to study and analyze land use land cover changes specially scenarios of deforestation as it simulates the spatial patterns of changes through time[55].

Table.4 Cellular Approach models example

Model name	Application	Integration with other	References
------------	-------------	------------------------	------------

		models	
IDRISI's CA-MARKOV	to predict and simulate the future land use and land cover changes	-	[56]
	to simulating urban expansion/growth	-	[57]
	To simulate developments in future land use that will affect the future of wild ungulates	-	[58].
	to produce maps projecting spatial and temporal changes of wetland	-	[51]
	the study and assessment of the biodiversity	-	[59]
	assist in the process of land resources management	-	[60]
	To identify changes to land use and climate that may eventually affect the water resources	climate change model	[61]
Clue	to predict land use changes in the future & the simulation of land cover changes but at a regional scale & to study rapid changes in phenomenon such as agriculture, urban growth, and trade sector at regional scale & land use transitions & farmland abandonment, deforestation, and carbon sequestration	-	[32], [62]–[66]
Sleuth	to study urban growth or development and to project future scenarios	-	[53], [67]–[78]
	to help in the process of assessing flood hazards	-	[79]
Geomod	to study and assess deforestation and to predict future	-	[80]
	to predict how one land state change from one state to another	-	[81]
	to assess and predict the changes of urban expansion scenarios and how it relates to floods and droughts	-	[82]
	predicting deforestation in the future	-	[83]

Table.5 Cellular Approach applications

4. Future prospects & trends

The Current and future direction all is focused on land change science and advancements in how to better detect past changes and how to accurately predict the future changes as close to reality as possible. The next few points will discuss future research trends.

- A major topic in advancing land change modeling is concerned with the development of the land change models validation, evaluation, and calibration practices, as they help both the model accuracy and advancement also they increase the degree of how well accepted the model to the user. More research must be conducted to further improve calibration practices specially the manual calibration and expert knowledge needs more attention also the way calibration is conducted in applications must be more defined[11], [16], [92].
- A number of researchers tried to facilitate the process of selecting the appropriate model for a given research application by assessing various types of LUCC models. Although stating the pros and cons of each land change model made it a bit easier to select the suitable model for a certain research, it is still pretty much foggy. To help researcher's choice the most suitable model a certain process needs to be defined or a set of steps to ease and improve the quality of the application since it relays on the chosen model[12], [93], [94]. This also can be done with help of expert systems or by using a multi-criteria decision making method to help make the decision.
- Assessing a geographic phenomenon using land use/cover change techniques, Due to the limited amount of studies that focused on comparing various land use/cover change models to asses change [95] upcoming studies and research should be focused on such subject.
- Focusing on local scale land change models and coming up with different ideas to manage to transmit them to global scale can help advancing land change modeling more and better results may appear from such models on a wider scale [47]
- The amount of remote sensing datasets and images used to run some land change models requires new computing technique and new approach deal and manage large datasets especially high resolution ones that can be large in both size and number of files and will consume more processing time to create a more accurate result[50]. Big data techniques can be integrated to manage and run such large datasets, some articles written about using big data to higher the performance of a land change model but more studies needs to be performed in the same topic[47][96].
- Datasets and sources that are used as input for land change models also effects the quality of results and also requires attention to enhance the quality of such datasets in terms of scale and temporal aspects, also the way those datasets are processed and managed before entered to the land change model needs more attention spatially while classifying images and selecting the best algorithms that suites the purpose of the research[97].

5. Conclusion

This paper has reviewed the current practices, applications, and future prospects for land change modeling, although plenty approaches are available each one meets a certain set of objectives. The future of land change modeling is in the integration of its models with other techniques to better understood and simulate changes.

References

- [1] J. Jokar Arsanjani, *Dynamic land use/cover change modelling*. 2012.
- [2] M. J. . D. and A. . V. Peter H . Verburg , Paul P . Schot, "Land use change modelling : current practice and research priorities," *GeoJournal*, vol. 61, no. 4, pp. 309–324, 2004.

- [3] C. Agarwal, G. L. Green, M. Grove, T. Evans, and C. Schweik, "A Review and Assessment of Land-Use Change Models : Dynamics of Space , Time , and Human Choice Chetan Agarwal," *Apollo Int. Mag. Art Antiq.*, vol. 1, pp. 812–855, 2002.
- [4] J. Li and X. Yang, "Monitoring and Modeling of Global Changes: A Geomatics Perspective," pp. 265–279, 2015.
- [5] E. Koomen, J. Stillwell, A. Bakema, and H. J. Scholten, *Modelling Land-Use Change: Progress and Applications*, vol. 90, no. 2. 2007.
- [6] P. H. Verburg, P. P. Schot, M. J. Dijst, and A. Veldkamp, "Land use change modelling: Current practice and research priorities," *GeoJournal*, vol. 61, no. 4, pp. 309–324, 2004.
- [7] J. Li and X. Yang, *Jonathan Li Xiaojun Yang*. .
- [8] D. G. Brown, R. Walker, S. Manson, and K. Seto, "Land Change Science," vol. 6, pp. 403–417, 2004.
- [9] J. Mas, A. P. Vega, K. Clarke, and D. México, "Assessing ' spatially explicit ' land use / cover change models Centro de Investigaciones en Geografía Ambiental - Universidad Nacional Autónoma Department of Geography – University of California - Santa Barbara , USA," *Landsc. Ecol.*, pp. 508–513, 2010.
- [10] D. G. Brown *et al.*, *Advancing Land Change Modeling*. 2014.
- [11] D. G. Brown *et al.*, "Opportunities to improve impact, integration, and evaluation of land change models," *Curr. Opin. Environ. Sustain.*, vol. 5, no. 5, pp. 452–457, 2013.
- [12] D. G. Brown, P. H. Verburg, R. G. Pontius, and M. D. Lange, "Opportunities to improve impact, integration, and evaluation of land change models," *Curr. Opin. Environ. Sustain.*, vol. 5, no. 5, pp. 452–457, 2013.
- [13] J.-F. Mas, M. Kolb, M. Paegelow, M. C. Olmedo, and T. Houet, "Modelling Land use / cover changes : a comparison of conceptual approaches and softwares," *Environ. Model. Software*, vol. 51, pp. 94–111, 2014.
- [14] J. Malczewski, "GIS-based land-use suitability analysis: A critical overview," *Prog. Plann.*, vol. 62, no. 1, pp. 3–65, 2004.
- [15] Committee on Needs and Research Requirements for Land Change Modeling, Geographical Sciences Committee, Board on Earth Sciences and Resources, Division on Earth and Life Studies, and National Research Council, *Advancing Land Change Modeling: Opportunities and Research*. 2014.
- [16] C. on N. and R. R. for L. C. M. G. S. C. B. on E. S. and R. D. on E. and L. S. N. R. Council, *Advancing Land Change Modeling: Opportunities and Research Requirements*. .
- [17] J. van V. and P. Verburg, "Land Use Modelling," 2015.
- [18] C. on N. and R. R. for L. C. M. G. S. C. B. on E. S. and R. D. on E. and L. S. N. R. Council, *Advancing Land Change Modeling: Opportunities and Research Requirements*. 2014.
- [19] T. GUO, "Dinamica_EGO_guidebook.pdf," Jan. 2015.
- [20] L. C. Modeler, "Geospatial Monitoring and Modeling System Habitat and Biodiversity Modeler Ecosystem Services Modeler Earth Trends Modeler Climate Change Adaptation Modeler TerrSet Geospatial Monitoring and Modeling System," pp. 1–8.
- [21] B. C. Pijanowski, M. M. Building, D. Hyndman, E. Lansing, and M. M. Building, "The application of the land transformation, groundwater flow and solute transport models for michigan's grand traverse bay watershed," pp. 1–12.

- [22] B. C. PIJANOWSKI, K. T. ALEXANDRIDIS, and D. MULLUERE, “Modelling urbanization patterns in two diverse regions of the world,” vol. 1, no. December, 2006.
- [23] T. Centro, S. Remoto, U. Federal, and D. M. Gerais, “Modeling Environmental Dynamics with Dinamica EGO : Training Session in PORTUGAL How Dinamica EGO works ?”
- [24] A. B. Soares-filho *et al.*, “Role of Brazilian Amazon protected areas in climate change mitigation Role of Brazilian Amazon change mitigation protected areas in climate,” 2014.
- [25] K. M. Carlson, L. M. Curran, D. Ratnasari, A. M. Pittman, and B. S. Soares-filho, “Committed carbon emissions , deforestation , and community land conversion from oil palm plantation expansion in West Kalimantan , Indonesia,” vol. 109, no. 19, 2012.
- [26] H. T. L. Huong and A. Pathirana, “Urbanization and climate change impacts on future urban flooding in Can Tho city , Vietnam,” pp. 379–394, 2013.
- [27] S. Berbero, A. Akin, and K. C. Clarke, “Landscape and Urban Planning Cellular automata modeling approaches to forecast urban growth for adana , Turkey : A comparative approach,” vol. 153, pp. 11–27, 2016.
- [28] F. Hajek, M. J. Ventresca, J. Scriven, and A. Castro, “Regime-building for REDD + : Evidence from a cluster of local initiatives in south-eastern Peru,” *Environ. Sci. Policy*, vol. 14, no. 2, pp. 201–215, 2011.
- [29] R. Almeida *et al.*, “Simulating fire regimes in the Amazon in response to climate change and deforestation Michael Coe , Hermann Rodrigues and Renato Assunção Published by : Wiley Stable URL : <http://www.jstor.org/stable/23023102> Accessed : 09-08-2016 21 : 17 UTC Your use of ,” vol. 21, no. 5, pp. 1573–1590, 2016.
- [30] M. T. Coe *et al.*, “Deforestation and climate feedbacks threaten the ecological integrity of south – southeastern Amazonia,” no. figure 1, 2013.
- [31] E. E. Maeda, P. K. E. Pellikka, M. Siljander, and B. J. F. Clark, “Geomorphology Potential impacts of agricultural expansion and climate change on soil erosion in the Eastern Arc Mountains of Kenya,” *Geomorphology*, vol. 123, no. 3–4, pp. 279–289, 2010.
- [32] J. Mas, M. Kolb, M. Paegelow, M. Teresa, C. Olmedo, and T. Houet, “Environmental Modelling & Software Inductive pattern-based land use / cover change models : A comparison of four software packages,” *Environ. Model. Softw.*, vol. 51, pp. 94–111, 2014.
- [33] J. J. Schulz, L. Cayuela, C. Echeverria, and J. Salas, “Monitoring land cover change of the dryland forest landscape of Central Chile (1975 – 2008),” vol. 30, pp. 436–447, 2010.
- [34] M. Leh, S. Bajwa, and I. Chaubey, “IMPACT OF LAND USE CHANGE ON EROSION RISK : AN INTEGRATED REMOTE SENSING , GEOGRAPHIC INFORMATION SYSTEM AND MODELING METHODOLOGY,” 2011.
- [35] F. V Faleiro, R. B. Machado, and R. D. Loyola, “Defining spatial conservation priorities in the face of land-use and climate change,” *Biol. Conserv.*, vol. 158, pp. 248–257, 2013.
- [36] M. G. Tewolde, 1, * and Pedro Cabral, and 2, “Urban Sprawl Analysis and Modeling in Asmara, Eritrea,” *Remote Sens.*, pp. 2148–2165, 2011.
- [37] G. I. Díaz, L. Nahuelhual, C. Echeverría, and S. Marín, “Landscape and Urban Planning Drivers of land abandonment in Southern Chile and implications for landscape planning,” *Landsc. Urban Plan.*, vol. 99, no. 3–4, pp. 207–217, 2011.
- [38] C. O. Wilson and Q. Weng, “Science of the Total Environment Simulating the impacts of future land use and climate changes on surface water quality in the Des Plaines River watershed ,

- Chicago Metropolitan Statistical Area , Illinois,” *Sci. Total Environ.*, vol. 409, no. 20, pp. 4387–4405, 2011.
- [39] D. Geneletti, “Assessing the impact of alternative land-use zoning policies on future ecosystem services,” *Environ. Impact Assess. Rev.*, vol. 40, pp. 25–35, 2013.
- [40] A. Pérez-Vega, J. F. Mas, and A. Ligmann-Zielinska, “Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest,” *Environ. Model. Softw.*, vol. 29, no. 1, pp. 11–23, 2012.
- [41] A. Carmona, L. Nahuelhual, C. Echeverría, and A. Báez, “Linking farming systems to landscape change : An empirical and spatially explicit study in southern Chile,” *Agriculture, Ecosyst. Environ.*, vol. 139, no. 1–2, pp. 40–50, 2010.
- [42] L. Nahuelhual, A. Carmona, A. Lara, C. Echeverría, and M. E. González, “Landscape and Urban Planning Land-cover change to forest plantations : Proximate causes and implications for the landscape in south-central Chile,” *Landsc. Urban Plan.*, vol. 107, no. 1, pp. 12–20, 2012.
- [43] K. S. Kumar, P. U. Bhaskar, and K. Padmakumari, “APPLICATION OF LAND CHANGE MODELER FOR PREDICTION OF FUTURE LAND USE LAND COVER A CASE STUDY OF VIJAYAWADA CITY,” pp. 2571–2581.
- [44] Y. Megahed, “LAND COVER MAPPING ANALYSIS AND URBAN GROWTH MODELING USING REMOTE SENSING TECHNIQUES Case Study : Greater Cairo Region - Egypt Case Study : Greater Cairo Region - Egypt.”
- [45] R. B. Thapa and Y. Murayama, “Computers , Environment and Urban Systems Urban growth modeling of Kathmandu metropolitan region , Nepal,” *Comput. Environ. Urban Syst.*, vol. 35, no. 1, pp. 25–34, 2011.
- [46] M. I. Mahmoud, A. Duker, C. Conrad, M. Thiel, and H. S. Ahmad, “Analysis of settlement expansion and urban growth modelling using geoinformation for assessing potential impacts of urbanization on climate in Abuja City, Nigeria,” *Remote Sens.*, vol. 8, no. 3, 2016.
- [47] B. C. Pijanowski, A. Tayyebi, J. Doucette, B. K. Pekin, D. Braun, and J. Plourde, “A big data urban growth simulation at a national scale: Configuring the GIS and neural network based Land Transformation Model to run in a High Performance Computing (HPC) environment,” *Environ. Model. Softw.*, vol. 51, pp. 250–268, 2014.
- [48] S. Park, S. Jeon, S. Kim, and C. Choi, “Landscape and Urban Planning Prediction and comparison of urban growth by land suitability index mapping using GIS and RS in South Korea,” *Landsc. Urban Plan.*, vol. 99, no. 2, pp. 104–114, 2011.
- [49] U. Projection, “Land Transformation Model,” pp. 11–20, 2000.
- [50] G. Newman, J. Lee, and P. Berke, “Using the land transformation model to forecast vacant land,” *J. Land Use Sci.*, vol. 4248, no. March, pp. 1–26, 2016.
- [51] V. M. Bacani, A. Y. Sakamoto, H. Quénol, C. Vannier, and S. Corgne, “Markov chains–cellular automata modeling and multicriteria analysis of land cover change in the Lower Nhecolândia subregion of the Brazilian Pantanal wetland,” *J. Appl. Remote Sens.*, vol. 10, no. 1, p. 016004, 2016.
- [52] P. Verburg, “The CLUE model,” no. Hands-on exercises, p. 53, 2010.
- [53] X. Li and P. Gong, “Urban growth models: progress and perspective,” *Sci. Bull.*, vol. 61, no. 21, pp. 1637–1650, 2016.
- [54] K. C. Clarke, “Cellular Automata and Agent-Based Models,” pp. 1217–1233.

- [55] R. G. Pontius Jr. and H. Chen, "GEOMOD Modeling," *Clark Labs*, no. January 2006, pp. 1–44, 2006.
- [56] H. Mahmoud and P. Divigalpitiya, "Modeling Future Land Use and Land-Cover Change in the Asyut Region Using Markov Chains and Cellular Automata," *Smart Sustain. Plan. Cities Reg.*, pp. 1–435, 2017.
- [57] M. M. Aburas, Y. M. Ho, M. F. Ramli, and Z. H. Ash'aari, "Improving the capability of an integrated CA-Markov model to simulate spatio-temporal urban growth trends using an Analytical Hierarchy Process and Frequency Ratio," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 59, pp. 65–78, 2017.
- [58] P. Acevedo, M. Á. Farfán, A. L. Márquez, M. Delibes-Mateos, R. Real, and J. M. Vargas, "Past, present and future of wild ungulates in relation to changes in land use," *Landsc. Ecol.*, vol. 26, no. 1, pp. 19–31, 2011.
- [59] P. Mondal and J. Southworth, "Evaluation of conservation interventions using a cellular automata-Markov model," *For. Ecol. Manage.*, vol. 260, no. 10, pp. 1716–1725, 2010.
- [60] S. S. Palmate, A. Pandey, and S. K. Mishra, "Modelling spatiotemporal land dynamics for a trans-boundary river basin using integrated Cellular Automata and Markov Chain approach," *Appl. Geogr.*, vol. 82, pp. 11–23, 2017.
- [61] S. T. Y. Tong, Y. Sun, T. Ranatunga, J. He, and Y. J. Yang, "Predicting plausible impacts of sets of climate and land use change scenarios on water resources," *Appl. Geogr.*, vol. 32, no. 2, pp. 477–489, 2012.
- [62] A. Tayyebi, B. K. Pekin, B. C. Pijanowski, D. James, J. S. Doucette, and D. Braun, "Hierarchical modeling of urban growth across the conterminous USA : developing meso-scale quantity drivers for the Land Transformation Model," *J. Land Use Sci.*, no. February 2015, pp. 37–41, 2012.
- [63] L. D. A. Campania, S. Pindozi, E. Cervelli, P. F. Recchi, and A. Capolupo, "Predicting land use change on a broad area: Dyna-CLUE model application to the Litorale Domizio-Agro Aversano (Campania, South Italy)," vol. XLVIII, pp. 27–35, 2017.
- [64] J. Fox, J. B. Vogler, O. L. Sen, T. W. Giambelluca, and A. D. Ziegler, "Simulating land-cover change in Montane mainland southeast Asia," *Environ. Manage.*, vol. 49, no. 5, pp. 968–979, 2012.
- [65] G. Luo, C. Yin, X. Chen, W. Xu, and L. Lu, "Combining system dynamic model and CLUE-S model to improve land use scenario analyses at regional scale: A case study of Sangong watershed in Xinjiang, China," *Ecol. Complex.*, vol. 7, no. 2, pp. 198–207, 2010.
- [66] E. E. Maeda, C. M. de Almeida, A. de Carvalho Ximenes A., A. R. Formaggio, Y. E. Shimabukuro, and P. Pellikka, "Dynamic modeling of forest conversion: Simulation of past and future scenarios of rural activities expansion in the fringes of the Xingu National Park, Brazilian Amazon," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 13, no. 3, pp. 435–446, 2011.
- [67] F. Sunar and S. Berbero, "Urban change analysis and future growth of Istanbul," no. April, 2015.
- [68] A. A. A. Alsharif and B. Pradhan, "Urban Sprawl Analysis of Tripoli Metropolitan City (Libya) Using Remote Sensing Data and Multivariate Logistic Regression Model," vol. 42, no. March, pp. 149–163, 2014.
- [69] A. J. Terando, J. Costanza, C. Belyea, R. R. Dunn, A. McKerrow, and J. A. Collazo, "The southern megalopolis: Using the past to predict the future of urban sprawl in the Southeast U.S.," *PLoS One*, vol. 9, no. 7, 2014.

- [70] H. Sha and M. Helbich, "Spatiotemporal urbanization processes in the megacity of Mumbai , India : A Markov chains-cellular automata urban growth model," vol. 40, pp. 140–149, 2013.
- [71] M. A. L. Billa, "Modelling urban growth evolution and land-use changes using GIS based cellular automata and SLEUTH models : the case of Sana ' a metropolitan city , Yemen," 2012.
- [72] D. Triantakonstantis, "Urban growth prediction: a review of computational models and human perceptions," *J. Geogr. Inf. Syst.*, vol. 04, no. December, pp. 555–587, 2012.
- [73] C. R. Jacobson, "Identification and quantification of the hydrological impacts of imperviousness in urban catchments: A review," *J. Environ. Manage.*, vol. 92, no. 6, pp. 1438–1448, 2011.
- [74] S. Maithani, "Application of cellular automata and GIS techniques in urban growth modelling: A new perspective," *Inst. T. Planners, India J.*, vol. 7–1, no. 1, pp. 36–49, 2010.
- [75] M. L. Mann *et al.*, "Modeling residential development in California from 2000 to 2050: Integrating wildfire risk, wildland and agricultural encroachment," *Land use policy*, vol. 41, no. June 2009, pp. 438–452, 2014.
- [76] A. D. Syphard, K. C. Clarke, J. Franklin, H. M. Regan, and M. McGinnis, "Forecasts of habitat loss and fragmentation due to urban growth are sensitive to source of input data," *J. Environ. Manage.*, vol. 92, no. 7, pp. 1882–1893, 2011.
- [77] I. Sekovski, C. Armaroli, L. Calabrese, F. Mancini, F. Stecchi, and L. Perini, "Coupling scenarios of urban growth and flood hazards along the Emilia-Romagna coast (Italy)," *Nat. Hazards Earth Syst. Sci.*, vol. 15, no. 10, pp. 2331–2346, 2015.
- [78] V. Srinivasan, K. C. Seto, R. Emerson, and S. M. Gorelick, "The impact of urbanization on water vulnerability: A coupled human-environment system approach for Chennai, India," *Glob. Environ. Chang.*, vol. 23, no. 1, pp. 229–239, 2013.
- [79] E. O. Neill, M. Brennan, F. Brereton, and H. Shahumyan, "Exploring a spatial statistical approach to quantify flood risk perception using cognitive maps," pp. 1573–1601, 2015.
- [80] M. K. Rashmi and N. Lele, "Spatial modeling and validation of forest cover change in Kanakapura region using GEOMOD," *J. Indian Soc. Remote Sens.*, vol. 38, no. 1, pp. 45–54, 2010.
- [81] R. C. Estoque and Y. Murayama, "Quantifying landscape pattern and ecosystem service value changes in four rapidly urbanizing hill stations of Southeast Asia," *Landsc. Ecol.*, vol. 31, no. 7, pp. 1481–1507, 2016.
- [82] B. Guneralp, I. Guneralp, and Y. Liu, "Changing global patterns of urban exposure to flood and drought hazards," *Glob. Environ. Chang.*, vol. 31, pp. 217–225, 2015.
- [83] O. S. Kim, "An Assessment of Deforestation Models for Reducing Emissions from Deforestation and Forest Degradation (REDD)," *Trans. GIS*, vol. 14, no. 5, pp. 631–654, 2010.
- [84] M. Brady and E. Irwin, "Accounting for Spatial Effects in Economic Models of Land Use : Recent Developments and Challenges Ahead," pp. 487–509, 2011.
- [85] A. J. Plantinga and D. J. Lewis, "Landscape simulations with econometric-based land-use models," *Oxford Handb. L. Econ.*, pp. 380–401, 2014.
- [86] H. A. Klaiber and N. V Kuminoff, "Equilibrium Sorting Models of Land Use and Residential Choice," *Oxford Handb. L. Econ.*, 2014.
- [87] R. B. Matthews, N. G. Gilbert, A. Roach, J. G. Polhill, and N. M. Gotts, "Agent-based land-use models: a review of applications," *Landsc. Ecol.*, vol. 22, pp. 1447–1459, 2007.
- [88] R. G. Pontius, D. Huffaker, and K. Denman, "Useful techniques of validation for spatially explicit land-change models," *Ecol. Modell.*, vol. 179, no. 4, pp. 445–461, 2004.

- [89] T. Filatova, P. H. Verburg, D. C. Parker, and C. A. Stannard, "Spatial agent-based models for socio-ecological systems: Challenges and prospects," *Environ. Model. Softw.*, vol. 45, pp. 1–7, 2013.
- [90] A. R. Bell, D. T. Robinson, A. Malik, and S. Dewal, "Modular ABM development for improved dissemination and training," *Environ. Model. Softw.*, vol. 73, pp. 189–200, 2015.
- [91] J. J. Arsanjani, *Dynamic Land-Use/Cover Change Simulation: Geosimulation and Multi Agent-Based Modelling*. 2012.
- [92] J. van Vliet, A. K. Bregt, D. G. Brown, H. van Delden, S. Heckbert, and P. H. Verburg, "A review of current calibration and validation practices in land-change modeling," *Environ. Model. Softw.*, vol. 82, pp. 174–182, 2016.
- [93] H. Geist, W. McConnell, E. F. Lambin, E. Moran, D. Alves, and T. Rudel, *Land-Use and Land-Cover Change Local Processes and Global Impacts*. 2006.
- [94] R. J. Kavanagh *et al.*, "Endocrine disruption and altered gonadal development in white perch (*Moreno americana*) from the lower Great Lakes region," *Environ. Health Perspect.*, vol. 112, no. 8, pp. 898–902, 2004.
- [95] A. Pérez-Vega, J. F. Mas, and A. Ligmann-Zielinska, "Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest," *Environ. Model. Softw.*, vol. 29, no. 1, pp. 11–23, 2012.
- [96] L. An *et al.*, "Space–Time Analysis: Concepts, Quantitative Methods, and Future Directions," *Ann. Assoc. Am. Geogr.*, vol. 105, no. 5, pp. 891–914, 2015.
- [97] C. Paper *et al.*, "Land Cover projection based on Chain Markov and Cellular Automata : Case study of Pampulha Land Cover projection based on Chain Markov and Cellular," no. November 2016, 2015.

Cairo University
Institute of Statistical Studies and Research

**The 52nd Annual Conference on Statistics, Computer
Sciences and Operations Research**

Operations Research

25-27 Dec. 2017

Index
Operations Research

1	Measuring the Performance of Airport Gates Using Data Envelopment Analysis Naglaa Ragaa Saeid, Baraka Ahmed Galal, Mohamed Ali, Ibrahim Raslan	1-14
2	A Distance to the Ideal Alternative Approach for Group Decision Making Mohamed F. El-Santawy	15-24
3	Joint Chance Constrained Programming with Dependent Parameters Nada Hafez, Afaf El-Dash, Nagwa Albehery	25-31

MEASURING THE PERFORMANCE OF AIRPORT GATES USING DATA ENVELOPMENT ANALYSIS

Naglaa Ragaa Saeid ⁽¹⁾, Baraka Ahmed Galal ⁽²⁾, Mohamed Ali ⁽³⁾, Ibrahim Raslan⁽⁴⁾

Abstract

Airports all over the world are becoming busier and many of them are facing capacity problems. The actual airport capacity strongly depends on the efficiency of the resource utilization. This paper focuses on analysis the performance of gates by data envelopment analysis method to determine the less efficiency gates, and introduce simple heuristic to improve the efficiency of gates.

Keywords: Airport Gate Assignment, Data envelopment analysis, Heuristic.

1- Introduction

The airport gate assignment problem (AGAP) is a critical and essential issue for daily airport operations planning. Airports today become much busier and more complicated than previous days. Aircraft on the ground requires all kinds of diverse services, like the aircraft need to be refueled, new passengers need to board; new supplies have to be put on board. The aircraft has to get cleaned. All the actions take place while the aircraft is standing at a gate (Lim & Wang 2005).

Growing flights congestion makes it necessary and compulsory to find ways to increase the airport operation efficiency. Research on airport gate assignment problem (AGAP) appears extremely significant on facilitate airlines to assess how many gates they should rent from airports to serve their own aircrafts.

Recently AGAP becomes one of core components in the field of airport resource management and naturally appeals the close concentration of current researchers. AGAP can be described as follows: Suppose an airline company owns the business of hosting a certain number of flights every day and to run the business smoothly it must lease a certain number of gates from an airport (Chun etc. 1999).

The main mechanism of flight-to-gate assignment is to assign aircrafts to suitable gates so that not only passengers can conveniently embark or disembark but also the airline companies can minimize the cost in the whole operational process. Efficient airport operation largely depends on how to gate aircrafts in a smooth flow of arriving and departing flights (Chendong Li 2005).

(1) Associated Professor of Operations Researches, Institute of Statistical Studies and Researches, Cairo University.

(2) Pre Master student in Operations Researches, Institute of Statistical Studies and Researches, Cairo University

(3) Pre Master student in Operations Researches, Institute of Statistical Studies and Researches, Cairo University

(4) Pre Master student in Operations Researches, Institute of Statistical Studies and Researches, Cairo University

In this paper, we use Data Envelopment Analysis (DEA) to measure the performance of gates in Cairo airport and then, a simple heuristic algorithm is suggested to raise the performance of airport gates. We present a real case problem about "Cairo airport gates".

2- Problem Definition

Before presenting our case study, we have to explain some definitions as follows:

Aircraft stand: A designated area on an apron intended to be used for parking an aircraft.

Runway: A defined rectangular area on a land aerodrome prepared for the landing and take-off of aircraft.

Taxiway: A defined path on a land aerodrome established for the taxiing of aircraft and intended to provide a link between one part of the aerodrome and another.

Movement area: That part of an aerodrome to be used for the take-off, landing and taxiing of aircraft, consisting of the maneuvering area and the apron(s).

Ground control: is responsible for the airport "movement" areas, as well as areas not released to the airlines or other users. This generally includes all taxiways, inactive runways, holding areas, and some transitional aprons or intersections where aircraft arrive.

After the aircraft is landed on the runway, the pilot tunes ground control for clearance to taxi to the gate. The ground control tells him the taxiway and the suitable gate according to the aircraft type. The International Civil Aviation Organization (ICAO) classified aircrafts A, B, C, D, E and F according to wingspan.

Table 2.1 classification of aircraft wingspan

Code Letter	Wing Span
A	Up to 15 m
B	15 m up to 24 m
C	24 m up to 36 m
D	36 m up to 52 m
E	52 m up to 65 m
F	65 m up to 80 m

The common aircraft types in the Cairo international airport terminal 3 are classified to narrow and wide wingspan aircraft. The narrow wingspan aircraft are A320, A321, B738 the wide wingspan aircraft are A333, A332, B772, B773 (A is the first letter of Airbus and B is first letter of Boeing).

Table2.2 aircraft wingspan

Aircraft Type	Wingspan (M)	Aircraft Type	Wingspan (M)
A320	34.09 m	A332	60.30 m
A321	34.09 m	A333	60.30 m
B738	35.79 m	B772	60.93 m
B773	60.93 m		

The Cairo international airport terminal No. 3 has 15 gates called: (F1, F2, F3, F5, F7, F8, F9, G1, G2, G3, G4, G5, G7, G8, G9). Every gate has specifications. Some gates can take narrow wingspan aircraft and another gate gets wide wingspan aircraft. For example: F1 can assign one aircraft from type C and F2 can assign one aircraft from type E...and so on.

Gates (F5, F7, G5, G8) can assign wide body aircraft or two narrow body aircraft to become (F5A, F5B, F7A, F7B, G5A, G5B, G8A, G8B).

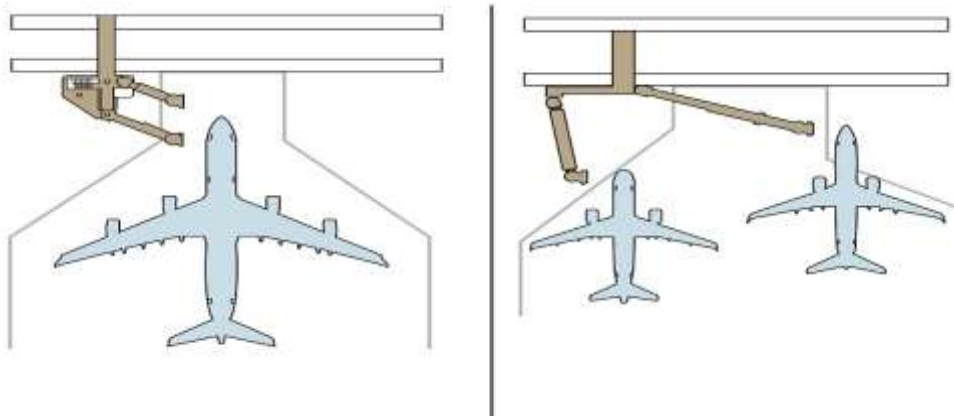


Figure2.1 Large gates can be used as two small gates by small aircraft

Applying DEA to measure performance of each gate aiming to improve its capability, starting from analyzing the process of the gate assignment and, Based on the real data, the gates assignment of Cairo International Airport (CIA) is analyzed to know: which gates have a maximum efficiencies and what are the efficiencies of other gates relative to the maximum efficiencies gates and a heuristic algorithm of AGAP is built to maximize the number of flights assigned to the gates.

2-1- The Objectives

The main objective is to measure the performance of airport gate through applying DEA. We also suggest a heuristic approach to raise the performance and the capability of these gates to improve the efficiency.

2-2- Methodology

The DEA is comparative approach for identify the performance. DEA is commonly used to evaluate the efficiency of a number of producers. A typical statistical approach is characterized as a central tendency approach and it evaluates producers relative to an average producer In contrast, DEA compares each

producer with only the "best" producers. By the way, in the DEA literature, a producer is usually referred to as a decision making unit or DMU (c.Ray, 2004).

DEA models will have two orientations: input-oriented and output-oriented, Input-oriented models are used to test if a DMU under evaluation can reduce its inputs while keeping the outputs at their current levels. Output-oriented models are used to test if a DMU under evaluation can increase its outputs while keeping the inputs at their current levels (Milan M, 2009).

In this paper, CRS (constant-returns to scale) output-oriented model is used because the output of the process increases or decreases simultaneously and in step with increase or decrease in the inputs.

Let x_{ij} - denote the observed magnitude of i - type input for gate j ($x_{ij} > 0, i = 1, 2, \dots, m, j = 1, 2, \dots, n$) and y_{rj} - the observed magnitude of r -type output for gate j ($y_{rj} > 0, r = 1, 2, \dots, s, j = 1, 2, \dots, n$). Then, the Charnes-Cooper-Rhodes (CCR) model is formulated in the following form for the selected gate k :

$$\text{Min } q = \sum_{i=1}^m v_i x_{ik}$$

subject to

$$\sum_{r=1}^s u_r y_{rk} = 1$$

$$\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} \geq 0, \quad (j = 1, 2, \dots, n)$$

$$u_r \geq \varepsilon, \quad (r = 1, 2, \dots, s)$$

$$v_i \geq \varepsilon, \quad (i = 1, 2, \dots, m)$$

Where:

- v_i is the weights to be determined for input i ;
- m is the number of inputs;
- u_r is the weights to be determined for output r ;
- s is the number of outputs;
- q is the relative efficiency of DMU $_k$;
- n is the number of entities;
- ε is a small positive value.

The variables selected for this research are most commonly used for input and output variables affecting airport gate efficiency as found in the published research. In our case we have one input and two outputs. We consider the design of each gate as input variable, the number of aircrafts and the time as output variables.

3- Literature Review

The multi-objective nature of the problem has been gradually acknowledged by researchers working on this problem. Initially research focused mainly on various passenger comfort objectives. For example, (Haghani, 1998) and (Xu, 2001) minimized the distance a passenger walks inside a terminal to reach a departure gate. (Yan, 2001) Introduced an integer programming model to minimize the total passenger walking distance and the total passenger waiting time. (Ding, 2004) Shifted the research interest slightly from passenger comfort to airport operations and solved a multi-objective IP formulation of the GAP with an additional objective: minimization of the number of ungated flights. (Dorndorf, 2007) Went further and claimed that not enough attention has been given to the airport side when solving the GAP. They focused on three objectives: maximization of the total flight to gate preferences, minimization of the number of towing moves and minimization of the absolute deviation from the original schedule. Perhaps surprisingly at the time, in the context of many previous publications, they omitted the walking distance objective, arguing that the airport managers do not consider it an important aspect of the GAP. Our experience with airports indicates that this is probably correct, which is why the objective function presented in Section 4 does not consider passenger walking distance but aims instead to reduce the conflicts by the gates, to allocate gates so that the airline and the size preferences are maximized and to ensure that the time gaps between two adjacent allocations are large enough. The constraints that we have modelled in Section 4 include other aspects of the problem which have already been discussed by other researchers.

It is discussed relative sizes of gates and aircraft, airline preferences and shadowing restrictions. Similarly the importance of the maximization of time gaps between allocations and the robustness of solutions was discussed by (Bolat, 2000). This is also included in our objective function. We aim to avoid small gaps by maximizing the time gaps within a defined time window. (Kumar, 2011) Included towing in their models. They modelled the towing using three flights: (arrival) + (tow away); (tow away) + (tow back); and (tow back) + (departure).

4- Experiment Analysis

Data about performance in Cairo airport was collected to analyze the efficiency and determine the improvement opportunity, our study involved 13 working days at Cairo international airport, terminal 3 that contains 14 gates.

We calculated the number of aircrafts (A/Cs) arriving in this particular period of time, we found them to be 1304 A/Cs, 974 A/Cs of them could be served at terminal 3 because there are types that can't be served in terminal 3 like E170.

Table 4.1 shows the number of planes served at gates in the study time.

Table 4.1 A/C serviced by all Gates

Total number of A/C	A/C serviced by Gates	Percentage
974	668	68.58%

The percentage of A/Cs that was served was 68.58% from the total number of A/Cs that can be served in the terminal, they were distributed over gates as follows in table 4.2

Table 4.2 A/C serviced by each Gate

Gate		Type		Total
		Dom	Int	
F1	Count	6	58	64
	%	9.4%	90.6%	100.0%
F2	Count	0	29	29
	%	0.0%	100.0%	100.0%
F3	Count	3	50	53
	%	5.7%	94.3%	100.0%
F5	Count	4	41	45
	%	8.9%	91.1%	100.0%
F7	Count	0	41	41
	%	0.0%	100.0%	100.0%
F8	Count	8	48	56
	%	14.3%	85.7%	100.0%
G1	Count	9	60	69
	%	13.0%	87.0%	100.0%
G2	Count	0	28	28
	%	0.0%	100.0%	100.0%
G3	Count	11	50	61
	%	18.0%	82.0%	100.0%
G4	Count	0	33	33
	%	0.0%	100.0%	100.0%
G5	Count	1	34	35
	%	2.9%	97.1%	100.0%
G7	Count	6	46	52
	%	11.5%	88.5%	100.0%
G8	Count	9	44	53
	%	17.0%	83.0%	100.0%
G9	Count	5	44	49
	%	10.2%	89.8%	100.0%
Total	Count	62	606	668
	%	9.3%	90.7%	100.0%

The hours in which each gate was in use during the period covered by the study in figure 4.1

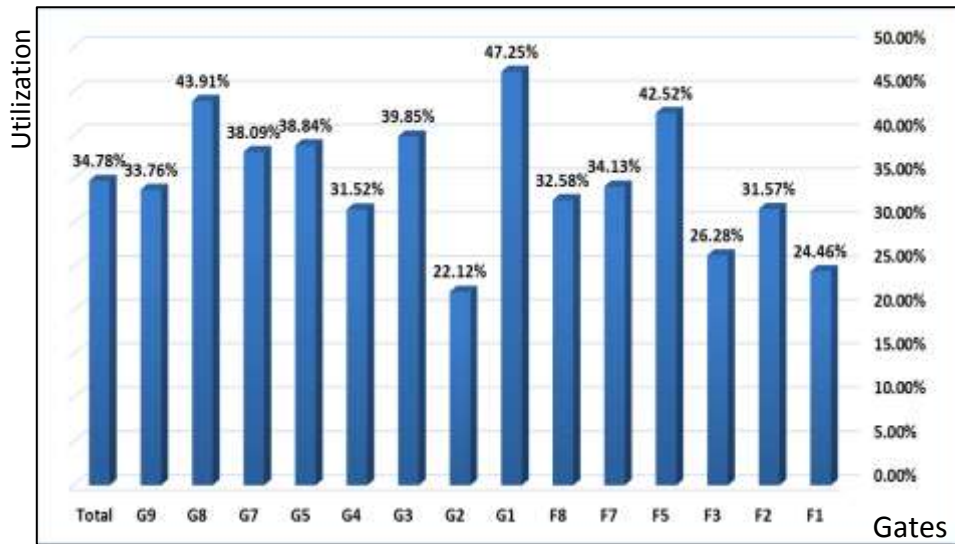


Figure 4.1 Time of use each gate

Utilization of gates was 22.12%-47.25% of Study time which are poor percentages indicating defects in management and operating the gates which requires improving their efficiency.

5- Calculate Efficiency of Gates Using DEA

Output-oriented DEA using CRS model was used for determining the overall technical efficiency for each gate in terminal 3. The model has the following components:

Input

- Width of gate which determine the allowable wingspan of aircraft.

Outputs

- Number of served A/Cs.
- Time in use for each gate.

Table 5.1 shows the values of input and outputs variables for each gate in terminal 3 within the time of study.

Table 5.1 the values of variables for each gate

Gates	Input	Output	
	Width	Number of A/C	Time in Use (hrs.)
F1	36	64	76.33
F2	65	29	98.50
F3	52	53	82.00
F5	80	45	132.67
F7	65	41	106.50
F8	52	56	101.66
G1	36	69	147.42
G2	65	28	87.58
G3	52	61	124.33
G4	65	33	98.93
G5	80	35	121.17
G7	52	52	118.83
G8	65	53	137.00
G9	36	49	105.33

Using WinDeap program and analyzing this data, the results of technical efficiency for each gate in terminal 3 were as follows in figure 5.1

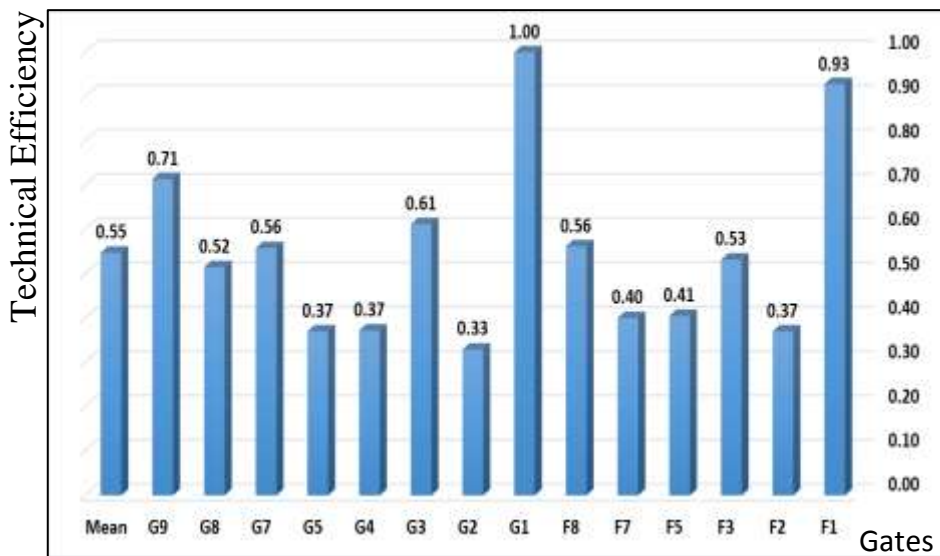


Figure 5.1 Technical efficiency of each gate

Figure 5.1 shows that technical efficiency of gates begins with 0.33, and shows that the mean of efficiency of terminal was 0.55.

It can be deduced that gate G1 is the most efficient, it can be considered as a "peer" to every other gate in the terminal, the Peers weights of each gate was demonstrated in table 5.2.

Table 5.2 Peers weights of each gate

Gate	Peers	peer weights
F1	G1	1.00
F2	G1	1.81
F3	G1	1.44
F5	G1	2.22
F7	G1	1.81
F8	G1	1.44
G1	G1	1.00
G2	G1	1.81
G3	G1	1.44
G4	G1	1.81
G5	G1	2.22
G7	G1	1.44
G8	G1	1.81
G9	G1	1.00

The improvement opportunity for each gate can be determined using DEA, the study choose the outputs orientation to improve the efficiency because it is available to change the values of outputs, and there is not available to change the

value of input of any gate, figure 5.2 shows the target number of A/Cs and the total time in use for each gate according to output-oriented DEA.

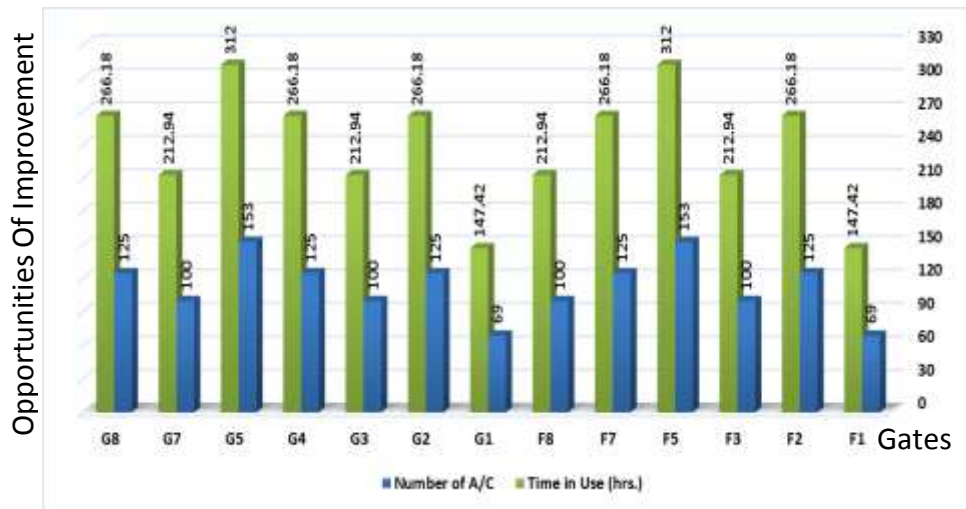


Figure 5.2 Summary of outputs target

Using DEA Shows that it is possible to improve the efficiency of the terminal and each gate, but DEA does not provide a way to achieve the target outputs, so the study proposes a simple heuristic algorithm for improve the efficiency of gate assignment.

6- Suggested heuristic algorithm

The word “heuristic” comes from the Greek word “eurisko”, which means “I find”. It has been used in the last century or so, to indicate a practical decision rule or a practical way to find a solution to a problem, relying upon experience and common sense.

These rules do not aim at satisfying any formal or theoretical property; sometimes “heuristics” have been defined just as practical alternatives in contrast to formal mathematical techniques.

Heuristic algorithms are used to solve large instances of computationally difficult problems, because the computation of an exact solution would require an excessive amount of computing time.

These algorithms can be roughly classified into two types:

- Specific algorithms for specific problems **Heuristics**
- General ideas for almost any problem **Meta-heuristics**
(Pearl, 1984)

6-1- Prioritization of gates

As a rule of thumb the gates will be divided into three types, first one for narrow wing span aircraft only, second for wide wing span aircraft only, and third is multi for both narrow and wide wing span aircraft.

1. If the narrow wing span aircraft arrive and all multi gates are not assigned; the priority for narrow wing span gate.
2. If the narrow wing span aircraft arrive and there is multi gate are assigned with narrow wing span aircraft; the priority for this gate.
3. If the wide wing span aircraft arrive; the priority for wide wing span gate only

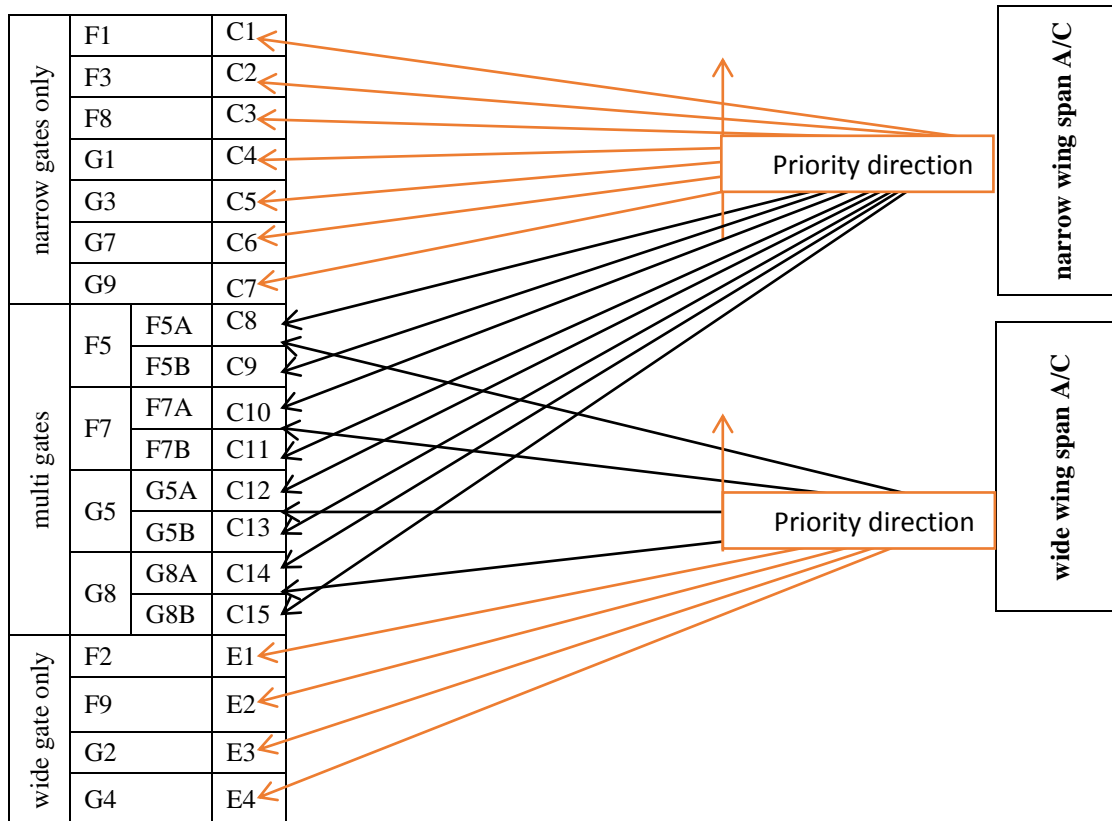


Figure 6-1-1 Prioritization of gates

6-2- Flow chart

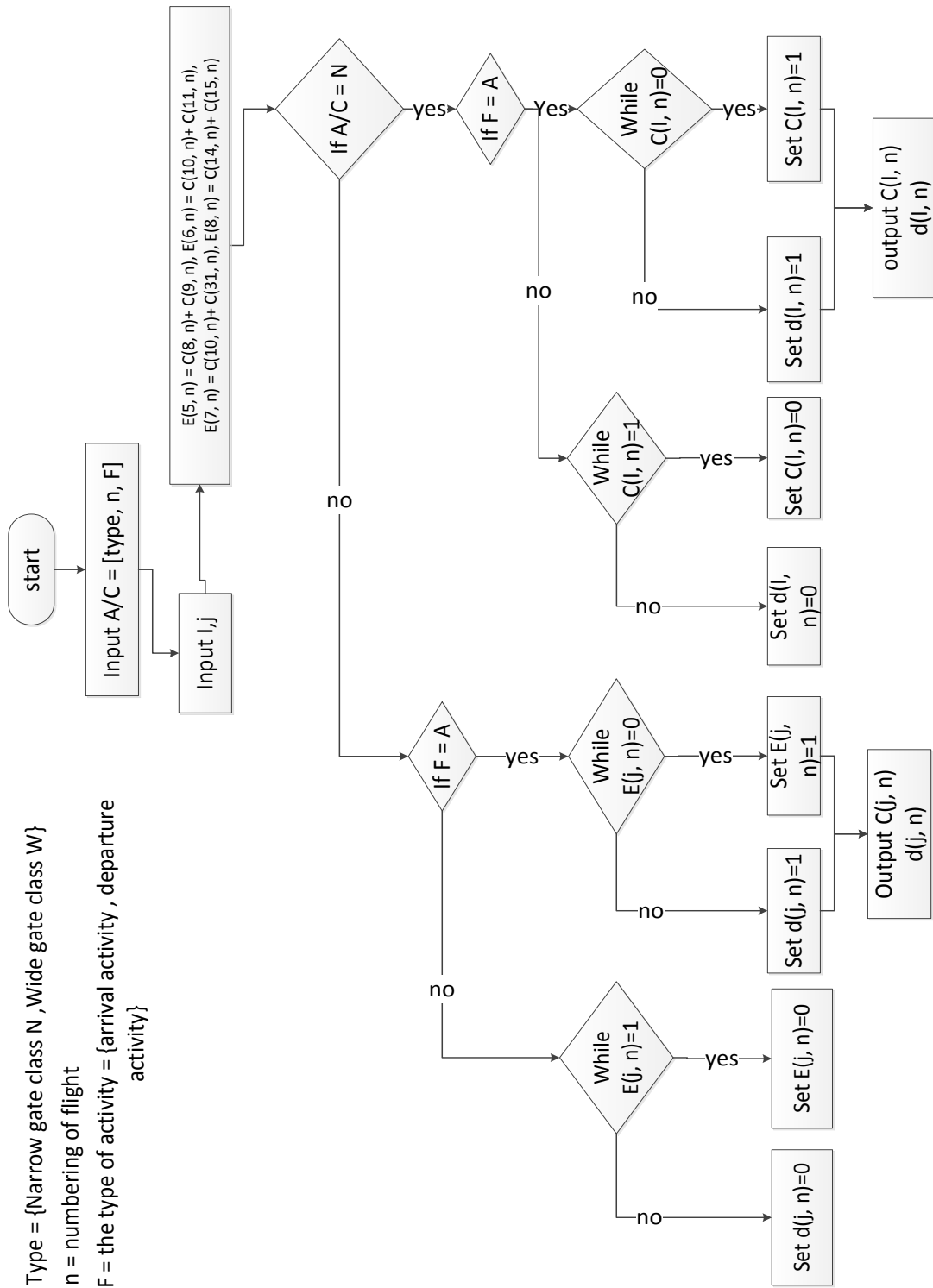


Figure 6-2 showing the flow chart of simple heuristic algorithm for assigning A/C in gates

6-3- Results of applying heuristic algorithm

This part will present the results for three randomly day from historical data. For each day the results of applying simple heuristic algorithm will be compared to the original gate assignment (before).

In figure 6-3-1 showing the served time in hour for each gate in first day before and after applying simple heuristic algorithm, the total served time increased 1.6 from the original assignment.

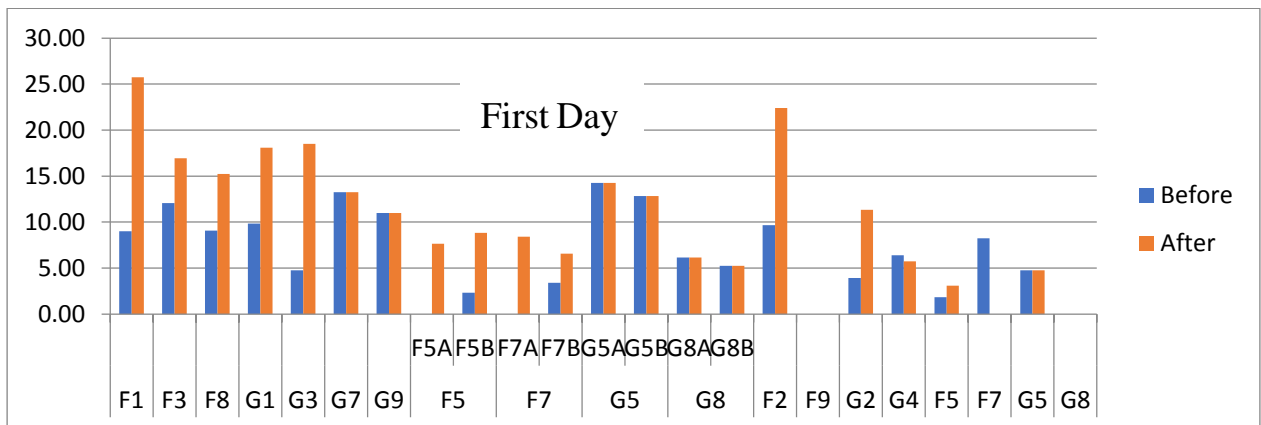


Figure 6-3-1 comparisons between served time in hour before and after applying algorithm in first day

In figure 6-3-2 showing the served time in hour for each gate in second day before and after applying simple heuristic algorithm, the total served time increased 1.97 from the original assignment

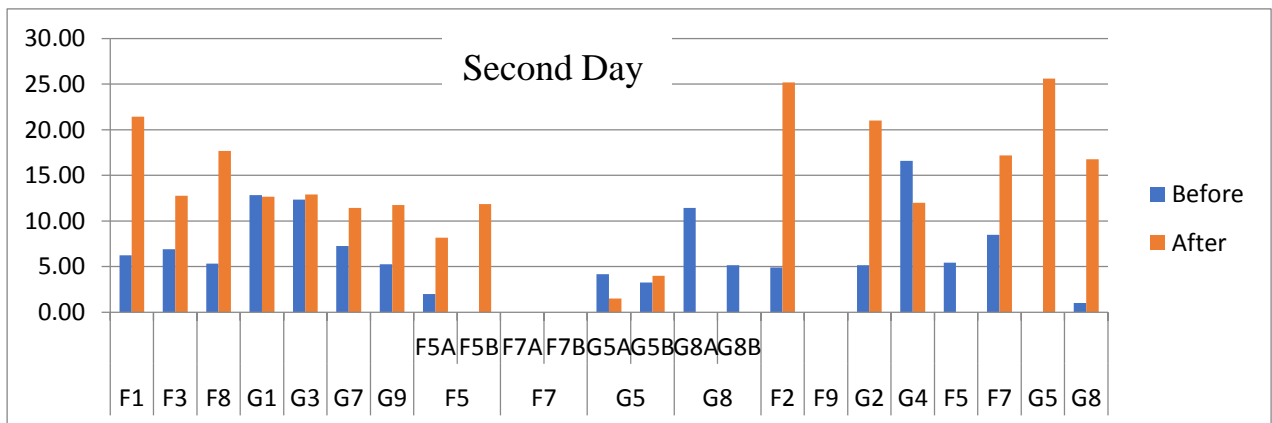


Figure 6-3-2 comparisons between served time in hour before and after applying algorithm in second day

In figure 6-3-3 showing the served time in hour for each gate in third day before and after applying simple heuristic algorithm, the total served time increased 1.96 from the original assignment

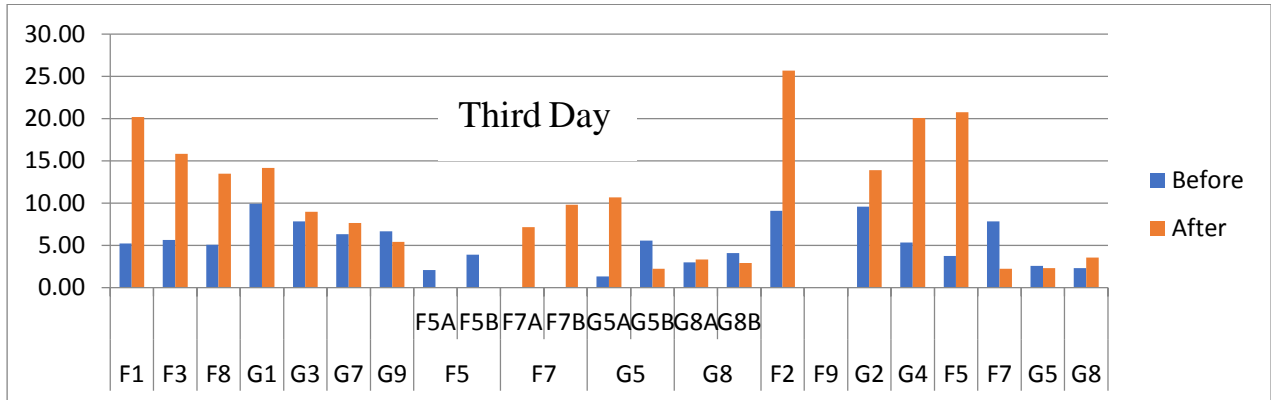


Figure 6-3-3 comparisons between served time in hour before and after applying algorithm in third day

7- Conclusion

This paper attempts to identify the key performance indicators to assess the efficiency of Cairo airport gates. Given sample of preschedule information about airport gates, in our case we have one input and two outputs, we consider the design of each gate as input variable, and the number of aircrafts and the using time of each gate as output variables (most commonly used in the published researches).

The case study involved 13 working days at Cairo international airport, terminal 3 that contains 14 gates.

We calculated the number of aircrafts (A/Cs) arriving in this particular period of time, we found them 1304 A/Cs, and 974 A/Cs of them could be served at terminal 3.

Utilization of gates located in the interval 22.12%-47.25% of study time, which are poor percentages, indicating defects in management and operating the gates, which requires improving their efficiency.

Using data envelopment analysis shows that the mean of efficiency of terminal was 0.55.

Finally, we suggested a simple heuristic algorithm to improve the efficiency of gate and increasing utilization, the total utilization of all gates in terminal 3 increased in three day of applying algorithm are 1.6, 1.97, and 1.96 of the actual assignment.

References

- A., Chen, M.C Haghani(1998) .*Optimizing gate assignments at airport terminals* .springer.
- Bolat (2000) .*Procedures for providing robust gate assignments for arriving aircraft* .European Journal of Operational Research.
- Subhash C.Ray (2004) .Data Envelopment Analysis: Theory and Techniques in economies and operations research, CAMBRIDGE Press, 2-14.
- Charnes A., Cooper W. and Rhodes E. (1978) Measuring the efficiency of decision making units, *Eur. J. Opl. Res* 2, 429-444.
- Chun A., Chan S., Tsang F., and Yeung D. Hkia sas (1999.) A constraint-based airport stand allocation system developed with software components, Proceeding of AAAI/IAAI 786–793.
- H., Lim & Rodrigues, B., Zhu, Y Ding ..(2004) *Aircraft and gate scheduling optimization at airport* .Annual Hawaii International Conference.
- J., Bailey, G .(2001) .*The airport gate assignment problem: Mathematical model* . Hawaii International Conference on System Sciences.
- Judea Pearl .(1984) .*heuristic intelligent search strategies for computer solving problem* .Los Angeles: ADDISON-WESLEY PUBLISHING COMPANY. 1-31
- Lim, A., and Wang, F. (2005). Robust Airport Gate Assignment. ICTAI 2005, 74-81.
- S., Huo and C.M Yan .(2001) .*Optimization of multiple objective gate assignments* .Transportation Research Part A.

A Distance to the Ideal Alternative Approach for Group Decision Making

Mohamed F. El-Santawy *

Abstract

In Group Decision Making, a compromise strategy is presented to compile the conflicting opinions of decision makers. Many parameters are relevant in solving this category of problems. The Distance to Ideal Alternative (DIA) technique is an efficient method to tackle such category because of its mechanism and robust performance. In this paper, a DIA approach is proposed to solve the Group Decision Making problems and to deal with the ranking abnormalities existed in the ranking orders produced by the Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) and Multi-Objective Optimization on the basis of Ratio Analysis (MOORA) methods. A Numerical example is adopted for testing the proposed approach.

Keywords: Distance to the Ideal Alternative; Group Decision Making; Multi-Criteria Decision Making; MOORA; TOPSIS.

1. Introduction

A group is defined by set of members gathered for an aim or specific goal. Daily there are a lot groups constructed for certain purposes whether the members of the group are collected in a specific place or time. In management, many examples can be mentioned; a company board of directors having set of actions to choose from, hiring a new employee, facility location to be determined, expansion in markets, and other examples beyond listing here. Many other areas like politics, economics, engineering, and other disciplines are applications for this type. The members of the group might be humans or others races but still there is an ultimate goal to be reached. Birds' swarms seeking their food in groups, fish flocks searching destinations, and insects going shelters are various examples for groups with ultimate collective goals. Group members are sharing information and discussions to reach this ultimate goal. Despite of the might be existence of conflicting interests among the members but still there must be a compromise to be done. Discussions are conformed by various types like voting, brainstorming, analysis to construct unique parameters for the whole group. This process of narrowing the conflicting points of view among group members is studied by many researchers. Some researchers involve some managerial techniques like Delphi technique and SWOT analysis for this aim.

* PhD in Operations Research, Financial Researcher at Ministry of Finance, Egypt
(E-mail: lost_zola@yahoo.com)

Every member of the group should present his interests, priorities, hazards, and conflicts clearly in advance before the group goal's construction. After, a procedure is done for reaching a decision or set of decisions to reach the ultimate group goal under the group individuals' parameters mentioned. Problems with many decision makers are complicated because it involves many opposing factors, such as: conflicting individual goals, inefficient knowledge, individual motivation, personal opinion, and power. In brief, MCDM is the art of compromising conflicting criteria for a single decision maker, while GDM is the art of compromising different opinions of group members [10]. In both Multi-Criteria Decision Making (MCDM) and Group Decision Making (GDM), there are two major steps: aggregation and exploitation [4]. Both steps are defined in a different manner relative to the method employed and application area [9]. In a sense, GDM can be a multi-dimensional MCDM problem, but more issues should be studied to treat the conflicting factors aforementioned.

The group members have different intentions, aspects, backgrounds, and preferences. Although they can be clustered into specific sets but still there must be a compromise should be done to reach satisfactory decision for all decision makers. In MCDM, many popular techniques were proposed since mid 1960s [5]. Many researchers contribute in MCDM field and after they extend their work to GDM. Over two decades TOPSIS method had explored many applications and problems [6]. TOPSIS suffers from a crucial drawback so-called abnormalities of ranking orders. Distance to Ideal Alternative (DIA) is a modified version stemmed out from TOPSIS method [8]. DIA outperforms TOPSIS concerning robustness and efficiency. Basically, the present paper is to extend DIA to problems with many decision makers and set relevant modifications and enhancements to the original method. Additional to this goal, the introduced approach get benefit from employing DIA to tackle the well known abnormalities problem. The validation tests show how superior is the proposed approach over TOPSIS and MOORA.

This paper is organized as follows: Section 2, "Group Decision Making Problem" which contains techniques that used in comparison to the introduced approach Section 3, "Solution Methods & Techniques" presenting quick review about "MOORA", "TOPSIS" and "DIA". Section 4 "The proposed method" is illustrated. In section 5, "a numerical example" is given for validation, and finally section 6 is made for "Conclusion".

2. Group Decision Making Problem

A problem with many decision makers involved can be seen from two major points of view. Firstly, set of independent MCDM models are developed based on individuals' preferences. For each decision maker there will be a separate MCDM problem from his point of view. Each decision maker is responsible to define the relevant parameters according to him and solves the separate problem to get a solution. Then, the independent solutions are compiled using specific operators to get the group solution. This can be

briefly described by compromising resulted solutions (ranks). Secondly, the aggregation is done to the individuals' preference, providing finally a set of group parameters.

For this resulted MCDM problem, it can be solved as a unique problem and the solution (rank) produced is for the whole group. Solving a group decision making problem begins with the construction of decision making matrices. Let $D = \{1,2, \dots,K\}$ a set of decision makers or experts. The Group Decision Making problem can be expressed in k-matrix format in the following way

$$\begin{array}{c|cccc}
 & C_1 & C_2 & \dots & C_n \\
 \hline
 A_1 & x_{11}^k & x_{12}^k & \dots & x_{1n}^k \\
 A_2 & x_{21}^k & x_{22}^k & \dots & x_{2n}^k \\
 \dots & \dots & \dots & \dots & \dots \\
 A_m & x_{m1}^k & x_{m2}^k & \dots & x_{mn}^k
 \end{array} \tag{1}$$

where:

- A_1, A_2, \dots, A_m are alternatives that decision makers have to choose from
- C_1, C_2, \dots, C_n are the criteria for which the alternative performance is measured
- x_{ij}^k is the k -decision maker rating of alternative A_i with respect to the criterion C_j
- for m alternatives and n criteria we have matrix $X^k = (x_{ij}^k)$ where x_{ij}^k is the performance rating of the i^{th} alternative A_i , with respect to the j^{th} criterion C_j for the k^{th} decision maker
- $i = 1,2,\dots,m, \quad j = 1,2,\dots,n, \quad \text{and } k = 1,2,\dots,K.$

The importance of each criterion is given by $W^k = [w_1^k, w_2^k, \dots, w_n^k]$ a weight vector for the

$$k^{th} \text{ decision maker, } \sum_{j=1}^n w_j^k = 1.$$

3. Solution Methods & Techniques

This section contains three well known MCDM techniques which will be illustrated for being used in solving GDM in next section. The three chosen methods are MOORA for being a famous aggregating method, TOPSIS, and the proposed DIA approach. Both TOPSIS and DIA are so similar in their early steps but both of them have a distinct philosophy in reaching a solution.

3.1 MOORA

As shown in Eq. (1) for a single decision maker, the procedure of MOORA for ranking alternatives can be described as following [2]:

Step 1: Normalized decision matrix is computed as in Eq. (2)

$$x_{ij}^* = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}}, \quad i = 1, \dots, m; j = 1, \dots, n. \quad (2)$$

Step 2: The composite score is calculated as illustrated in Eq. (3)

$$z_i = \sum_{j=1}^b x_{ij}^* - \sum_{j=b+1}^n x_{ij}^*, \quad i = 1, \dots, m. \quad (3)$$

where $\sum_{j=1}^b x_{ij}^*$ and $\sum_{j=b+1}^n x_{ij}^*$ are for the Max (benefit) and Min (cost) criteria, respectively.

For different weights of criteria, the composite score becomes

$$z_i = \sum_{j=1}^b w_j \cdot x_{ij}^* - \sum_{j=b+1}^n w_j \cdot x_{ij}^*, \quad i = 1, \dots, m. \quad (4)$$

where W_j is the weight of j^{th} criterion.

Step 3: The alternatives are ranked in descending order.

The simplicity of MOORA is the main reason of its widely application to various disciplines. The philosophy of aggregating score for alternatives makes the method robust and easy to use. Recently, MOORA has been explored many problems of various fields, such as contractor selection [2], and climate [7].

3.2 TOPSIS

TOPSIS method is relying on the idea of comparing set of alternatives to an ideal composite vector constructed from the best performance alternatives under all criteria. Being closer to this ideal vector is better alternatives. In the latest decades, TOPSIS have been extended according to the requirements of different real-world decision making problems; it has been extended to various optimization areas like robot selection [3], and operating systems [1]. For one decision maker a decision matrix is to be constructed as shown in Eq. (1), the TOPSIS solution method consists of the following steps [6]:

Step 1: *The decision matrix is normalized:* it can be calculated as in Eq. (2)

Step 2: *The normalized weighted decision is constructed:* The columns of the normalized decision matrix are multiplied by the associated weights as follows:

$$v_{ij} = w_j \cdot x_{ij}^* , \quad i = 1, \dots, m; \quad j = 1, \dots, n, \quad \text{and} \quad \sum_{j=1}^n w_j = 1 \quad (5)$$

Step 3: *The positive and negative ideal solutions are computed:* The positive and negative ideal value sets are determined, respectively, as follows:

$$A^+ = (v_1^+, v_2^+, \dots, v_n^+) = \{ \max v_{ij} \text{ if } j \in \Omega_b \} \\ \text{or } \{ \min v_{ij} \text{ if } j \in \Omega_c \} \quad (6)$$

$$A^- = (v_1^-, v_2^-, \dots, v_n^-) = \{ \min v_{ij} \text{ if } j \in \Omega_b \} \\ \text{or } \{ \max v_{ij} \text{ if } j \in \Omega_c \} \quad (7)$$

Where Ω_b is the set of benefit criteria and Ω_c is the set of cost criteria.

Step 4: *The distance from positive and negative ideal solutions are calculated:* Two Euclidean distances for each alternative are calculated as follows:

$$S_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2}, \quad i = 1, \dots, m \quad (8)$$

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}, \quad i = 1, \dots, m \quad (9)$$

Where S_i^+ and S_i^- represents the distance of alternative A_i from the positive and negative ideal solutions, respectively.

Step 5: *The relative closeness to the ideal solution is calculated :* The Relative Closeness (RC) of the i^{th} alternative to the ideal solution is defined as follows:

$$RC_i = \frac{S_i^-}{S_i^+ + S_i^-}, \quad i = 1, 2, \dots, m. \quad 0 \leq RC_i \leq 1 \quad (10)$$

Step 6: *The alternatives are ranked:* Alternatives must be ranked based on RC_i in which the highest score is the best alternative.

3.3 DIA

The distance to the Ideal Alternative (DIA) algorithm belonging to the MCDM was developed [8] to select dynamically the best network inter-face. It can be seen as an edited version from classical TOPSIS. The philosophy behind DIA is ranking the alternatives based on distances in-between them after taking a snap shot reflecting their locations from

positive and negative ideal vectors. The main advantage of DIA is the method's employment of the distances between alternatives and capturing a rank based on the alternatives' relative locations after shifting from the distances to ideal vectors like TOPSIS. This crucial modification done gives a comparative advantage for DIA over TOPSIS as being robust and overcoming the abnormalities of TOPSIS. After decision matrices construction for all decision makers as in Eq. (1), the DIA is identical to the TOPSIS method in the first three steps. In a sense the DIA enhances TOPSIS as shown in steps (4-6). DIA method consists of the following steps for a single decision maker problem [8]:

Steps (1-3): *The same as in TOPSIS method mentioned before in previous subsection*

Step 4: *Calculate the distance to the positive and negative criteria:* as shown in the following two equations

$$D_i^+ = \sqrt{\sum_{j=1}^n (v_j^+ - v_{ij}^-)^2}, \quad i = 1, 2, \dots, m \quad (11)$$

$$D_i^- = \sqrt{\sum_{j=1}^n (v_{ij}^+ - v_j^-)^2}, \quad i = 1, 2, \dots, m \quad (12)$$

Step 5: *Determine the Positive Ideal Alternative:* (PIA) has minimum D^+ and maximum D^- .

$$PIA = \{\min(D_i^+), \max(D_i^-)\}, \quad i = 1, 2, \dots, m \quad (13)$$

Step 6: *Calculate the distance of an alternative to the PIA:*

$$R_i = \sqrt{(D_i^+ - \min(D_i^+))^2 + (D_i^- - \max(D_i^-))^2} \quad (14)$$

A set of alternatives is ranked according to the increasing order of R_i

4. Proposed Approach

In the previous section DIA was illustrated for solving MCDM problem with single decision maker. However, the extension of DIA to tackle group decision making problems will differ from single decision maker version. Many aspects should be adopted when introducing a group decision making approach like how the approach will compile different opinions of decision makers, how the proposed approach will compromise different ranks originated. The main reason behind the allocation of DIA to be extended to group decision making is its outperformance to TOPSIS as well as its overcoming the drawback of abnormalities found in classical TOPSIS. Construct the decision matrices for the k -decision makers and determine their weights which reflect their priorities. $X^k = (x_{ij}^k)$,

$W^k = [w_1^k, w_2^k, \dots, w_n^k]$ a weight vector for the k^{th} decision maker, $\sum_{j=1}^n w_j^k = 1$, for $k =$

1, 2, ..., K.

Step 1: Construct the normalized decision matrix for each decision maker:

$$x_{ij}^{k*} = \frac{x_{ij}^k}{\sqrt{\sum_{i=1}^m (x_{ij}^k)^2}}, \quad i = 1, \dots, m; j = 1, \dots, n; k = 1, 2, \dots, K \quad (15)$$

Step 2: Construct the weighted normalized decision matrices for all decision makers:

$$v_{ij}^k = w_j^k \cdot x_{ij}^{k*}, \quad i = 1, \dots, m; j = 1, \dots, n, k = 1, 2, \dots, K \quad (16)$$

Step 3: Determine the positive and negative ideal solutions for all decision makers as following:

$$A^{k+} = (v_1^{k+}, v_2^{k+}, \dots, v_n^{k+}) = \{ \max v_{ij}^k \text{ if } j \in \Omega_b \} \\ \text{or } \{ \min v_{ij}^k \text{ if } j \in \Omega_c \} \quad (17)$$

$$A^{k-} = (v_1^{k-}, v_2^{k-}, \dots, v_n^{k-}) = \{ \min v_{ij}^k \text{ if } j \in \Omega_b \} \\ \text{or } \{ \max v_{ij}^k \text{ if } j \in \Omega_c \} \quad (18)$$

Where Ω_b is the set of benefit criteria and Ω_c is the set of cost criteria.

$i = 1, \dots, m; j = 1, \dots, n$, and $k = 1, 2, \dots, K$

Step 4: Calculate the distance to the positive and negative criteria for all alternatives under all decision makers: as shown in the Eqs. (19-22)

$$D_i^{k+} = \sqrt{\sum_{j=1}^n (v_j^{k+} - v_{ij}^k)^2}, \quad i = 1, 2, \dots, m \text{ and } k = 1, 2, \dots, K \quad (19)$$

$$D_i^{k-} = \sqrt{\sum_{j=1}^n (v_{ij}^k - v_j^{k-})^2}, \quad i = 1, 2, \dots, m \text{ and } k = 1, 2, \dots, K \quad (20)$$

$$D_i^+ = \frac{\sum_{k=1}^K D_i^{k+}}{K}, \quad i = 1, 2, \dots, m \quad (21)$$

$$D_i^- = \frac{\sum_{k=1}^K D_i^{k-}}{K}, \quad i = 1, 2, \dots, m \quad (22)$$

Step 5: Determine the Positive Ideal Alternative for the group:

$$PIA = \{ \min(D_i^+), \max(D_i^-) \}, \quad i = 1, 2, \dots, m \quad (23)$$

Step 6: Calculate the distance of an alternative to the PIA for the group:

$$R_i = \sqrt{(D_i^+ - \min (D_i^+))^2 + (D_i^- - \max (D_i^-))^2} \quad (24)$$

A set of alternatives is ranked according to the increasing order of R_i

5. Numerical Example

A numerical example is introduced in this section to stand over the proposed approach. The group involved three experts (decision makers) with four criteria to rank five alternatives. C_1 and C_2 are benefit criteria (the maximum is better) while the other two criteria are cost type (the minimum is better). The first three criteria ratings are subjective; the ratings are allocated by experts (decision makers) while the last criterion (C_4) is measured by fixed rating. As shown in Table 1. Three decision matrices are obtained and put into this augmented table. The allocated criteria weights for each decision maker are found in Table 2.

Table 1. Criteria rating values for the three decision makers

Utility type	Max	Max	Min	Min
	C_1	C_2	C_3	C_4
D1				
Alternative 1	1.1	0.2	8	20
Alternative 2	2	0.3	16	50
Alternative 3	2.5	0.3	10	44
Alternative 4	2	0.53	10	29
Alternative 5	1.9	0.6	15	25
D2				
Alternative 1	0.9	0.3	12	20
Alternative 2	3	0.5	12	50
Alternative 3	1	1.7	10	44
Alternative 4	1.5	0.5	15	29
Alternative 5	1.9	0.6	15	25
D3				
Alternative 1	1.3	0.1	4	20
Alternative 2	1	0.1	20	50
Alternative 3	4	0.43	10	44
Alternative 4	2.5	0.56	5	29
Alternative 5	1.9	0.6	15	25

Table 2. Criteria weights for the three decision makers

	C_1	C_2	C_3	C_4
D1	0.1	0.15	0.25	0.5
D2	0.08	0.18	0.24	0.5
D3	0.12	0.12	0.26	0.5

Using the solution methods illustrated before in previous sections, the resulted scores and ranking orders are shown in Table 3. The rank originated from the three methods are identical based on their different scores and techniques (in DIA the less R score the more preferable).

Table 3. Ranking orders for the full problem

	MOORA		TOPSIS		DIA	
	Z	Rank	RC	Rank	R	Rank
Alternative 1	-0.1417	1	0.73739	1	0	1
Alternative 2	-0.3672	5	0.11151	5	0.318835	5
Alternative 3	-0.2630	4	0.32157	4	0.17148	4
Alternative 4	-0.1428	2	0.71422	2	0.001478	2
Alternative 5	-0.1543	3	0.70211	3	0.01785	3

In this simulation, the ranking abnormality problem is illustrated and the “robustness” of the algorithms to removal of an alternative of the worst alternatives. Alternative 2 (the worst) is removed from the alternatives candidate list. Table 4 presents the overall score of MOORA, the relative closeness to the ideal solution of TOPSIS and the distance to PIA of DIA.

Table 4. Ranking orders after removal of the worst alternative

	MOORA		TOPSIS		DIA	
	Z	Rank	RC	Rank	R	Rank
Alternative 1	-0.1901	1	0.73111	1	0	1
-----	-----	----	-----	----	-----	-----
Alternative 3	-0.3542	4	0.25466	4	0.23204	4
Alternative 4	-0.2061	2	0.65269	3	0.02267	2
Alternative 5	-0.2209	3	0.65294	2	0.04347	3

In this situation, the results show that a removal of an alternative causes a change in the ranking order of TOPSIS. The ranking order of MOORA, and DIA remains the same. We continue removing an alternative (i.e. alternative 3) from the alternatives candidate list. The results, in Table 5, show that the ranking order in MOORA and DIA is always stable, but the top ranked alternative in TOPSIS has changed again. In Table 3, all methods determine that Alternative 1 is the best. When we remove the worst Alternative (i.e. Alternative 1) out of the candidates list, this does not influence the ranking order of other Alternatives of MOORA and DIA. However, the second best Alternative in TOPSIS changes (i.e. from Alternative 4 to Alternative 5 in Table 4). When another worst Alternative (i.e. Alternative 3) is removed, the second best Alternative in TOPSIS also changes (see Table 5).

Table 5. Ranking orders after removal of the worst two alternatives

	MOORA		TOPSIS		DIA	
	Z	Rank	RC	Rank	R	Rank
Alternative 1	-0.2595	1	0.63465	1	0	1
-----	-----	----	-----	----	-----	----
-----	-----	----	-----	----	-----	----
Alternative 4	-0.2987	2	0.46110	2	0.055482	2
Alternative 5	-0.3065	3	0.46007	3	0.066401	3

The tables' results highlight the ranking abnormality problem of TOPSIS and show that DIA provides a more efficient behavior in this situation.

5. Conclusion

In this paper, a proposed approach based on DIA algorithm is allocated to select the best alternative in a group decision making. The proposed approach extended the advantage of DIA algorithm to explore the group decision making. The solution improves the limitations of the MCDM approach, by producing more robust and efficient ranking orders. The experiments results show that the drawback aforementioned is overcome by using the introduced approach.

References

1. Balli, S. and Korukoglu, S. (2009), "Operating system selection using fuzzy AHP and TOPSIS methods", *Mathematical & Computational Applications*, 14, pp. 119–130.
2. Brauers, W. K. M. (2008), "Multi-objective contractor's ranking by applying the MOORA method", *Journal of Business Economics and Management*, 2, pp. 245–255.
3. Chu, T. C. and Lin, Y. C. (2003), "A fuzzy TOPSIS method for robot selection", *International Journal of Advanced Manufacturing Technology*, 21, pp. 284–290.
4. El-Santawy, M. F. (2015), "A CV-COPRAS Approach for Solving Multi-Criteria Decision Making Problems", *Computing and Information Systems Journal*, 19(2), pp. 15–18.
5. El-Santawy, M. F. (2012), "A VIKOR Method for Solving Personnel Training Selection Problem", *International Journal of Computing Science*, ResearchPub, 1(2), pp. 9–12.
6. Hwang, C.L. and Yoon, K. (1981), *Multiple Attributes Decision Making Methods and Applications*, Heidelberg: Springer, Berlin.
7. Kalibatas, D. and Turskis, Z. (2008), "Multi-criteria evaluation of inner climate by using MOORA method", *Information Technology and Control*, 37, pp. 79–83.
8. Lahby, M., Cherkaoui, L. and Adib A. (2012), "New Multi Access Selection Method Based on Mahalanobis Distance", *Applied Mathematical Sciences*, 6(55), pp. 2745-2760.
9. Shih, H.S., Shyur, H.J. and Lee E.S. (2007), "An Extension of TOPSIS for Group Decision Making", *Mathematical and Computer Modelling*, 45, pp. 801-813.
10. Saaty, T.L. (1980), *The Analytic Hierarchy Process*, McGraw-Hill, New York.

Joint Chance Constrained Programming with Dependent Parameters

Nada Hafez¹

Afaf El-Dash²

Nagwa Albehery³

Abstract

In this paper, we consider joint chance constrained programming (JCCP) technique, where two probabilistic constraints are required to jointly satisfy at least a tolerance measure α . We introduce a suggested approach to obtain an equivalent deterministic model for probabilistic model with joint chance constraints (JCC's) when the RHS parameters are dependent random variables, and distributed as (i) single-parameter exponential distributions (S-PED), (ii) two-parameter exponential distributions (T-PED), which is more applicable in most real life applications than the S-PED; since it avoids having its mode at the origin. The joint density function of random RHS parameters is assumed to be the Downton bivariate exponential distribution.

It is shown that the suggested equivalent deterministic model under the assumption of S-PED is a special case of the corresponding equivalent model under the assumption of T-PED when the second parameter is zero. Also the equivalent deterministic model under the assumption of independence between random parameters is a special case of the suggested equivalent deterministic model under the assumption of dependence when the correlation coefficient is zero.

Keywords: Stochastic Programming, probabilistic programming (PP), Joint chance constrained programming (JCCP), Downton bivariate exponential distribution, Non-linear programming, convex model.

1. Introduction

Stochastic Programming approach is one of the Mathematical Programming (MP) approaches where some or all of the model parameters are random variables, taking into account the probability distributions of the random parameters in the underlying problem (Prékopa, 1995). Randomness may exist in the RHS parameter or in some (or all) LHS input coefficients, or in some (or all) objective function coefficients.

Chance constrained programming (CCP) technique is one of the stochastic programming techniques developed by Charnes and Cooper (1959, 1962, 1963) which offers powerful means of modeling stochastic decision problems with assumption that the stochastic constraints will hold with probability at least α , where α refers to the confidence level provided as an appropriate safety margin by the decision-maker and is called tolerance measure. CCP aims to transform the stochastic model into a deterministic one, then solving the deterministic model using a suitable mathematical programming technique. For theoretical background see (Prékopa, 2003).

¹ **Teaching Assistant**, Department of Mathematics , Insurance and Applied Statistics, Faculty of commerce and Business Administration, Helwan University

² **Professor of operations Research**; Department of Mathematics , Insurance and Applied Statistics, Faculty of commerce and Business Administration, Helwan University

³ **Assistant Professor**; Department of Mathematics , Insurance and Applied Statistics, Faculty of commerce and Business Administration, Helwan University

Several contributions have been suggested by many researchers assuming different probability distributions for random parameters, for example; the Normal distribution (Charnes and Cooper, 1962; Symonds, 1967; Prekopa, 1974; Jagannathan, 1974; Ackooij et al., 2011; Houda and Lisser, 2015), the Chi-square distribution (Sengupta, 1972; El-Dash, 1984), the Gamma distribution (Lingaraj and Wolfe, 1974; Atalay and Apaydin, 2011), and the Dirichlet distribution (Gouda and Szántai 2010).

The exponential distribution is our concern in this paper, and it is applicable for a wide class of economic, demographic and reliability models where most of the input coefficients and resources have to be nonnegative, which require distributions with nonnegative range such as demand, supply, prices,...etc (Hafez et al., 2017). Some attempts Gamma diswere introduced in literature to deal with exponentially distributed random parameters (Sengupta, 1972; El-Dash, 1984; Biswal et al., 1998; Hafez et al., 2017). However; they assumed Individual chance constraints (ICC's) with independent parameters where the input coefficients are random; and only Hafez et al. (2017) introduced also the case where the input coefficients are dependent exponential random variables based on the Downton bivariate exponential distribution.

In this paper; we propose an equivalent deterministic model of joint chance constrained (JCC) model where JCC's are two constraints with random RHS parameters assuming that RHS parameters are dependent exponential random variables based on the Downton bivariate exponential distribution.

This paper is organized as follows; Section 2 "Problem definition" presenting the original mathematical model. Section 3 "The Suggested approach" which includes the proposed approach for obtaining the equivalent deterministic model of JCCP models with two joint constraints and the RHS parameters are exponentially distributed based on the two-parameter Downton bivariate exponential distribution. Section 4 illustrates some "Special cases" of the proposed equivalent deterministic model. Section 5 shows a "Numerical example", and finally Section 6 "Conclusion"

2. Problem definition:

Consider the following programming model:

$$Max. (Min.) z = f(x_1, x_2, \dots, x_m) \quad (2.1)$$

$$S.t.; \sum_{j=1}^m a_{ij}x_j \leq \tilde{b}_i \quad ; i = 1,2 \quad (2.2)$$

$$\sum_{j=1}^m a_{ij}x_j \leq b_i \quad ; i = 3, \dots, n \quad (2.3)$$

$$x_j \geq 0 \quad (2.4)$$

Where a_{ij} ($i = 1, \dots, n; j = 1, \dots, m$) and b_i ($i = n' + 1, \dots, n$) are constants, and \tilde{b}_i ($i = 1,2$) are continuous random variables with a known joint distribution function, and x_j are decision variables. we refer to the stochastic parameters as \tilde{b}_i .

There are two major types of chance constraints: individual chance constraints (ICCs) and joint chance constraints (JCCs). In the former case, we put probabilistic requirements on an individual constraint. In the latter case, probabilistic constraints are required to jointly satisfy at least the tolerance measurement α . That is the solution will strongly depend on how the individual random variables depend on each other, and on the tolerance measure α as well (Miller and Wagner, 1965; Liu, 2002; Prékopa, 2003; King, 2012). Constraints (2.5) represent JCC's

$$P_r \left(\sum a_{ij}x_j \leq \tilde{b}_i \quad ; i = 1,2, \dots, n \right) \geq \alpha \quad (2.5)$$

While the ICCs are separately considered as in constraints (2.6).

$$P_r \left(\sum a_{ij}x_j \leq \tilde{b}_i \right) \geq \alpha_i \quad ; i = 1,2, \dots, n \quad (2.6)$$

In this paper; we are concerned with JCCP since JCCP is more preferable on ICCP and appropriate in contexts in which;

- i) There may be dependence between random parameters in different constraints, for example; demand and supply are both functions of price which means that there is a dependence relation between quantity demanded and quantity supplied of certain product (Luedtke, 2007).
- ii) It is important to have all constraints satisfied with single probability in the model. That is; ICCP constrain the individual probability of constraint violations independently; however, considering these constraints separately neglects to explicitly consider simultaneous violation conditions in the overall system. For example, in reliability systems, specifically in a power flow problem; where the system consists of several nodes and we need to maximize the voltage of nodes. If the ICCP approach is considered, it may be independently guaranteed that the voltage at each node in the system is under its maximum with a high probability, but this could result in a situation where multiple nodes simultaneously have overvoltage conditions. Thus, JCCP is more appropriate in this context to restrict the probability of all voltages, line flows, states of charge, etc. being within prescribed limits throughout the system (Kyri and Bridget, 2017).

Miller and Wagner (1965) are the first to consider the mathematical properties of JCC's of the form of (2.5), where RHS parameters are random and independent. They show that under certain restrictions the model can be viewed as a deterministic nonlinear programming problem, and by assuming that RHS parameters are independent with single-parameter exponential marginal; the equivalent deterministic constraint of the JCC (2.5) could be linear after applying a logarithmic transformation. In this paper; we propose an equivalent deterministic model of JCC model where JCC's are two constraints with random RHS parameters assuming that RHS parameters are dependent and distributed with i) single-parameter exponential distributions, and ii) two-parameter exponential distributions, which is more applicable in most real life applications than the one with single-parameter since it avoids having its mode at the origin.

As for the bivariate exponential distribution, many different forms were presented in literature. Gumbel (1960) presented the first type of bivariate exponential distribution, followed by Freund (1961), then many other forms were presented after that, see (Kotz et al., 2000). In this paper; Downton bivariate exponential distribution, introduced by Downton (1970), is used to model the dependence between $\tilde{b}_i, i = 1,2$ with correlation coefficient $\rho ; 0 \leq \rho < 1$ in a similar fashion as presented by Hafez et al. (2017) for random \tilde{a}_i for both single-parameter and two-parameter exponential distribution.

3. The Suggested approach

In this section, we present the equivalent deterministic model of the probabilistic linear programming model with JCC's under the assumption that random RHS parameters are dependent with \tilde{b}_1 and \tilde{b}_2 (i.e only two joint dependent constraints), where \tilde{b}_1 and \tilde{b}_2 are distributed as Downton bivariate exponential distribution with two-parameter. The approximated Pdf of the Downton bivariate exponential distribution introduced by Hafez et al. (2017) is;

$$f(\tilde{b}_1, \tilde{b}_2) = \frac{\lambda_1 \lambda_2}{(1-\rho)^2} \exp\left\{\frac{\sum_{i=1}^2 \lambda_i \gamma_i}{1-\rho}\right\} \exp\left\{-\frac{\lambda_1 \tilde{b}_1 + \lambda_2 \tilde{b}_2}{1-\rho}\right\}; 0 \leq \gamma_i \leq \tilde{b}_i; \lambda_i > 0; 0 \leq \rho < 1, \quad i = 1, 2 \quad (3.1)$$

where ρ is the correlation between \tilde{b}_1 and \tilde{b}_2 .

Theorem (3.1): Let the CCP model with JCC's be of the form of

$$\text{Max. (Min.) } z = f(x_1, x_2, \dots, x_m) \quad (3.2)$$

$$\text{S.t.; } P_r\left(\sum_{j=1}^m a_{ij} x_j \leq \tilde{b}_i \quad ; i = 1, 2\right) \geq \alpha \quad (3.3)$$

$$\sum_{j=1}^m a_{ij} x_j \leq b_i \quad ; i = 3, \dots, n \quad (3.4)$$

$$x_j \geq 0 \quad (3.5)$$

where a_{ij} and b_i are constants, and \tilde{b}_1, \tilde{b}_2 are continuous random variables and distributed as the Downton bivariate two-parameter exponential distribution with joint density function $f(\tilde{b}_1, \tilde{b}_2)$ defined in (3.1), then the deterministic equivalent model is:

$$\text{Max. (Min.) } z = f(x_1, x_2, \dots, x_m) \quad (3.6)$$

S.t.;

$$\sum_{i=1}^2 \sum_{j=1}^m \lambda_i a_{ij} x_j \leq \sum_{i=1}^2 \lambda_i \gamma_i - (1-\rho) \ln \alpha \quad (3.7)$$

$$\sum_{j=1}^m a_{ij} x_j \geq \gamma_i \quad ; i = 1, 2 \quad (3.8)$$

$$\sum_{j=1}^m a_{ij} x_j \leq b_i \quad ; i = 3, \dots, n \quad (3.9)$$

$$x_j \geq 0 \quad (3.10)$$

Proof: Since the LHS of JCC's (3.3) is:

$$\begin{aligned} & P_r \left(\tilde{b}_1 \geq \sum_{j=1}^m a_{1j} x_j, \tilde{b}_2 \geq \sum_{j=1}^m a_{2j} x_j \right) = \\ & \int_{\sum_{j=1}^m a_{2j} x_j}^{\infty} \int_{\sum_{j=1}^m a_{1j} x_j}^{\infty} \frac{\lambda_1 \lambda_2}{(1-\rho)^2} \exp\left\{\frac{\sum_{i=1}^2 \lambda_i \gamma_i}{1-\rho}\right\} \exp\left\{-\frac{\lambda_1 \tilde{b}_1 + \lambda_2 \tilde{b}_2}{1-\rho}\right\} d\tilde{b}_1 d\tilde{b}_2 \\ & = \exp\left\{\frac{\sum_{i=1}^2 \lambda_i \gamma_i}{1-\rho}\right\} \cdot \exp\left\{-\frac{\lambda_1 \sum_{j=1}^m a_{1j} x_j + \lambda_2 \sum_{j=1}^m a_{2j} x_j}{1-\rho}\right\} \end{aligned} \quad (3.11)$$

Then the equivalent deterministic constraint of the JCC's (3.3) is:

$$\exp\left\{\frac{\sum_{i=1}^2 \lambda_i \gamma_i}{1-\rho}\right\} \cdot \exp\left\{-\frac{\lambda_1 \sum_{j=1}^m a_{1j} x_j + \lambda_2 \sum_{j=1}^m a_{2j} x_j}{1-\rho}\right\} \geq \alpha \quad (3.12)$$

By applying a logarithmic transformation for constraint (3.12), constraint (3.12) becomes equivalent to:

$$\sum_{i=1}^2 \sum_{j=1}^m \lambda_i a_{ij} x_j \leq \sum_{i=1}^2 \lambda_i \gamma_i - (1-\rho) \ln \alpha \quad (3.13)$$

And the equivalent deterministic model of the probabilistic model (3.2)- (3.5) using (3.13) is given by (3.6)-(3.10)

Note that constraints (3.7)-(3.10) are linear constraints; then the deterministic model (3.6)-(3.10) is a convex model if the objective function (3.6) is convex function. Therefore; a global optimal solution is guaranteed.

4. Special Cases

1) By setting the $\gamma_i = 0$; $i = 1, 2$ constraints (3.7) and (3.8) are reduced to

$$\sum_{i=1}^2 \sum_{j=1}^m \lambda_i a_{ij} x_j \leq -(1 - \rho) \ln \alpha \quad (4.1)$$

$$\sum_{j=1}^m a_{ij} x_j \geq 0 ; i = 1,2 \quad (4.2)$$

which is the case when \tilde{b}_1 and \tilde{b}_2 are distributed as Downton bivariate single-parameter exponential distribution with pdf defined as (Hafez et al., 2017)

$$f(\tilde{b}_1, \tilde{b}_2) = \frac{\lambda_1 \lambda_2}{(1 - \rho)^2} \exp \left\{ -\frac{\lambda_1 \tilde{b}_1 + \lambda_2 \tilde{b}_2}{1 - \rho} \right\}; \tilde{b}_i \geq 0 ; \lambda_1, \lambda_2 > 0 ; \\ 0 \leq \rho < 1, i = 1,2 \quad (4.3)$$

2) by setting $\mathbf{p} = \mathbf{0}$, constraint (3.7) is reduced to

$$\sum_{i=1}^2 \sum_{j=1}^m \lambda_i a_{ij} x_j \leq \sum_{i=1}^2 \lambda_i \gamma_i - \ln \alpha \quad (4.4)$$

which is the case under the assumption that \tilde{b}_1 and \tilde{b}_2 are independent and distributed as two-parameter exponential distributions. Similarly, constraint (4.1) is reduced to constraint

$$\sum_{i=1}^2 \sum_{j=1}^m \lambda_i a_{ij} x_j \leq -\ln \alpha \quad (4.5)$$

The above case is under the assumption that \tilde{b}_1 and \tilde{b}_2 are independent and distributed as single-parameter exponential distributions (Hafez et al., 2017).

5. Numerical Example

In this section, we demonstrate the suggested approach for obtaining an equivalent deterministic model of JCCP models presented in section 3 through the following example.

Consider the following linear programming model

$$\text{Max } Z = 5x_1 + 6x_2 \quad (5.1)$$

S.t.;

$$P(5x_1 + 4x_2 \leq \tilde{b}_1 ; 10x_1 + 2x_2 \leq \tilde{b}_2) \geq 0.90 \quad (5.2)$$

$$2x_1 + x_2 \leq 5 \quad (5.3)$$

$$x_1, x_2 \geq 0 \quad (5.4)$$

Where \tilde{b}_1, \tilde{b}_2 are distributed as the Downton bivariate two-parameter exponential distribution defined in (3.1), where $\lambda_1 = 0.5, \lambda_2 = 0.3, \gamma_1 = 5, \gamma_2 = 7$ and $\rho = 0.8$. Then the equivalent deterministic model of the JCCP model (5.1)-(5.4) could be obtained using theorem (3.1) as follows:

$$\text{Max. } Z = 5x_1 + 6x_2 \quad (5.5)$$

S.t.;

$$27.5x_1 + 13x_2 \leq 23.021 \quad (5.6)$$

$$5x_1 + 4x_2 \geq 5 \quad (5.7)$$

$$10x_1 + 2x_2 \geq 7 \quad (5.8)$$

$$2x_1 + x_2 \leq 5 \quad (5.9)$$

$$x_1, x_2 \geq 0 \quad (5.10)$$

Model (5.5)-(5.10) is a linear model and the optimal solution is obtained using the Simplex method as $z^* = 6.013, x_1^* = 0.599, x_2^* = 0.503$.

6 Conclusion

In this paper, we introduced the equivalent deterministic model of JCCP model when the RHS parameters are random variables and dependent. We focused on the exponential distribution since previous researches are limited in case of non-negative random parameters. We assumed that the JCC's consists of two constraints and we introduced a suggested approach to obtain the equivalent deterministic constraint of such joint constraints assuming that \tilde{b}_1, \tilde{b}_2 are distributed as bivariate Downton exponential distribution with

- i) Two-parameters (through theorem (3.1)) and
- ii) Single-parameter (as a special case of theorem (3.1)).

By setting the second parameter $\gamma_i = 0, i = 1, 2$; It was shown that the equivalent deterministic model of theorem (3.1) for Downton T-PED could be reduced to the equivalent deterministic model for Downton S-PED. And for both deterministic models we obtain a linear constraint replacing the JCC's, which is convex, so that a global optimal solution could be obtained when the objective function is convex.

By setting the correlation coefficient $\rho = 0$; It was shown that the equivalent deterministic model of theorem (3.1) for dependent Downton T-PED or S-PED could be reduced to the equivalent deterministic model for independent two-parameter or single-parameter exponential random parameters, respectively.

References

- Ackooij, W. van ; Henrion, R. ; Moller, A. ; Zorgati, R. 2011. "On Joint Probabilistic Constraints with Gaussian Coefficient Matrix." *Operations Research Letters*, vol. 39, pp.99–102.
- Atalay, D. Kumru; Apaydin, A. 2011. "Gamma Distribution Approach in Chance-Constrained Stochastic Programming Model." *Journal of Inequalities and Applications*, vol. 108.
- Biswal, M.; Biswal, N.; Li, D. 1998. "Probabilistic Linear Programming Problems with Exponential Random Variables: A Technical Note." *European Journal of Operational Research*, vol. 111, pp. 589-597.
- Charnes, A.; Cooper, W. 1959. "Chance Constrained Programming." *Management Science*, vol. 6, pp. 73 – 79.
- Charnes, A. and Cooper, W. 1962 "Chance Constraints and Normal Deviates"; *Journal of the American Statistical Association*, vol.57, pp. 138-148.
- Charnes, A. and Cooper, W. 1963. "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints." *Operations Research*, vol.11, pp.18-39.
- Downton, F. 1970. "Bivariate Exponential Distributions in Reliability Theory." *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 32(3), pp. 408-417.
- El-Dash, A. A. 1984. *Chance-Constrained and Nonlinear Goal Programming*. Ph.D. thesis, University of North Wales, Bangor.
- Freund, John E. 1961."A Bivariate Extension of the Exponential Distribution" *Journal of the American Statistical Association*, vol. 56, no. 296, pp. 971–977.
- Gouda, A., Szántai, T. 2010. "On numerical calculation of probabilities according to Dirichlet distribution" *Annals of Operations Research*, 177: 185–200

- Gumbel, E. 1960. "Bivariate Exponential Distributions." *Journal of the American Statistical Association*, vol. 55(292), pp.698-707.
- Hafez; N. Mohammed, 2017. A joint Chance Constrained Programming Approach and its Applications. unpublished thesis; Department of Mathematics, Insurance and Applied Statistics; Faculty of commerce and Business administration, Helwan University.
- Hafez; N. Mohammed, El-Dash; A. Ali, and El Behery, N. Mohammed, 2017. "Chance constraint programming with Exponential Input Coefficients." *The Egyptian Statistical Journal*, ISSR, in press
- Houda, M., and Lisser, A. 2015. "Archimedean Copulas in Joint Chance Constrained Programming." *Operations Research and Enterprise Systems Third International Conference*, Angers, France, March 6–8, 2014.
- Jagannathan, R. 1974. "Chance-Constrained Programming with Joint Constraints." *Operations Research*, vol. 22(2), pp. 358-372.
- King, J. Alan, Wallace, W. Stein. 2012. *Modeling with Stochastic Programming*, Springer Science+Business Media, New York
- Kotz, S. Balakrishnan and Johnson, L. Norman (2000) "Continuous Multivariate Distributions" Volume 1: Models and Application; 2nd ed.; John Wiley & Sons, Inc.
- Kyri; B & Bridget; T., 2017. "Efficient Relaxations for Joint Chance Constrained AC Optimal Power Flow" *Electric Power Systems Research*, 148, pp. 230–236
- Lingaraj, B. and Wolfe, H. 1974. "Certainty Equivalent of a Chance Constraint if the Random Variable Follows a Gamma Distribution." *Sankhya: The Indian Journal of Statistics*, Vol. 36, Series B, Pt. 2, pp. 204-208.
- Liu, B. (2002) "Theory and Practice of Uncertain Programming" *Physica-Verlag*, Heidelberg.
- Luedtke, J., Ahmed, S.,and Nemhauser, G. 2007. "An Integer Programming Approach For Linear Programs With Probabilistic Constraints." In: *Proceedings of the 12th Conference for Integer Programming and Combinatorial Optimization (IPCO 2007)*, pp. 410–423.
- Miller, L. Bruce and Wagner, M. Harvey 1965. "Chance Constrained Programming With Joint Constraints". *Oper. Res.* 13, 930–945.
- Prékopa, A. (1974) "Programming under Probabilistic Constraints with a Random Technology Matrix" *Math.; Operations research*; 5, pp. 109–116.
- Prékopa, A. (1995) "Stochastic Programming" *Kluwer Academic Publishers*, Dordrecht; The Netherlands.
- Prékopa, A. (2003) "Probabilistic programming" In: A. Ruszczyński, A. Shapiro (eds.) *Stochastic Programming, Handbooks in Operations Research and Management Science*, vol. 10, pp. 267–351.
- Sengupta, K. Jati 1972. "Chance-Constrained Linear Programming with Chi-Square Type Deviates." *Management Science*, Vol. 19 (3) p. 337.
- Symonds, G. 1967. "Deterministic Solutions for a Class of Chance Constrained Programming Problems." *Operations Research*, vol. 15(3), pp. 495-512.

**The 52nd Annual Conference on Statistics, Computer
Sciences and Operation Research**

Mathematical Statistics

25-27 Dec. 2017

Index
Mathematical statistics

1	Identification of Double Seasonal Autoregressive Models: A Bayesian Approach Ayman A Amin	1-10
2	On Expressing Continuous Distributions With Discrete Distributions (A Review Of Literature) Mohamed. S. A. Muiftah	11-23
3	Estimation for Birnbaum-Saunders Distribution in Constant- Stress Accelerated Life Test with Hybrid Censoring Mohamed S. Hamouda	24-44

Identification of Double Seasonal Autoregressive Models: A Bayesian Approach

Ayman A Amin^{*}

Abstract

In this paper we present a Bayesian methodology to identify the order of double seasonal autoregressive (DSAR) models. Assuming the model errors are normally distributed and using conjugate priors on the model parameters, we derive the joint posterior mass function of the model order in a closed-form. Accordingly, the posterior mass function can be investigated and the best order of DSAR model is chosen as a value with the highest posterior probability. We evaluate the proposed Bayesian methodology using simulation study and real-world hourly internet amount of traffic datasets.

Key Words: Bayesian analysis; DSAR models; Posterior distribution; Double seasonality; Hourly internet traffic.

1 Introduction

High frequency time series that are observed at small time units may be characterized by exhibiting multiple seasonal patterns. For example, hourly internet traffic data can exhibit intraday and intraweek seasonal patterns. Other examples of high frequency time series contain multiple seasonal patterns include daily hospital admissions, daily usage of water and natural gas, hourly volumes of call center arrivals, hourly traffic on a road, hourly access to computer web sites, and hourly electricity load. The notion of modelling multiple seasonalities is not new and it can be traced back to 1971 when Thompson and Tiao (1971) showed that monthly disconnections of the Wisconsin telephone company have annual and quarterly (double) seasonal patterns. Accordingly, seasonal autoregressive moving average (SARMA) models being widely applied to analyze time series with single seasonal pattern need to be modified and extended to accommodate multiple seasonalities, see for example Box et al. (1994) and Taylor (2003). In addition to SARMA models, other techniques have been extended to fit multiple seasonal time series, which include exponential smoothing methods, neural networks, and innovation state models (Feinberg and Genethliou, 2005).

In particular, the double SARMA (DSARMA) models have been the subject of interest of many researchers and extensively studied and employed in modeling and forecasting double seasonal time series data. Taylor (2003) showed that electricity load in England and Wales features daily (within day) and weekly (within week) seasonal patterns. Taylor et al. (2006) compared the forecast accuracy of six univariate models including DSARMA for electricity demand forecasting in Brazil and in England and Wales. Cruz et al. (2011) empirically compared the predictive accuracy of a set of methods for day-ahead spot price forecasting in the Spanish electricity market.

^{*} Assistant Professor of Statistics, Department of Statistics, Mathematics, and Insurance, Faculty of Commerce, Menoufia University, Egypt. Email: ayman.aamin@gmail.com.

Other references may include Taylor (2008a), Taylor (2008b), Caiado (2008), Baek (2010), Mohamed et al. (2011), and Kim (2013) and references therein, among others.

Bayesian analysis of SARMA model for single seasonality has been well established, and different approaches have been developed in literature, including Markov Chain Monte Carlo (MCMC) methods and analytical approximations. Chib and Greenberg (1994) and Marriott et al. (1996) developed a Bayesian analysis of ARMA models using MCMC methods, and Barnett et al. (1996,1997) used MCMC methods to estimate seasonal autoregressive and ARMA models. Ismail (2003) used Gibbs sampling algorithm, as one of MCMC methods, to ease the Bayesian analysis of seasonal autoregressive models. This work was extended by Ismail and Amin (2014) to SARMA models. The main criticism that can be raised against the use of MCMC methods for Bayesian analysis of time series models is that they are highly iterative and computationally intensive. On the other hand, analytical approximations work on approximating the posterior and predictive densities to be standard closed-form distributions that are analytically tractable, see for example Broemeling and Shaarawy (1984), Shaarawy and Ismail (1987), and Shaarawy and Ali (2003).

With respect to the Bayesian analysis of double seasonal autoregressive (DSAR) models, it is immature and few work have been introduced in literature. Recently, Amin and Ismail (2015) used Gibbs sampling algorithm to develop a Bayesian estimation of DSAR models, and Amin (2017) used the analytical approximations to develop a Bayesian inference of DSARMA models. Based on our review for the existing Bayesian analysis of DSAR models, there is no work that considers the problem of DSAR models identification. Since in reality the order of the DSAR models is unknown and needs to be identified or estimated. In order to fill this gap, in this paper we propose a Bayesian method to identify the order of the DSAR models. The main idea of the proposed Bayesian identification method is that we assume the order of the DSAR model is a random variable with a known maximum, and we derive its posterior mass function to select the order as a value with a maximum posterior probability.

The remainder of this paper is organized as follows: Section 2 presents the double seasonal autoregressive (DSAR) models. Section 3 is devoted to summarizing the proposed Bayesian identification method for DSAR models. The simulation study and real application of the proposed Bayesian identification method are given in Section 4. Finally, the conclusions are given in Section 5.

2 Double Seasonal Autoregressive (DSAR) Models

A time series $\{y_t\}$ is said to be generated by a double seasonal autoregressive model of order p , P_1 , and P_2 , denoted by DSAR(p)(P_1) s_1 (P_2) s_2 , if it satisfies

$$\phi_p(B)\Phi_{P_1}(B^{s_1})\Pi_{P_2}(B^{s_2})y_t = \varepsilon_t \quad (1)$$

where $\{\varepsilon_t\}$ is a sequence of independent normal variates with zero mean and variance σ^2 . The backshift operator B is defined as $B^k y_t = y_{t-k}$, s_1 and s_2 are the seasonal periods. The non seasonal autoregressive polynomial is $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ with order p . In addition, $\Phi_{P_1}(B^{s_1}) = (1 - \Phi_1 B^{s_1} - \Phi_2 B^{2s_1} - \dots - \Phi_{P_1} B^{P_1 s_1})$ with order P_1 and $\Pi_{P_2}(B^{s_2}) = (1 - \Pi_1 B^{s_2} - \Pi_2 B^{2s_2} - \dots - \Pi_{P_2} B^{P_2 s_2})$ with order P_2 are the seasonal

autoregressive polynomials. Finally, the non seasonal and seasonal autoregressive coefficients are $\phi = (\phi_1, \phi_2, \dots, \phi_p)^T$, $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_{P_1})^T$ and $\Pi = (\Pi_1, \Pi_2, \dots, \Pi_{P_2})^T$, respectively.

It should be noted that the DSAR model (1) has an extra terms compared with the usual single SAR model. The new term is $\Pi_{P_2}(B^{S_2})$ that accommodates the second seasonal pattern. Accordingly, the model (1) can be written as

$$y_t = \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^{P_1} \Phi_j y_{t-j_{S_1}} + \sum_{\tau=1}^{P_2} \Pi_{\tau} y_{t-\tau_{S_2}} - \sum_{i=1}^p \sum_{j=1}^{P_1} \phi_i \Phi_j y_{t-i-j_{S_1}} - \sum_{i=1}^p \sum_{\tau=1}^{P_2} \phi_i \Pi_{\tau} y_{t-i-\tau_{S_2}} - \sum_{j=1}^{P_1} \sum_{\tau=1}^{P_2} \Phi_j \Pi_{\tau} y_{t-j_{S_1}-\tau_{S_2}} + \sum_{i=1}^p \sum_{j=1}^{P_1} \sum_{\tau=1}^{P_2} \phi_i \Phi_j \Pi_{\tau} y_{t-i-j_{S_1}-\tau_{S_2}} + \varepsilon_t \quad (2)$$

The matrix form of this model can be written as:

$$y = X\beta + \varepsilon, \quad (3)$$

where $y = (y_1, y_2, \dots, y_n)^T$, $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$, X is an $n \times p^*$ design matrix, where $p^* = (1 + p)(1 + P_1)(1 + P_2) - 1$, with the t^{th} row:

$$X_t = (y_{t-1}, \dots, y_{t-p}, y_{t-S_1}, y_{t-S_1-1}, \dots, y_{t-S_1-p}, \dots, y_{t-P_1 S_1}, y_{t-P_1 S_1-1}, \dots, y_{t-P_1 S_1-p}, y_{t-S_2}, y_{t-S_2-1}, \dots, y_{t-S_2-p}, y_{t-S_1-S_2}, y_{t-S_1-S_2-1}, \dots, y_{t-S_1-S_2-p}, \dots, y_{t-P_1 S_1-S_2}, y_{t-P_1 S_1-S_2-1}, \dots, y_{t-P_1 S_1-S_2-p}, \dots, y_{t-P_2 S_2}, y_{t-P_2 S_2-1}, \dots, y_{t-P_2 S_2-p}, y_{t-S_1-P_2 S_2}, y_{t-S_1-P_2 S_2-1}, \dots, y_{t-S_1-P_2 S_2-p}, \dots, y_{t-P_1 S_1-P_2 S_2}, y_{t-P_1 S_1-P_2 S_2-1}, \dots, y_{t-P_1 S_1-P_2 S_2-p}),$$

$$\beta = (\phi_1, \dots, \phi_p, \Phi_1, -\phi_1 \Phi_1, \dots, -\phi_p \Phi_1, \dots, \Phi_{P_1}, -\phi_1 \Phi_{P_1}, \dots, -\phi_p \Phi_{P_1}, \Pi_1, -\phi_1 \Pi_1, \dots, -\phi_p \Pi_1, -\Phi_1 \Pi_1, \phi_1 \Phi_1 \Pi_1, \dots, \phi_p \Phi_1 \Pi_1, \dots, -\Phi_{P_1} \Pi_1, \phi_1 \Phi_{P_1} \Pi_1, \dots, \phi_p \Phi_{P_1} \Pi_1, \dots, \Pi_{P_2}, -\phi_1 \Pi_{P_2}, \dots, -\phi_p \Pi_{P_2}, -\Phi_1 \Pi_{P_2}, \phi_1 \Phi_1 \Pi_{P_2}, \dots, \phi_p \Phi_1 \Pi_{P_2}, \dots, -\Phi_{P_1} \Pi_{P_2}, \phi_1 \Phi_{P_1} \Pi_{P_2}, \dots, \phi_p \Phi_{P_1} \Pi_{P_2})^T \quad (4)$$

Equation (2) shows that the DSAR model can be written as an autoregressive model of order $(1 + p)(1 + P_1)(1 + P_2) - 1$ with some coefficients that are products of non seasonal and seasonal coefficients. The DSAR model (2) is stationary if the roots of the polynomials $\phi(B)$, $\Phi_{P_1}(B^{S_1})$ and $\Pi_{P_2}(B^{S_2})$ lie outside the unit circle. For more details about the properties of seasonal AR models see Box et al. (1994).

It is worth noting that the design matrix X becomes a function of p , P_1 , and P_2 when the DSAR model order is unknown. In this case we can assume that the model order p , P_1 , and P_2 is a random variable with known maximum values of k_1 , k_2 , and k_3 respectively. The prior information about p , P_1 , and P_2 can be represented in terms of a prior mass function $\zeta(p, P_1, P_2)$ that can have different forms such as uniform or geometric.

3 Proposed Bayesian Identification for DSAR Models

Bayesian analysis of the DSAR models is based on Bayes' theorem that combines the prior distribution of the model parameters with the likelihood function of observed sample to get the posterior distribution.

Regarding the prior specification, we consider the natural conjugate prior. In case of the DSAR model with normally distributed errors, the natural conjugate prior is normal-gamma. So, suppose $\tau = 1/\sigma^2 \sim G(\frac{\nu}{2}, \frac{\lambda}{2})$ and $\beta \sim N_{p^*}(\mu_\beta, \tau^{-1}\Sigma_\beta)$, the joint natural conjugate prior distribution of β and τ can be written as:

$$\zeta_n(\beta, \tau) \propto \tau^{\left(\frac{\nu+p^*}{2}-1\right)} \exp\left\{-\frac{\tau}{2}\left[\lambda + (\beta - \mu_\beta)^T \Sigma_\beta^{-1}(\beta - \mu_\beta)\right]\right\}, \quad (5)$$

where $\mu_\beta, \Sigma_\beta, \nu$ and λ are hyperparameters need to be estimated, and the products of non seasonal and seasonal coefficients are considered as free coefficients.

In case of no or little information is available about the model parameters β and τ Jeffreys' prior can be employed and it is given as:

$$\zeta_j(\beta, \tau) \propto \tau^{-1}, \tau > 0 \quad (6)$$

Employing a straightforward random variable transformation from ε to y , the conditional likelihood function (on the first p^* initial values) is given by

$$\begin{aligned} L(\beta, \tau, p, P_1, P_2 | y) &\propto \tau^{\frac{n-p^*}{2}} \exp\left\{-\frac{\tau}{2} \varepsilon^T \varepsilon\right\}, \\ &\propto \tau^{\frac{n-p^*}{2}} \exp\left\{-\frac{\tau}{2} (y - X\beta)^T (y - X\beta)\right\}, \end{aligned} \quad (7)$$

Based on the likelihood function (7), we update the information about the DSAR model order p, P_1 , and P_2 by deriving the posterior probability mass function. To derive this posterior mass function, we need first to obtain the joint posterior of the model parameters β, τ, p, P_1 , and P_2 and then integrate out the parameters β and τ .

For the natural conjugate prior, the joint posterior of the model parameters is obtained as:

$$\zeta_n(\beta, \tau, p, P_1, P_2 | y) \propto \zeta(p, P_1, P_2) \tau^{\left(\frac{n+\nu}{2}-1\right)} \exp\left\{-\frac{\tau}{2}\left[\lambda + (\beta - \mu_\beta)^T \Sigma_\beta^{-1}(\beta - \mu_\beta) + (y - X\beta)^T (y - X\beta)\right]\right\}. \quad (8)$$

For Jeffreys' prior, the resulting joint posterior of β and τ is:

$$\zeta_j(\beta, \tau, p, P_1, P_2 | y) \propto \zeta(p, P_1, P_2) \tau^{\left(\frac{n-p^*}{2}-1\right)} \exp\left\{-\frac{\tau}{2} (y - X\beta)^T (y - X\beta)\right\}. \quad (9)$$

In the following theorem we use the natural conjugate prior to obtain the joint posterior mass function of the DSAR model order.

Theorem 1 Using the joint natural conjugate prior distribution of β and τ (5) and the

conditional likelihood function (7), the joint posterior mass function of the DSAR model order p , P_1 , and P_2 is given by

$$\zeta_n(p, P_1, P_2 | y) \propto \zeta(p, P_1, P_2) \left[\frac{|\Sigma_\beta^{-1}|}{|A_n|} \right]^{1/2} [y^T y + \lambda + \mu_\beta^T \Sigma_\beta^{-1} \mu_\beta - B_n^T A_n^{-1} B_n]^{-\frac{n+\nu}{2}}$$

$$\forall p = 1, \dots, k_1, P_1 = 1, \dots, k_2, P_2 = 1, \dots, k_3. \quad (10)$$

Where $A_n = (X^T X + \Sigma_\beta^{-1})$, $B_n = (X^T y + \Sigma_\beta^{-1} \mu_\beta)$, and $\zeta(p, P_1, P_2)$ is a prior mass function of the model order.

Proof. The proof follows by completing the square in the exponent of (??) with respect to β and then integrating over β and τ respectively results in the given joint posterior mass function of the DSAR model order.

Based on Theorem 1, we can straightforwardly identify the order of the DSAR model by computing all possible posterior probabilities from the joint posterior mass function (??) and choose the values of p , P_1 , and P_2 that have the highest posterior probabilities to be the appropriate DSAR model order for the time series under study. In case of no or little information is available a priori about the model parameters β and τ , in the following corollary we obtain the joint posterior mass function of the DSAR model order.

Corollary 1 Using Jeffreys' prior of β and τ (6) and the conditional likelihood function (7), the joint posterior mass function of the DSAR model order p , P_1 , and P_2 is given by

$$\zeta_j(p, P_1, P_2 | y) \propto \zeta(p, P_1, P_2) \frac{\Gamma(\frac{n-p^*}{2})}{\pi^{\frac{n-p^*}{2}} |X^T X|^{1/2}} [y^T y - y^T X (X^T X)^{-1} X^T y]^{-\frac{n-p^*}{2}}$$

$$\forall p = 1, \dots, k_1, P_1 = 1, \dots, k_2, P_2 = 1, \dots, k_3. \quad (11)$$

Proof. The proof follows directly from Theorem 1 by setting $\lambda = 0$, $\Sigma_\beta^{-1} = 0$, and $\nu = -p^*$.

4 Simulation Study and Real Application

In this section we first present a simulation study to evaluate the efficiency of the proposed Bayesian identification method for DSAR models, and then we apply the proposed method to real world time series dataset of internet traffic.

4.1 Simulation Study

In this simulation study, we generate 100 time series of size n (from 200 to 500 with an increment of 100 observations) from different DSAR models. Table 0 shows these DSAR models and their corresponding true parameters values. By these DSAR models we try to cover different seasonality patterns with different data type.

Model	ϕ_1	ϕ_2	Φ_1	Φ_2	Π_1	Π_2	σ^2
I.DSAR(1)(1) _{4(1)₂₈}	0.80	-	0.20	-	0.30	-	1.00
II. DSAR(2)(1) _{4(1)₂₈}	-0.40	0.30	0.40	-	0.40	-	1.00
III.DSAR(2)(2) _{4(1)₂₈}	0.5	-0.40	-0.50	0.40	0.40	-	1.00
IV.DSAR(2)(2) _{4(2)₂₈}	0.4	0.40	0.60	-0.50	0.30	0.50	1.00

Table 1: Simulation design.

Once the time series datasets are generated from these DSAR models, the Bayesian analysis is performed by assuming Jeffreys’ prior for the model parameters β and τ , and employing three prior distributions for the model order given as:

$$\begin{aligned} \zeta_1(p, P_1, P_2) &= \frac{1}{k_1} \times \frac{1}{k_2} \times \frac{1}{k_3}, \forall p = 1, \dots, k_1, P_1 = 1, \dots, k_2, P_2 = 1, \dots, k_3 \text{ (Uniform prior)} \\ \zeta_2(p, P_1, P_2) &= 0.5^{p+P_1+P_2}, \forall p = 1, \dots, k_1, P_1 = 1, \dots, k_2, P_2 = 1, \dots, k_3 \text{ (Geometric prior)} \\ \zeta_3(p, P_1, P_2) &= \frac{k_1 - p + 1}{k_1 + 1} \times \frac{k_2 - P_1 + 1}{k_2 + 1} \times \frac{k_3 - P_2 + 1}{k_3 + 1}, \\ &\quad \forall p = 1, \dots, k_1, P_1 = 1, \dots, k_2, P_2 = 1, \dots, k_3 \text{ (Arithmetic prior)} \end{aligned} \quad (12)$$

For each time series, we compute the posterior mass functions of the DSAR model order, $\zeta_1(p, P_1, P_2|y)$, $\zeta_2(p, P_1, P_2|y)$, and $\zeta_3(p, P_1, P_2|y)$ resulting from the employed priors in eq(12) respectively with assuming the maximum values of the model order are $k_1 = k_2 = k_3 = 3$, and then we identify the model order as a value with a maximum posterior probability. For all simulated time series, we compute the percentage of correctly identified models by comparing the identified order with the true value of p , P_1 , and P_2 used to generate the time series. Results of the simulation study are presented in Table (1).

From the simulation results we can observe some remarks. First, the larger sample size the higher percentage of correctly identified models is obtained. Second, the employed priors for the model order result in very similar posteriors, and their impact in the percentage of correctly identified models is small and can be ignored especially for sample size $n \geq 300$. Third, a higher percentage of correctly identified models is obtained for simple DSAR models (Models I and II) compared to complicated DSAR models (Models III and IV). In general, all these simulation results confirm the efficiency of the proposed Bayesian method for identifying the DSAR models.

4.2 Application to Real-World Internet Traffic Time Series

The internet traffic time series represents real-world hourly internet amount of traffic (in bits) dataset collected from private Internet service provider (ISP) with centers in 11 European cities. In particular, Cortez et al. (2012) collected this time series dataset from 06:57 hours on 7 June to 11:17 hours on 29 July 2005, which is available online and can be downloaded [†].

[†] <http://www3.dsi.uminho.pt/pcortez/series/>

n	$\zeta_1(p, P_1, P_2 \gamma)$	$\zeta_2(p, P_1, P_2 \gamma)$	$\zeta_3(p, P_1, P_2 \gamma)$
Model I			
200	94	97	96
300	100	100	100
400	99	100	100
500	100	100	100
Model II			
200	79	72	77
300	96	94	94
400	99	99	99
500	97	99	99
Model III			
200	55	52	54
300	85	84	87
400	95	94	94
500	100	100	100
Model IV			
200	21	16	19
300	64	62	62
400	91	90	90
500	96	96	96

Table 2: Simulation results for percentage of correctly identified models.

Figure (1) shows how the daily and weekly seasonal patterns are represented in the internet traffic time series which reveal clearly the double seasonality. Accordingly, it is strongly recommended to consider two seasonal components for this dataset, where the first seasonal component is due to the intraday cycle ($s_1 = 24$) and the second seasonal component is due to the intraweek cycle ($s_2 = 168$).

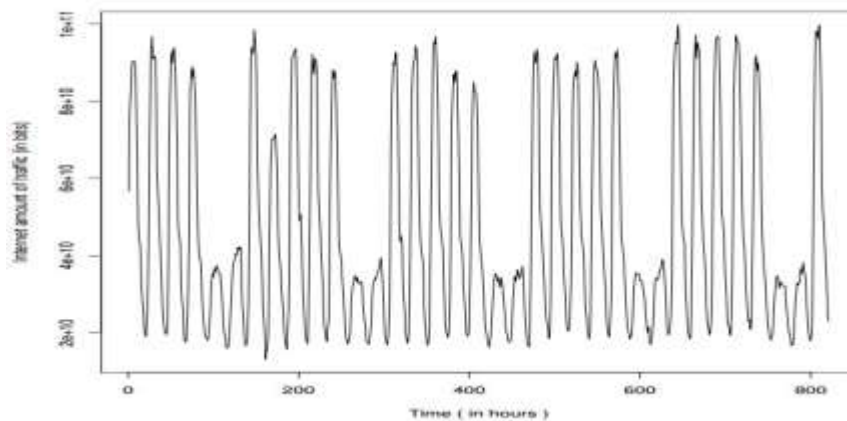


Figure 1: Time plot of the internet traffic time series.

Cortez et al. (2012) tested a collection of DSAR models using different combinations of the

p , P_1 , and P_2 values up to a maximum order of 2 in order to identify the best DSAR model that can fit the internet amount of traffic dataset. They decided that DSAR(2)(2) $_{24}(2)_{168}$ is the best model that has smallest value of Bayesian information criterion (BIC).

We can apply our proposed Bayesian methodology to identify the best order of DSAR model to fit the internet amount of traffic dataset. We first employ the three prior mass functions of the model order introduced in eqn (12) and then compute the posterior mass functions of the DSAR model order, $\zeta_1(p, P_1, P_2|y)$, $\zeta_2(p, P_1, P_2|y)$, and $\zeta_3(p, P_1, P_2|y)$ with assuming the maximum values of the model order are 3. Results of the posterior probabilities are presented in Table (2). Accordingly, we identify the model order as a value with the maximum posterior probability as highlighted in Table (2). From these results, we can observe that the employed three priors for the model order result in very similar identified DSAR models, which are also similar to that identified by Cortez et al. (2012).

Model Order (p)(P ₁)(P ₂)	Priors of Model Order		
	$\zeta_1(p, P_1, P_2 y)$	$\zeta_2(p, P_1, P_2 y)$	$\zeta_3(p, P_1, P_2 y)$
(1)(1)(1)	1.3E-109	7.0E-109	4.0E-109
(2)(1)(1)	3.6E-03	9.9E-03	7.5E-03
(3)(1)(1)	1.1E-03	1.5E-03	1.1E-03
(1)(2)(1)	3.6E-112	1.0E-111	7.5E-112
(2)(2)(1)	2.1E-06	2.8E-06	2.8E-06
(3)(2)(1)	9.4E-08	6.5E-08	6.5E-08
(1)(3)(1)	2.0E-113	2.8E-113	2.1E-113
(2)(3)(1)	2.3E-13	1.6E-13	1.6E-13
(3)(3)(1)	6.2E-16	2.1E-16	2.1E-16
(1)(1)(2)	1.7E-107	4.8E-107	3.6E-107
(2)(1)(2)	4.4E-01	6.1E-01	6.1E-01
(3)(1)(2)	4.0E-02	2.7E-02	2.8E-02
(1)(2)(2)	1.5E-113	2.0E-113	2.0E-113
(2)(2)(2)	5.4E-06	3.7E-06	5.0E-06
(3)(2)(2)	1.3E-11	4.5E-12	6.0E-12
(1)(3)(2)	2.4E-116	1.7E-116	1.7E-116
(2)(3)(2)	2.9E-18	1.0E-18	1.4E-18
(3)(3)(2)	3.4E-23	5.9E-24	7.9E-24
(1)(1)(3)	7.0E-104	9.6E-104	7.2E-104
(2)(1)(3)	5.1E-01	3.5E-01	3.5E-01
(3)(1)(3)	5.7E-05	2.0E-05	2.0E-05
(1)(2)(3)	1.2E-112	8.0E-113	8.1E-113
(2)(2)(3)	3.1E-11	1.1E-11	1.4E-11
(3)(2)(3)	1.3E-19	2.2E-20	2.9E-20
(1)(3)(3)	9.2E-115	3.2E-115	3.2E-115
(2)(3)(3)	6.3E-26	1.1E-26	1.4E-26
(3)(3)(3)	3.9E-37	3.3E-38	4.5E-38
Identified	DSAR(2)(1)(3)	DSAR(2)(1)(2)	DSAR(2)(1)(2)

Table 3: Posterior probability and identified models for the real-world time series.

5 Conclusions

In this paper we presented a Bayesian methodology to identify the order of DSAR models. We derived the joint posterior mass function of the model order in a closed-form by combining the conjugate priors on the model parameters with the conditional likelihood function, assuming the model errors are normally distributed. Therefore, we can easily inspect the posterior mass function of the DSAR model order and choose a value with the highest posterior probability as the best model order for the time series being analyzed. We conducted a simulation study with different priors for the model order and the results confirmed the efficiency of the proposed Bayesian methodology. In addition, we applied our proposed methodology to real-world hourly internet amount of traffic dataset. Future work may investigate model identification and outliers detection for multivariate double seasonal models.

References

1. Amin, A.A. Bayesian Inference for Double SARMA Models. *Journal of Communications in Statistics - Theory and Methods*, DOI: 10.1080/03610926.2017.1390132, 2017.
2. Amin, A.A. and Ismail, M.A. (2015). Gibbs Sampling for Double Seasonal Autoregressive Models. *Communications for Statistical Applications and Methods*, 22(6), 557-573.
3. Baek, M. (2010). Forecasting Hourly Electricity Loads of South Korea: Innovations State Space Modeling Approach. *The Korean Journal of Economics* 17(2), 301-317.
4. Barnett, G., Kohn, R., and Sheather, S. (1996). Bayesian estimation of an autoregressive model using Markov Chain Monte Carlo. *Journal of Econometrics*, 74, 237-254.
5. Barnett, G., Kohn, R., and Sheather, S. (1997). Robust Bayesian estimation of an autoregressive moving-average models. *Journal of Time Series Analysis*, 18, 11-28.
6. Box, G.E.P, Jenkins, G.M and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control*, 3rd ed, Prentice-Hall, Inc.:NJ.
7. Broemeling, L. D. and Shaarawy, S. (1984). Bayesian inferences and forecasts with moving average processes. *Communications in Statistics: Theory and Methods*, 13, 1871-1888.
8. Caiado, J (2008). Forecasting Water Consumption in Spain Using Univariate Time Series Models. MPRA Paper no. 6610.
9. Chib, S. and Greenberg, E. (1994). Bayes inference in regression models with ARMA(p,q) errors. *Journal of Econometrics*, 64, 183-206.
10. Cortez, P., Rio, M., Rocha, M. and Sousa, P. (2012). Multi-scale Internet Traffic Forecasting Using Neural Networks and Time Series Methods. *Expert Systems* 29 (2), 143-155.
11. Cruz, A., Munoz,A., Zamora, J.L., Espinola, R. (2011). The effect of Wind Generation and Weekday on Spanish Electricity Spot price Forecasting. *Electric Power Systems Research* 81(10), 1924-1935.
12. Feinberg, E. and Genethliou, D. (2005). Load forecasting. In F. W. J. Chow, *Applied Mathematics for Restructured Electric Power Systems: Control and Computational Intelligence* (pp. 269-285). Springer.
13. Ismail, M. A. (2003). Bayesian Analysis of Seasonal Autoregressive Models. *Journal of Applied Statistical science*, 12(2), 123-136.

14. Ismail, M.A. and Amin, A.A. (2014). Gibbs Sampling for SARMA Models. *Pakistan Journal of Statistics*, 30(2), 153-168.
15. Kim, M.S. (2013). Modeling Special-day Effects for Forecasting Intraday Electricity demand, *European Journal of Operational Research* 230(1), 170-180.
16. Marriott, J., Ravishanker, N., Gelfand, A. and Pai, J. (1996). Bayesian analysis of ARMA processes: Complete sampling based inference under full likelihoods. In *Bayesian Statistics and Econometrics: Essays in honor of Arnold Zellner, D., Berry, K., Chaloner, and J. , Geweke, (eds)*. New York, Wiley.
17. Mohamed, N, Ahmad, MH and Suhartono (2011). Forecasting Short Term Demand Using Double Seasonal ARIMA Model. *World Applied Sciences Journal* 13, 27-35.
18. Shaarawy, S. and Ismail, M. A. (1987). Bayesian Inference for Seasonal ARMA Models. *The Egyptian Statistical Journal*, 31, 323-336.
19. Shaarawy, S., and Ali, S. (2003). Bayesian Identification of Seasonal Autoregressive Models. *Communications in Statistics: Theory and Methods* 32 (5), 1067-1084.
20. Taylor, J.W. (2003). Short-Term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing. *The Journal of the Operational Research Society* 54(8), 799-805.
21. Taylor, J. W., de Menezes, L. M. and McSharry, P. A (2006). Comparison of Univariate Methods for Forecasting Electricity Demand up to a Day ahead. *International Journal of Forecasting* 22, 1-16.
22. Taylor, J. W. (2008a). An Evaluation of Methods for very Short-term Load Forecasting using Minute-by-Minute British Data. *International Journal of Forecasting* 24: 645- 658.
23. Taylor, J.W. (2008b). A Comparison of Univariate Time Series Methods for Forecasting Intraday Arrivals at a Call Center. *Management Science* 54, 253-265.
24. Thompson, H.E, and Tiao, G.C. (1971). Analysis of Telephone Data. *The Bell Journal of Economics and Management Science* 2, 514-541.

On Expressing Continuous Distributions With Discrete Distributions (A Review Of Literature)

Mohamed. S. A. Muiftah¹

Abstract:

In this paper, the motivation of expressing continuous distributions via discrete ones is discussed, discretizing methods are introduced, and the discretized distributions obtained using these methods are also reviewed.

Key Words: (Continuous distribution, discretization, survival function, density function, distribution function, failure rate).

Introduction:

It is sometimes inconvenient or difficult to get samples from a continuous distribution in real life. The observed values are discrete because they are usually measured to only a finite number of decimal places and cannot really constitute all points in a continuum. Even if the measures are taken on a continuous (ratio or interval) scale the observations may be recorded in a way making discrete model more appropriate. Therefore, it is reasonable to consider the observations as coming from a discretized distribution generated from the underlying continuous model.

The reliability (survival) function is sometimes assumed to be a function of a discrete random variable instead of being a function of continuous time random variable. For example, the reliability of an on/off - switching device is a function of the number of operating times, and the reliability of an airplane tire is a function of the number of landings. Similarly, the length of stay in an observation ward is usually measured as number of days, and survival time of leukemia or lung cancer patients may be measured in number of weeks. Many discrete distributions are available to model such mentioned situations, for example, the geometric distribution is a

¹ Department of Mathematical Statistics, ISSR, Cairo University, Egypt.

discrete alternative of the exponential distribution and the negative binomial distribution is a discrete alternative of the gamma distribution, but it is also well known that these discrete distributions are unsuitable for some situations because of their monotonic hazard rate functions [Jazi et al. (2010)], therefore, research is ongoing to find new discrete distributions suitable to treat different conditions and to fit to various types of lifetime data. In the light of the above contexts, the study of the discretization of continuous distributions is meaningful.

General approaches of discretization:

Discrete analogues of continuous distributions may be obtained using some methods of discretization. Only four methods will be introduced in this work:

Approach 1: Discretizing using the survival function:

Suppose that the continuous random variable X has the survival function (sf) $S(x) = \Pr(X \geq x)$, then the random variable $Y = [X]$ (where, $[X]$ is the largest integer less than or equal to X) will have the probability mass function (pmf):

$$\begin{aligned} P(Y = y) &= P(y \leq X < y+1) \\ &= P(X \geq y) - P(X \geq y+1) \\ &= S(y) - S(y+1), \quad y = 0, 1, 2, \dots \end{aligned} \quad (1)$$

The pmf of the random variable Y may be considered as a discrete concentration of the probability density function (pdf) of the continuous random variable X . (Roy, 2004).

Using formula (1), it is possible to generate a discrete distribution corresponds to a given continuous distribution. For example, If X is an exponential random variable with the sf $S(x) = \Pr(X \geq x) = e^{-\theta x}$, then the pmf of its discrete version is given by:

$$P(Y = y) = e^{-\theta y} - e^{-\theta(y+1)} = q^y - q^{y+1} = (1-q)q^y, \quad y = 0, 1, 2, \dots$$

where, $q = e^{-\theta}$. This is the geometric distribution.

Discretization of continuous distributions using this method retains the same functional form of the survival function. So, many reliability characteristics remain unchanged [Krishna and Pundir (2009)]. There is enough motivation to use this technique of generating discretized version of continuous distribution with this approach to develop new discrete lifetime models corresponding to the existing continuous one.

Following this approach, Many researchers studied discretization of some known continuous distributions for use as lifetime distributions. Nakagawa and Osaki (1975) proposed a discrete Weibull distribution. Xie et al. (2002) introduced the modified Weibull extension. Bracquemond and Gaudion (2003) introduced the discrete geometric Weibull distribution and investigated the discrete truncated logistic distribution. Roy (2003) derived the PMF of the discrete analogue of the normal distribution and elaborated its application for evaluating the reliability of complex systems as an alternative to simulation methods. Roy (2004) proposed a discrete Rayleigh version as a particular case of the discrete Weibull distribution and used it in approximating the reliability value of a complex system to examine probability integrals arising out of a continuous setting. Krishna and Pundir (2007) derived the discrete Maxwell distribution. The modified Weibull Type I distribution was introduced by Sarhan and Zaindin (2009). Krishna and Pundir (2009) derived the discrete Burr XII distribution and applied it to fit the reliability in series system and a set of real data, they also derived the discrete Pareto distribution as a special case of the discrete Burr XII distribution. The discrete extended exponential (telescopic) distribution was introduced by Roknabadi et al. (2009). Jazi et al. (2010) introduced the discrete inverse Weibull distribution. Gómez-Déniz (2010) used the Marshal-Olkin scheme to generalize the SF of the exponential distribution and then got the discrete generalized exponential distribution. Gómez-Déniz and Calderin-Ojeda (2011) introduced the discrete Lindley distribution and used it as an alternative to the Poisson distribution to model automobile claim frequency data. Nooghabi et al. (2011) got the modified Weibull Type II distribution as a special case of the telescopic distribution. Bebbington et al. (2012) explored a discrete analogue of the additive Weibull distribution. Bakouch et al. (2012) proposed another version of the discrete Lindley distribution. Chakraborty and Chakravarty (2012) proposed the discrete gamma distribution and applied it to fit the reliability function of real discrete failure time data and also for modeling probability distribution of count data. Al-Huniti and Al-Dayian (2012) introduced the discrete Burr III distribution and discussed its distributional properties and

reliability characteristics along with a graphical description. Nekoukhou et al. (2013) got the discrete generalized exponential distribution, estimated its parameters, and examined its model using a real set of data related to accidents of women working on Shells for five weeks. Khorashadizadeh (2013) proposed the discrete log-logistic distribution as a special case of the discrete Burr XII distribution. Hussain and Ahmad (2014) introduced the discrete inverse Rayleigh distribution. Chakraborty and Chakravarty (2014) derived the discrete Gumbel distribution and applied it to fit three real life count data in relation to maximum flood discharges and annual maximum wind speeds. Almalki and Nadarajah (2014) got the discrete reduced modified Weibull Type II distribution as a special case of the telescopic distribution. Barbiero (2014) proposed the discrete skew Laplace distribution and applied its model to artificial and real data. Gómez-Déniz et al. (2014) generalized the SF of the normal case when $\mu = 0$ and $\sigma > 0$ using the Marshal-Olkin scheme and got the discrete half normal distribution which can also be regarded as a particular case from the discrete generalized gamma distribution of Chakraborty (2015a)[see, Chakraborty (2015b)]. Chakraborty (2015a) derived the discrete generalized gamma distribution, discussed the failure rate (fr) behavior for the distribution and demonstrated its application in two real life count data sets. Another version of the discrete Pareto distribution was given by Kozubowski et al. (2015) from which they obtained the discrete truncated Pareto distribution. Chakraborty and Chakravarty (2015) introduced the two-sided power distribution, studied its distributional and reliability properties, and estimated its parameters using ML and moment methods. Alamatsaz et al. (2016) proposed the pmf of the discrete generalized Rayleigh distribution. Chakraborty and Chakravarty (2016) proposed the discrete logistic distribution and applied it to model real life count data.

Approach 2: Discretizing using the density function:

Suppose that X is a continuous random variable with pdf $f(x)$, $-\infty < x < \infty$, then, the pmf of the discretized random variable where, $Y = [X]$ is derived as:

$$P(Y = y) = \frac{f(y)}{\sum_{y=-\infty}^{\infty} f(y)}, \quad y = 0, \pm 1, \pm 2, \dots \quad \dots (2)$$

For example, If X is an exponential random variable with the pdf $f(x) = \theta e^{-\theta x}$, $x > 0$, then the pmf of its discrete version is given by:

$$P(Y = y) = \frac{f(y)}{\sum_{y=-\infty}^{\infty} f(y)} = \frac{\theta e^{-\theta y}}{\sum_{y=0}^{\infty} \theta e^{-\theta y}} = \frac{e^{-\theta y}}{\sum_{y=0}^{\infty} e^{-\theta y}}, \quad y = 0, 1, \dots$$

But, $\sum_{k=0}^{\infty} x^k = \frac{1}{(1-x)}$, $\Rightarrow \sum_{y=0}^{\infty} e^{-\theta y} = \frac{1}{(1-e^{-\theta})} = (1-e^{-\theta})^{-1}$, then:

$$P(Y = y) = \frac{e^{-\theta y}}{(1-e^{-\theta})^{-1}} = e^{-\theta y} (1-e^{-\theta}), \quad y = 0, 1, \dots$$

therefore, $P(Y = y) = (1-e^{-\theta}) e^{-\theta y} = (1-q)q^y$, $y = 0, 1, 2, \dots$

This is the geometric distribution where, $q = e^{-\theta}$.

Another example is the discrete analogue of the normal distribution with mean μ and variance σ^2 which was derived by Kemp (1997) is given by the following pmf:

$$P(Y = y) = P_y = \frac{\theta^y q^{[y(y-1)/2]}}{\sum_{y=-\infty}^{\infty} \theta^y q^{[y(y-1)/2]}}, \quad y = 0, \pm 1, \pm 2, \dots$$

where, $\theta = e^{(2C_3\mu - C_2 - C_3)} > 0$, $0 < q = e^{-2C_3} < 1$, and the normalizing constant $\frac{1}{\sum_{y=-\infty}^{\infty} \theta^y q^{[y(y-1)/2]}} = e^{(-1 - C_1 - C_3\mu^2)}$, where, C_1 , C_2 , and C_3 are the constants in the Lagrangian

L in the Lagrange's method of undetermined multipliers:

$$L \equiv - \sum_{y=-\infty}^{\infty} P_y \ln P_y - C_1 \left[\sum_{y=-\infty}^{\infty} P_y - 1 \right] - C_2 \left[\sum_{y=-\infty}^{\infty} y P_y - \mu \right] - C_3 \left[\sum_{y=-\infty}^{\infty} (y-\mu)^2 P_y - \sigma^2 \right],$$

and $\ln P_y + 1 + C_1 + C_2 y + C_3 (y - \mu)^2 = 0$

Using formula (2), Siromoney (1964) suggested a discrete general Dirichlet distribution and applied it to model the frequency distribution of the length of wet spells during the period 1932-

1962 in Tambaram in southern India. Lai and Wang (1995) suggested the discrete power function distribution and used it for fitting discrete failure times and a type of mortality data. The discrete analogue of the normal distribution with mean μ and variance σ^2 was derived by Kemp (1997) and was applied by Harris et al.(2001) in dynamic analysis of rural establishment count data. Sato et al. (1999) proposed the discrete exponential distribution and applied it to model defect count distribution in semiconductor deposition equipment and defect count distribution per chips. The convolution of the discrete exponential distribution was used by Sato et al. (1999) to give a discrete gamma distribution which was applied again by them to model defect count distribution in semiconductor deposition equipment and defect count distribution per chips. The discrete Gaussian exponential (also called the discrete log normal) distribution was proposed by Bi et al. (2001). The discrete half normal distribution was derived by Kemp (2006) as the maximum entropy distribution on $(0, 1, \dots)$ with specified mean and variance. It may be regarded as a special case of the discrete normal distribution of Kemp (1997). A discrete version of the Laplace distribution was derived by Inusah and Kozubowski (2006) and introduced by Andersen et al. (2013a) to estimate Y-STR haplotype frequencies, as described by Andersen et al. (2013b), demonstrating a number of examples using R package. Kozubowski and Inusah (2006) proposed the discrete skew Laplace distribution. Nekoukhou, et al. (2012) introduced a discrete analogue of the generalized exponential distribution and fitted its model to rank frequencies of graphemes in Slovene (a Slavic language). Lekshmi and Sebastian (2014) introduced the discrete generalized Laplace distribution. Nekoukhou et al. (2015) proposed the discrete analogue of the beta-exponential distribution and fitted it again to rank frequencies of graphemes in Slovene.

Approach 3: Discretizing by shifting the cumulative distribution function:

If X is a continuous random variable having a cumulative distribution function (cdf) $F_X(x)$, then the pmf of the discrete analogue Y is given by:

$$P(Y = y) = P(y) = F_T(y + \delta) - F_T[y - (1 - \delta)], y = 0, \pm 1, \pm 2, \dots, \pm i; \quad 0 < \delta \leq 1,$$

this pmf retains the form of the original cdf, except for a shift in the location by δ . [Roy and Dasgupta (2001)]. For example, If X is an exponential random variable with the cdf $F(x) = 1 - e^{-\theta x}$, $x > 0$, then the pmf of its discrete version is given by:

$$\begin{aligned}
 P(Y = y) &= F_T(y + \delta) - F_T[y - (I - \delta)] \\
 &= [1 - e^{-\theta(y+\delta)}] - \{1 - e^{-\theta[y-(I-\delta)]}\} \\
 &= e^{-\theta[y-(I-\delta)]} - e^{-\theta(y+\delta)} \\
 &= e^{-\theta[y-(I-\delta)]}(1 - e^{-\theta}), \quad y = 0, 1, \dots
 \end{aligned}$$

For $\delta = I$, this pmf reduces to the geometric distribution.

A standard form of this pmf is considered to reduce the range of Y . For example if X^* is a random variable having a normal distribution with mean μ and standard deviation σ , that is, $X^* \sim N(\mu, \sigma^2)$, then, the discrete counterpart of X^* is: $Y^* = \mu + \sigma \cdot Z^*$, where, Z^* is the discrete counterpart of $Z \sim N(0, 1)$.

Using this method of discretizing, Perry and Taylor (1985) proposed the discrete analogue of the Ade's distribution and fitted it to 22 entomological data sets. Roy and Dasgupta (2001) proposed this method in order to evaluate the reliability of complex systems under stress-strength model where available analytic approaches fail to provide a closed solution. According to this method, the cdf of Y (the resulted discrete random variable) retains the form of the cdf of X (the original continuous random variable) and support of Y is determined from a subset of the range of X .

Approach 4: Discretizing using the failure rate:

Stein and Dattero (1984) proposed this method to obtain a discrete version of the Weibull distribution. According to this method, the fr of Y retains the form of the fr of X .

Given the fr of the continuous random variable X , $h_X(x)$, then the pmf of Y is:

$$\begin{aligned}
 P(Y = y) &= [1 - h_X(0)][1 - h_X(I)] \cdots [1 - h_X(y - I)] \{1 - [1 - h_X(y)]\} \\
 &= [1 - h_X(0)][1 - h_X(I)] \cdots [1 - h_X(y - I)] h_X(y) \\
 &= \begin{cases} h_X(0), & y = 0 \\ [1 - h_X(0)][1 - h_X(I)] \cdots [1 - h_X(y - I)] h_X(y), & y = 1, 2, \dots, d \\ 0, & \text{otherwise} \end{cases}
 \end{aligned}$$

where, d is determined such that $0 \leq h_X(x) < 1$, and the sf of Y is given by:

$$S_Y(y) = [1 - h_X(0)][1 - h_X(I)] \cdots [1 - h_X(y - I)], \quad y = 1, 2, \dots$$

Salvia and Bollinger (1982) and Padgett and Spurrier (1985) gave the following illustrations:

1. For constant fr $[h_X(y) = c]$, $0 < c \leq 1$, $y = 0, 1, \dots$, (the exponential distribution),

$$\begin{aligned} P(Y = y) &= [1 - h_X(0)][1 - h_X(1)] \cdots [1 - h_X(y-1)]h_X(y) \\ &= [(1-c)(1-c) \cdots (1-c)]c = c(1-c)^y, \quad y = 0, 1, \dots \end{aligned}$$

this is the geometric distribution.

2. For decreasing fr $[h_X(y) = \frac{c}{y+1}]$, $0 < c \leq 1$, $y = 0, 1, \dots$, then:

$$\begin{aligned} P(Y = y) &= [1 - h_X(0)][1 - h_X(1)] \cdots [1 - h_X(y-1)]h_X(y) \\ &= [(1-c)(1-\frac{c}{2}) \cdots (1-\frac{c}{y})](\frac{c}{y+1}) \\ &= \frac{c(1-c)(2-c) \cdots (y-c)}{(y+1)!}, \quad y = 0, 1, \dots \end{aligned}$$

3. For increasing fr $[h_X(y) = 1 - \frac{c}{y+1}]$, $0 \leq c \leq 1$, $y = 0, 1, \dots$, then:

$$\begin{aligned} P(Y = y) &= [1 - h_X(0)][1 - h_X(1)] \cdots [1 - h_X(y-1)]h_X(y) \\ &= [c(\frac{c}{2}) \cdots (\frac{c}{y})][1 - (\frac{c}{y+1})] \\ &= \frac{(y-c+1)c^y}{(y+1)!}, \quad y = 0, 1, \dots \end{aligned}$$

To insure that $\sum_{y=0}^{\infty} P(Y = y) = 1$, each probability $P(Y = y)$ has to be multiplied by a positive normalizing constant. The functional form of the fr will not be affected by such a choice. [Chakraborty(2015b)].

Bracquemond and Gaudoin (2003) criticized the assumption of a bounded support and referred such models as not to be very useful practically, asking the question: How can one be sure that the underlying system will fail in less than c demands?.

Using this method of discretizing, Stein and Dattero (1984) discretize the Weibull distribution. Padgett and Spurrier (1985) obtained the discrete analogue of the Weibull distribution using the definition of the fr provided by Roy and Gupta (1992). Roy and Gosh (2009) got discrete analogues of the Rayleigh and the Lomax distributions and used the fr of the Lomax distribution to approximate the reliability of complex systems where exact determination of the survival probabilities are analytically intractable.

References:

- Alamatsaz, M. H., Dey, S., Dey, T., and Harandi, S. S. (2016). Discrete generalized Rayleigh distribution. *Pakistan Journal of Statistics*, **32**, **1**, 1-20.
- Al-Huniti, A., Al-Dayian, G. R. (2012). Discrete Burr Type III Distribution. *American Journal of Mathematics and Statistics*, **2**, **5**, 145-152.
- Almalki, S. J., and Nadarajah, S. (2014). A new discrete modified Weibull distribution. *IEEE Transactions on Reliability*, **63**, **1**, 68-80.
- Andersen, M. M., Ereksin, P. S., Morling, N. (2013a). A gentle introduction to the discrete Laplace method for estimating Y-STR haplotype frequencies. arXiv: 1304.2129v1 [stat. AP], 8. Apr.2013.
- Andersen, M. M., Eriksen, P. S., Morling, N. (2013b). The discrete Laplace exponential family and estimation of Y-STR haplotype frequencies. *Journal of Theoretical Biology*, **329**, 39-51.
- Bakouch, H. S., Jazi, M. A., and Nadarajah, S. (2012). A new discrete distribution. *Statistics*, **48**, **1**, 200-240.
- Barbiero, A. (2014). An alternative discrete skew Laplace distribution. *Statistical Methodology*, **16**, 47-67.
- Bebbington, M., Lai, C. D., Wellington, M., and Zitikis, R. (2012). The discrete additive Weibull distribution: a bathtub shaped hazard for discontinuous failure data. *Reliability Engineering and System Safety*, **106**, 37-44.

- Bi, Z., Faloutsos, C., Korn, F. (2001). The “DGX” distribution for mining massive skewed data. 7th Conference of Knowledge discovery and data mining, San Francisco.
- Bracquemond, C., and Gaudion, O. (2003). A Survey on discrete lifetime distributions. *International Journal of Reliability, Quality and Safety Engineering*, **10, 1**, 69-98.
- Chakraborty, S. (2015a). A new discrete distribution related to generalized gamma distribution and Its properties. *Communications in Statistics-Theory and Methods*, **44, 8**, 1691-1705.
- Chakraborty, S. (2015b). Generating discrete analogues of continuous probability distributions-A survey of methods and constructions. *Journal of Statistical Distributions and Applications*, **2, 6**. Published on: 5 August 2015.
- Chakraborty, S., Chakravarty, D. (2012). Discrete gamma distributions: properties and parameter estimations. *Communications in Statistics-Theory and Methods*, **41, 18**, 3301-3323
- Chakraborty, S., Chakravarty, D. (2016). A new discrete probability distribution with integer support on $(-\infty, \infty)$. *Communications in Statistics-Theory and Methods*, **45, 2**, 492-505.
- Gómez-Déniz, E. (2010). Another generalization of the geometric distribution. *Test*, **19, 2**, 399–415.
- Gómez-Déniz, E., Calderin-Ojeda. E. (2011). The discrete Lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation*, **81, 11**, 1405-1416.
- Gómez-Déniz, E., Vázquez-Polo, F. J., García-García, V. (2014). A discrete version of the half-normal distribution and its generalization with applications. *Statistical Papers*, **55, 2**, 497–511.
- Harris, T. R., Shonkwiler, J. S., Lin, Y. (2001). Application of discrete normal distribution for dynamic rural retail sector analysis. Selected paper at the annual AAEA meeting, Chicago, Illinois, August 5-8.
- Hussain, T., and Ahmad, M. (2014). Discrete inverse Rayleigh distribution. *Pakistan Journal of Statistics*. **30, 2**, 203-222.
- Inusah, S., Kozubowski, T. J. (2006). A discrete analogue of the Laplace distribution. *Journal of Statistical Planning and Inference*, **136**, 1090-1102.
- Jazi, M. A., Lai, C. D., Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. *Statistical Methodology*. **7**, 121-132.

- Kemp, A. W. (1997). Characterizations of a discrete normal distribution. *Journal of Statistical Planning and Inference*, **63**, 223-229.
- Kemp, A. W. (2006). The discrete half-normal distribution. *International Conference on Mathematical and Statistical Modeling in Honor of Enrique Castillo*, June 28-30.
- Khorashadizadeh, M., Roknabadi, A. H. R., and Borzadaran, G. R. M. (2013). Characterization of life distributions using log-odds rate in discrete ageing. *Communications in Statistics-Theory and Methods*, **42**, **1**, 67-87.
- Kozubowski, T. J., Inusah, S. (2006). A skew Laplace distribution on integers. *Annals of the Institute of Statistical Mathematics*, **58**, **3**, 555-571.
- Krishna, H., Pundir, P. S. (2007). Discrete Maxwell distribution. *InterStat*. Available at: <http://interstat.statjournals.net/YEAR/2007/articles/0711003.pdf>.
- Krishna, H., Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology*, **6**, 177-188.
- Lai, C. D., and Wang, D. Q. (1995). A finite range discrete life distribution. *International Journal of Reliability, Quality and Safety Engineering*, **2**, **2**, 147-160.
- Lekshmi, S., Sebastian, S. (2014). A skewed generalized discrete Laplace distribution. *International Journal of Mathematics and Statistics Invention (IJMSI)*, **2**, **3**, 95-102.
- Nakagawa, T., Osaki, S. (1975). The discrete Weibull distribution. *IEEE Transactions on Reliability*, **R- 24**, **5**, 300-301.
- Nekoukhou, V., Alamatsaz, M. H., Bidram, H. (2012). A discrete analog of the generalized exponential distribution. *Communications in Statistics-Theory and Methods*, **41**, **11**, 2000-2013.
- Nekoukhou, V., Alamatsaz, M. H., Bidram, H. (2013). Discrete generalized exponential distribution of a second type. *Statistics*, **47**, **4**, 876-887.
- Nekoukhou, V., Alamatsaz, M. H., Bidram, H., Aghajani, A. H. (2015). Discrete beta-exponential distribution. *Communications in Statistics-Theory and Methods*, **44**, **10**, 2079-2091.
- Nooghabi, M. S., Borzadaran, G. R. M., and Roknabadi, A. H. R. (2011). Discrete modified Weibull distribution. *Metron*, **69**, **2**, 207-222.
- Padgett, W. J., and Spurrier, J. D. (1985). Discrete failure models. *IEEE Transactions on Reliability* **34**, **3**, 253-256.

- Perry, J. N., Taylor, L. R. (1985). Adès: New ecological families of species-specific frequency distribution that describes repeated spatial samples with an intrinsic power law variance-mean property. *Journal of Animal Ecology*, **54**, **3**, 931-953.
- Roknabadi, A. H. R., Borzadaran, G. R. M., and Khorashadizadeh, M. (2009). Some aspects of discrete telescopic hazard rate functions in telescopic families. *Economic Quality Control*, **24**, **1**, 35–42.
- Roy, D. (2003). The discrete normal distribution. *Communications in Statistics-Theory and Methods*, **32**, **10**, 1871-1883.
- Roy, D. (2004). Discrete Rayleigh distribution. *IEEE Transactions on Reliability*, **53**, **2**, 255-260.
- Roy, D., and Dasgupta, T. (2001). A discretizing approach for evaluating reliability of complex Systems under stress-strength model. *IEEE Transactions on Reliability*, **50**, **2**, 145-150.
- Salvia, A. A., and Bollinger, R. C. (1982). On discrete hazard functions. *IEEE Transaction on Reliability*. **R-31**, **5**, 458-459.
- Sato, H., Ikota, M., Aritoshi, S., Masuda, H. (1999). A new defect distribution meteorology with a consistent discrete exponential formula and its applications. *IEEE Transactions on Semiconductor Manufacturing*, **12**, **4**, 409-418.
- Siromoney, G. (1964). The general Dirichlet's series distribution. *Journal of the Indian Statistical Association*, **2&3**, 69-74.
- Stein, W. E., Dattero, R. (1984). A new discrete Weibull distribution. *IEEE Transactions on Reliability*, **R- 33**, **2**, 196-197.
- Xie, M., Tang, Y., and Goh, T. N. (2002). A modified Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering & System Safety*, **76**, **3**, 279–285.

Estimation for Birnbaum-Saunders Distribution in Constant-Stress Accelerated Life Test with Hybrid Censoring

Mohamed S. Hamouda *

Abstract:

This paper introduces the maximum likelihood estimation for the Birnbaum-Saunders distribution based on hybrid censoring in the constant-stress accelerated life test with power law accelerated form. Maximum likelihood estimates (MLEs) of an accelerated life tests model are derived. Prediction of the scale parameter under usual conditions is obtained. Moreover, the reliability function at a certain mission time under the same conditions is also predicted. A Monte Carlo Simulation study is carried out to study the precision of the MLEs for the parameters involved.

Keywords and Phrases: Accelerated life testing; Reliability; Constant stress; Birnbaum-Saunders distribution, Maximum likelihood estimation; Inverse power law model; Hybrid censoring

1- Introduction

In industrial experiments, most of products and materials with high reliability are difficult to obtain information about their lifetime under normal conditions. It is often very costly and time consuming. To avoid these drawbacks and improve

* Department of Statistics, Faculty of Commerce, Aswan University Aswan ,Egypt
Email: msattar_statistic@yahoo.com

the efficiency of experiment, one of the choices may be to resort to Accelerated life testing (ALT) which is often used in order to induce failures of very high reliability devices. In such tests, the units are tested under accelerated conditions (stress) that is extremely more severe than normal use conditions encountered in practice. So, the time necessary to test a sample of such devices under normal conditions very costly, take a long time. So, ALT save time and cost. Stress would induce early failure of the tested units. Over stress may be in the form of temperature, voltage, pressure, load, humidity, vibration,...etc, or some combination of them.

As Nelson [1990] indicated, the stress can be applied in various ways, commonly used methods are constant-stress and step-stress. In a constant stress accelerated test, each unit is run at a prespecified constant stress level which does not vary with time. In use, most products such as semiconductors and microelectronics, capacitors, lamps, ...etc, run at a constant stress. This type of stress is widely used and preferred because the stress is constant in most applications, it is much easier to apply and quantify constant stress and models for constant stress are available, widely publicized and empirically verified.

Determining a relationship between stress and scale parameter of the lifetime distribution, extrapolation to the normal use conditions will be carried. This relationship is known as the acceleration model. It is assumed that changing the stress from one level to another affects the value of the parameter only and

not the functional form of the lifetime distribution. This is a major assumption of ALT.

Constant stress ALT, was applied by several authors. Among others, are Owen [1997], Aly [1997], Abdel-Ghaly et al. [1998] , Al-Hussaini and Abd-Hamid[2004,2006] andWatkins and John [2008].

Birnbaum and Saunders [1969] proposed a two-parameter failure time distribution for fatigue failure caused under cyclic loading. Fatigue failure based on a model which assumes that failure is due to the development and growth of a dominant crack. This distribution is known as the two-parameter Birnbaum-Saunders (BS) distribution or as the fatigue life distribution. As Artur and Silvia [2011] indicated that, this distribution is an attractive alternative to the Weibull, Gamma and Log-normal models, since its derivation considers the basic characteristics of the fatigue process.

Statistical analysis for the BS distribution has been limited to complete data by several authors,(for example, see Chang and Tang [1993], Owen [1997], Dupuis and Mills [1998], Xu and Tang [2010]).

little works have been done based on censored data, (for example, see Rieck [1995], Kevin [1999] ,Jeng [2003], Ng et al. [2006],Wang et al[2006], Xu and Tang [2011],Artur and Silvia [2011]) and Attia et al. [2013a,b] , Hamouda [2014,2015]) and Sun and Shi [2016].

The estimates from the censored data are less accurate than those from the complete data. However, this is more than offset by the reduced test time and expense. The most common censoring schemes are type-I and type-II censoring. Consider n units placed on life test, in conventional type-I censoring, the experiment continues up to a pre-specified time t . But in conventional type-II censoring, the experiment is terminated after a pre-chosen number of units $r < n$ of failures. The mixture of type-I and type-II censoring schemes is known as hybrid censoring scheme, which was introduced by Epstein (1954), and it can be described as follows. Consider the following life testing experiment in which n units are placed on test. The lifetimes of the sample units are assumed to be independent and identically (i.i.d) random variables. The test is terminated when a pre-chosen $r < n$, out of n units have failed, or when a prespecified time, η , has been reached. In other words, the life test is terminated at a random time $t_{(R_j)} = \min(t_{(r)}, \eta)$. It is also usually assumed that the failed units are not replaced in the experiment. From now on, we refer to this hybrid censoring scheme as type-I hybrid censoring scheme (Type-I HCS).

In this paper, we discuss the MLEs of constant-stress ALT for two-parameter BS distribution under Type-I HCS. This model is described in details in Section 2, and the value of the scale parameter and the reliability function at a mission time t_0 under usual stress are predicted. Section 3 explains the simulation studies for illustrating the theoretical results. Finally, conclusion are included in Section 4.

2. Model description and maximum likelihood estimation

Assume the lifetime T is assumed to have a two - parameter BS distribution with the shape and scale parameters α and β respectively. So, the probability density function of T is

$$f(t; \alpha, \beta) = \frac{1}{2\alpha\sqrt{2\pi\beta}} \frac{(t+\beta)}{\left(\frac{t}{\beta}\right)^{\frac{3}{2}}} \exp\left[\frac{-1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right], t > 0, \alpha, \beta > 0 \quad [1]$$

The scale parameter (β) is the median of the BS distribution which has a wide use in reliability studies. The cumulative distribution function is as follows:

$$F(t; \alpha, \beta) = \Phi\left(\frac{1}{\alpha}\sqrt{\left(\frac{t}{\beta}\right)} - \sqrt{\left(\frac{\beta}{t}\right)}\right), \quad [2]$$

where $\Phi(\cdot)$ denotes the standard normal distribution function. And the reliability function of BS distribution in (1) takes the form:

$$R(t) = 1 - F(t) = 1 - \Phi\left(\frac{1}{\alpha}\left(\sqrt{\left(\frac{t}{\beta}\right)} - \sqrt{\left(\frac{\beta}{t}\right)}\right)\right) \quad [3]$$

The scale parameters β and the stress V are related by the inverse power law model, which is described in Mann et al. [1974].

$$\beta_j = CV_j^{-P} \quad , C, P > 0 \quad j = 1, 2, \dots, k \quad [4]$$

Where C is the constant of proportionality, and P is the power of the applied stress, which are the parameters of this model.

Assume that the life testing experiment is conducted under K higher stress $V_j, j=1, \dots, k$ and assume that V_u be the normal use conditions such that $V_u < V_1 < V_2 < \dots < V_k$. At each V_j there are n_j units put on test. The total number of units in the experiments is $N = \sum_{j=1}^k n_j$. When a Type-I HCS is adopted at each stress level, the experiment terminates once the number of pre-chosen number r_j of failures, or pre-determined time η_j whichever is reached first.

When $t_{(1j)} \leq t_{(2j)} \leq \dots \leq t_{(r_j j)}$ denote the observed failure times if $t_{(r_j j)} < \eta_j$ and $t_{(1j)} \leq t_{(2j)} \leq \dots \leq t_{(d_j j)}$ denote the observed failure times if $t_{(d_j j)} < \eta_j, d_j < r_j$ and the $((d+1)_j)^{th}$ failure does not take place before time point η_j . So, sampling according to the type-I HCS is terminated at:

$$t_{(R_j)} = \min(t_{(r_j)}, \eta_j). \quad [5]$$

Such that : $R_j = r_j$ Case I, otherwise $R_j = d_j$ case II,

Where d_j denotes the number of failures observed before time point η_j .

The lifetime t_{ij} at stress $V_j, i=1, \dots, R_j$ and $j=1, \dots, k$, are assumed to be two-parameter BS distribution. The likelihood function of the experiment is considered to have the following form:

$$L(\alpha, C, P) = \prod_{j=1}^k \left[\frac{n_j!}{(n_j - R_j)!} \right] \left[\prod_{i=1}^{R_j} f(t_{ij}; \alpha, C, P) \right] \left[1 - F(t_{(R_j j)}) \right]^{n_j - R_j}$$

$$L(\alpha, C, P) = \prod_{j=1}^k \frac{n_j!}{(n_j - R_j)!} \left[\prod_{i=1}^{R_j} \frac{1}{2\alpha \sqrt{2\pi(CV_j^{-P})}} \frac{(t_{ij} + CV_j^{-P})}{(t_{ij})^2} \exp\left(\frac{-1}{2\alpha^2} \left(\frac{t_{ij}}{CV_j^{-P}} + \frac{CV_j^{-P}}{t_{ij}} - 2\right)\right) \right] \left[1 - \Phi\left(\frac{1}{\alpha} \left(\sqrt{\frac{t_{(R_jj)}}{CV_j^{-P}}} - \sqrt{\frac{CV_j^{-P}}{t_{(R_jj)}}}\right)\right) \right]^{n_j - R_j} \quad [6]$$

From a statistical point of view, the method of maximum likelihood yields estimators with good statistical properties. Since these estimators do not always exist in closed form, numerical techniques are used to compute them.

This section discusses the process of obtaining the maximum likelihood estimates of the parameters α , C and β based on type-I HCS. Both point and interval estimations of the parameters are derived.

Using $\ln L$ to denote the natural logarithm of $L(\alpha, C, P)$, then we have

$$\ln L = \text{constant} - \sum_{j=1}^k R_j \left[\ln \alpha + \frac{\ln C}{2} \right] + \frac{P}{2} \sum_{j=1}^k R_j \ln V_j + \sum_{j=1}^k \sum_{i=1}^{R_j} \ln(t_{ij} + CV_j^{-P}) - \frac{2}{3} \sum_{j=1}^k \sum_{i=1}^{R_j} \ln(t_{ij}) - \frac{1}{2\alpha^2} \sum_{j=1}^k \sum_{i=1}^{R_j} \left(\frac{t_{ij}}{CV_j^{-P}} + \frac{CV_j^{-P}}{t_{ij}} - 2 \right) + \sum_{j=1}^k (n_j - R_j) W(R_jj) \quad [7]$$

where $\text{constant} = \sum_{j=1}^k R_j (\ln n_j! - \ln(n_j - R_j)!) - k R_j \ln(2\sqrt{2\pi})$,

$$W(R_jj) = \ln \left[1 - \Phi \left(\omega_{(R_jj)} \right) \right], \quad \omega_{(R_jj)} = \frac{1}{\alpha} \left(\sqrt{\frac{t_{(R_jj)}}{CV_j^{-P}}} - \sqrt{\frac{CV_j^{-P}}{t_{(R_jj)}}} \right)$$

MLEs of α , C and P are obtained by getting the partial derivatives of $\ln L$ with respect to α , C and P respectively, as follows:

$$\frac{\partial \ln L}{\partial \alpha} = - \sum_{j=1}^k R_j \left(\frac{1}{\alpha} \right) + \left(\frac{1}{\alpha^3} \right) \sum_{j=1}^k \sum_{i=1}^{R_j} \left(\frac{t_{ij}}{CV_j^{-P}} + \frac{CV_j^{-P}}{t_{ij}} - 2 \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial W(R_jj)}{\partial \alpha} \quad [8]$$

$$\frac{\partial \ln L}{\partial C} = - \sum_{j=1}^k R_j \left(\frac{1}{2C} \right) + \sum_{j=1}^k \sum_{i=1}^{R_j} \frac{V_j^{-P}}{(t_{ij} + CV_j^{-P})} - \frac{1}{2\alpha^2} \sum_{j=1}^k \sum_{i=1}^{R_j} \left(\frac{V_j^{-P}}{t_{ij}} - \frac{t_{ij}}{C^2 V_j^{-P}} \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial W(R_jj)}{\partial C} \quad [9]$$

$$\frac{\partial \ln L}{\partial P} = \frac{1}{2} \sum_{j=1}^k R_j \ln V_j - \sum_{j=1}^k \ln V_j \sum_{i=1}^{R_j} \frac{CV_j^{-P}}{(t_{ij} + CV_j^{-P})} - \frac{1}{2\alpha^2} \sum_{j=1}^k \ln V_j \sum_{i=1}^{R_j} \left(\frac{t_{ij}}{CV_j^{-P}} - \frac{CV_j^{-P}}{t_{ij}} \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial W(R_{jj})}{\partial P} \quad [10]$$

For notational convenience, let us denote the hazard function of the standard normal distribution as: $H(y) = \frac{\phi(y)}{1-\Phi(y)}$

$$\frac{\partial W(R_{jj})}{\partial \alpha} = H(\omega_{(R_{jj})}) * \frac{1}{\alpha} (\omega_{(R_{jj})}) \quad [11]$$

$$\frac{\partial W(R_{jj})}{\partial C} = H(\omega_{(R_{jj})}) * \frac{1}{2\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{C^3 V_j^{-P}}} + \sqrt{\frac{V_j^{-P}}{C t_{(R_{jj})}}} \right) \quad [12]$$

$$\frac{\partial W(R_{jj})}{\partial P} = H(\omega_{(R_{jj})}) * \frac{-(\ln V_j)}{2\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{C V_j^{-P}}} + \sqrt{\frac{C V_j^{-P}}{t_{(R_{jj})}}} \right) \quad [13]$$

To obtain $\hat{\alpha}$, \hat{C} and \hat{P} at which the log likelihood function is maximized, equating the equations (8-10) to zero. Since the closed form solution to these equations do not exist. The numerical iteration will be used to solve these equations. To estimate the variance-covariance matrix of the estimated parameters, we use the second derivatives of the logarithm of the likelihood function defined in equation (6). The second derivatives are used to get the information matrix (F), and hence the inverse of F is the asymptotic variance-covariance matrix. The Fisher-information can be expressed in the form:

$$F = \begin{bmatrix} \frac{-\partial^2 \ln L}{\partial \alpha^2} & \frac{-\partial^2 \ln L}{\partial \alpha \partial C} & \frac{-\partial^2 \ln L}{\partial \alpha \partial P} \\ \frac{-\partial^2 \ln L}{\partial C \partial \alpha} & \frac{-\partial^2 \ln L}{\partial C^2} & \frac{-\partial^2 \ln L}{\partial C \partial P} \\ \frac{-\partial^2 \ln L}{\partial P \partial \alpha} & \frac{-\partial^2 \ln L}{\partial P \partial C} & \frac{-\partial^2 \ln L}{\partial C^2} \end{bmatrix} \quad [14]$$

The inverse of F is the asymptotic variance-covariance matrix of $(\hat{\alpha}, \hat{C}$ and $\hat{P})$.

$$v = F^{-1} \tag{15}$$

Then, the second partial derivates are given as follows:

$$\frac{\partial^2 \ln L}{\partial \alpha^2} = \sum_{j=1}^k R_j \frac{1}{\alpha^2} - \frac{3}{\alpha^4} \sum_{j=1}^k \sum_{i=1}^{R_j} \left(\frac{t_{ij}}{CV_j^{-P}} + \frac{CV_j^{-P}}{t_{ij}} - 2 \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial^2 W(R_{jj})}{\partial \alpha^2} \tag{16}$$

$$\frac{\partial^2 \ln L}{\partial C^2} = \sum_{j=1}^k R_j \frac{1}{2C^2} - \sum_{j=1}^k \sum_{i=1}^{R_j} \frac{V_j^{-2P}}{(t_{ij} + CV_j^{-P})^2} - \frac{1}{\alpha^2} \sum_{j=1}^k \sum_{i=1}^{R_j} \left(\frac{t_{ij}}{C^3 V_j^{-P}} \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial^2 W(R_{jj})}{\partial C^2} \tag{17}$$

$$\frac{\partial^2 \ln L}{\partial P^2} = C \sum_{j=1}^k (\ln V_j)^2 \sum_{i=1}^{R_j} \left(\frac{t_{ij} V_j^{-P}}{(t_{ij} + CV_j^{-P})^2} \right) - \frac{1}{2\alpha^2} \sum_{j=1}^k (\ln V_j)^2 \sum_{i=1}^{R_j} \left(\frac{t_{ij}}{CV_j^{-P}} + \frac{CV_j^{-P}}{t_{ij}} \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial^2 W(R_{jj})}{\partial P^2} \tag{18}$$

$$\frac{\partial^2 \ln L}{\partial \alpha \partial C} = \frac{1}{\alpha^3} \sum_{j=1}^k \sum_{i=1}^{R_j} \left(\frac{V_j^{-P}}{t_{ij}} - \frac{t_{ij}}{C^2 V_j^{-P}} \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial^2 W(R_{jj})}{\partial \alpha \partial C} \tag{19}$$

$$\frac{\partial^2 \ln L}{\partial \alpha \partial P} = \frac{1}{\alpha^3} \sum_{j=1}^k (\ln V_j) \sum_{i=1}^{R_j} \left(\frac{t_{ij}}{CV_j^{-P}} - \frac{CV_j^{-P}}{t_{ij}} \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial^2 W(R_{jj})}{\partial \alpha \partial P} \tag{20}$$

$$\frac{\partial^2 \ln L}{\partial C \partial P} = - \sum_{j=1}^k \ln V_j \sum_{i=1}^{R_j} \frac{t_{ij} V_j^{-P}}{(t_{ij} + CV_j^{-P})^2} + \frac{1}{2\alpha^2} \sum_{j=1}^k (\ln V_j) \sum_{i=1}^{R_j} \left(\frac{V_j^{-P}}{t_{ij}} + \frac{t_{ij}}{C^2 V_j^{-P}} \right) + \sum_{j=1}^k (n_j - R_j) \frac{\partial^2 W(R_{jj})}{\partial C \partial P} \tag{21}$$

Where $\frac{\partial^2 W(R_{jj})}{\partial \alpha^2} = H'(\omega_{(R_{jj})}) * \frac{1}{\alpha}(\omega_{(R_{jj})}) - H(\omega_{(R_{jj})}) * \frac{1}{\alpha^2}(\omega_{(R_{jj})})$ [22]

$$\frac{\partial^2 W(R_{jj})}{\partial C^2} = H'(\omega_{(R_{jj})}) \frac{1}{2\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{C^3 V_j^{-P}}} + \sqrt{\frac{V_j^{-P}}{C t_{(R_{jj})}}} \right) - H(\omega_{(R_{jj})}) \frac{1}{4\alpha} \left(3 \sqrt{\frac{t_{(R_{jj})}}{C^5 V_j^{-P}}} + \sqrt{\frac{V_j^{-P}}{C^3 t_{(R_{jj})}}} \right) \tag{23}$$

$$\frac{\partial^2 W(R_{jj})}{\partial P^2} = H'(\omega_{(R_{jj})}) \frac{-(\ln V_j)}{2\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{CV_j^{-P}}} + \sqrt{\frac{CV_j^{-P}}{t_{(R_{jj})}}} \right) + H(\omega_{(R_{jj})}) \frac{(\ln V_j)^2}{4\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{CV_j^{-P}}} - \sqrt{\frac{CV_j^{-P}}{t_{(R_{jj})}}} \right) \tag{24}$$

$$\frac{\partial^2 W(R_{jj})}{\partial \alpha \partial C} = H'(\omega_{(R_{jj})}) \frac{1}{\alpha}(\omega_{(R_{jj})}) - H(\omega_{(R_{jj})}) \frac{1}{2\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{C^3 V_j^{-P}}} + \sqrt{\frac{V_j^{-P}}{C t_{(R_{jj})}}} \right) \tag{25}$$

$$\frac{\partial^2 W(R_{jj})}{\partial \alpha \partial P} = H'(\omega_{(R_{jj})}) \frac{1}{\alpha}(\omega_{(R_{jj})}) + H(\omega_{(R_{jj})}) \frac{(\ln V_j)}{2\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{CV_j^{-P}}} + \sqrt{\frac{CV_j^{-P}}{t_{(R_{jj})}}} \right) \tag{26}$$

$$\frac{\partial^2 W(R_{jj})}{\partial C \partial P} = H'(\omega_{(R_{jj})}) \frac{1}{2\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{C^3 V_j^{-P}}} + \sqrt{\frac{V_j^{-P}}{C t_{(R_{jj})}}} \right) + H(\omega_{(R_{jj})}) \frac{(\ln V_j)}{4\alpha} \left(\sqrt{\frac{t_{(R_{jj})}}{C^3 V_j^{-P}}} - \sqrt{\frac{V_j^{-P}}{C t_{(R_{jj})}}} \right) \tag{27}$$

Getting Fisher-information matrix, it can be said that the MLE $\hat{\alpha}$, \hat{C} and \hat{P} have an asymptotic variance-covariance matrix defined by inverting the Fisher-information matrix.

The approximate confidence intervals of the parameters are derived based on the asymptotic distribution of the MLEs of the elements of the vector of unknown parameters $\Theta_{\mathcal{P}} = (\alpha, C, P)$. It is known that the asymptotic distribution of the MLEs of $\Theta_{\mathcal{P}} = (\alpha, C, P)$ is given as follows: $\left(\frac{(\hat{\alpha}-\alpha)}{\sqrt{\text{var}(\hat{\alpha})}}, \frac{(\hat{C}-C)}{\sqrt{\text{var}(\hat{C})}}, \frac{(\hat{P}-P)}{\sqrt{\text{var}(\hat{P})}} \right)$, which can be approximated by a standard normal distribution, see Miller [1981], where $\text{var}(\hat{\Theta})$ is estimated as the asymptotic variance, obtained from [15]. Then, the approximate 100 γ % two sided confidence interval for α , C , P are, respectively, given by:

$$\hat{\alpha} \pm Z_{\gamma/2} \sqrt{\text{var}(\hat{\alpha})}, \hat{C} \pm Z_{\gamma/2} \sqrt{\text{var}(\hat{C})} \text{ and } \hat{P} \pm Z_{\gamma/2} \sqrt{\text{var}(\hat{P})}, \quad [28]$$

where $Z_{\gamma/2}$ is the 100($\gamma/2$) standard normal percentile.

4. Simulation Studies

In this section simulation studies using the MathCAD(14) are conducted to investigate the performances of the maximum likelihood estimates through their estimators, their absolute relative bias (RABias), mean square error (MSE) and the confidence intervals of the MLEs. Using the invariance property of

MLEs, we can estimate the MLE of scale parameter β_j through the following equation

$$\hat{\beta}_j = \hat{C}V_j^{-\hat{P}} \quad , C, P > 0 \quad j = 1,2,3 \quad [29]$$

The Simulation procedures were described as follows:

Step1. 1000 random samples of sizes 90, 150, 210, 300 and 450 were generated from the BS distribution. Different initial values are selected for all sets of $(\alpha_0, C_0$ and $P_0)$. There are only three different levels of stress, $k=3$, the stress values are selected as $(V_1 = .9, V_2 = 1.4, V_3 = 1.9)$. $n_j = \frac{n}{3}$, and $r_j = 60\%n_j$ and $\eta_1 = 5, \eta_2 = 3, \eta_3 = 2$.

Step2. For all sample sizes, the parameters of the model are estimated under Type-I HCS, for all sets of $(\alpha_0, C_0$ and $P_0)$.

Step3. Equations 8,9 and 10 are solved by using the Numerical method. The estimate of scale parameter β_j was obtained from equation [29].

Step4. The estimators, RABias and MSE are tabulated for all sets of $(\alpha_0, C_0$ and $P_0)$ in Tables (1-3) based on Type-I HCS.

Step5. The confidence limit with confidence level $(1 - \gamma) = 0.95$ are tabulated for all sets of parameters in Tables (4-6) based on Type-I HCS.

Step6. Using the invariance property of MLEs, the MLE of the scale parameter β_u of BS distribution at usual stress $V_u = .5$, can be estimated by using the following equation:

$$\hat{\beta}_u = \hat{C} \cdot 5^{-\hat{P}} \quad , C, P > 0 \quad j = 1,2,3 \quad [30]$$

Also, the reliability function at mission time ($t_1 = 1.5, t_2 = 1.8, t_3 = 2.1$ and $t_4 = 2.4$) are predicted under the same usual condition for all sets of (α_0, C_0 and P_0) in Tables [7-9].

4. Conclusion

The performances of the estimators are investigated by simulation study and from results of Tables [1-6], it is observed the following :

- 1- The MLEs for all sets of initial values of parameters have good statistical properties for all sample sizes. As the sample size increases, the MSE of estimators decreases.
- 2- As the sample size increases, the interval of estimators decreases.
- 3-As the stress increases, it is evident that the MSE of the estimated scale parameter β_j tend to decrease.
- 4- As the value of α_0 decrease, it is evident that the MSE of estimates decrease.
- 5- It is evident that the MSE of the estimates do not rely on varying the values of P_0 and C_0 .

From results of Tables [7-9], it is observed the following:

- 1-As the sample size increases, the reliability function increases.
- 2- As the values of C_0 and P_0 increase at the same mission time, the reliability function increases.

3- There is an inverse proportional relationship between the shape parameter α_0 and the reliability function.

Table (1): The Estimates, RABias and MSE of the parameters under Type-I HCS

n	Parameters	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$					
		$\alpha = .25, C = 1.5, P = 1$			$\alpha = .5, C = 1.5, P = 1$		
		Estimators	RABias	MSE	Estimators	RABias	MSE
30	α	0.228	0.089	0.002	0.456	0.087	0.01
	C	1.49	0.006	0.011	1.49	0.005	0.045
	P	1.007	0.007	0.026	1.013	0.013	0.103
	β_1	1.66	0.004	0.019	1.667	0.003	0.077
	β_2	1.06	0.009	0.003	1.059	0.014	0.012
	β_3	0.788	0.009	0.003	0.783	0.007	0.014
90	α	0.242	0.033	0.0007	0.483	0.033	0.003
	C	1.49	0.002	0.004	1.49	0.003	0.014
	P	1.001	0.0006	0.008	0.993	0.007	0.033
	β_1	1.66	0.002	0.006	1.663	0.002	0.023
	β_2	1.07	0.003	0.001	1.068	0.002	0.004
	β_3	0.788	0.002	0.0004	0.792	0.003	0.005
150	α	0.246	0.016	0.002	0.492	0.016	0.002
	C	1.498	0.001	0.005	1.499	0.0009	0.008
	P	1.001	0.001	0.004	1.004	0.004	0.019
	β_1	1.665	0.0007	0.0006	1.667	0.0001	0.014
	β_2	1.069	0.001	0.0007	1.068	0.003	0.002
	β_3	0.788	0.002	0.0007	0.787	0.003	0.003
210	α	0.247	0.013	0.0003	0.494	0.012	0.001
	C	1.499	0.0009	0.002	1.499	0.0004	0.006
	P	1.001	0.0008	0.004	1.002	0.002	0.014
	β_1	1.665	0.0007	0.003	1.667	0.0003	0.010
	β_2	1.071	0.001	0.0004	1.07	0.002	0.001
	β_3	0.788	0.001	0.0005	0.788	0.002	0.002
300	α	0.248	0.009	0.0002	0.496	0.007	0.0008
	C	1.5	0.0002	0.001	1.5	0.0003	0.004
	P	1	0.0009	0.002	1.002	0.002	0.009
	β_1	1.001	0.0004	0.001	1.667	0.0003	0.007
	β_2	1.071	0.0002	0.0003	1.07	0.001	0.001
	β_3	0.789	0.00004	0.0003	0.788	0.002	0.001

Table (2): The Estimates, RABias and MSE of the parameters under Type-I HCS

n	Parameters	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$					
		$\alpha = .5, C = 1.5, P = 1$			$\alpha = .5, C = 2, P = 1$		
		Estimators	RABias	MSE	Estimators	RABias	MSE
30	α	0.456	0.087	0.01	0.453	0.094	0.009
	C	1.49	0.005	0.045	1.97	0.012	0.07
	P	1.013	0.013	0.103	1.009	0.009	0.104
	β_1	1.667	0.003	0.077	2.205	0.008	0.123
	β_2	1.059	0.014	0.012	1.402	0.019	0.021
	β_3	0.783	0.007	0.014	1.037	0.015	0.026
90	α	0.483	0.033	0.003	0.487	0.026	0.003
	C	1.49	0.003	0.014	1.99	0.0003	0.024
	P	0.993	0.007	0.033	1.002	0.002	0.033
	β_1	1.663	0.002	0.023	2.224	0.0009	0.123
	β_2	1.068	0.002	0.004	1.426	0.002	0.04
	β_3	0.792	0.003	0.005	1.052	0.0004	0.007
150	α	0.492	0.016	0.002	0.491	0.018	0.002
	C	1.499	0.0009	0.008	1.99	0.001	0.015
	P	1.004	0.004	0.019	1	0.00005	0.019
	β_1	1.667	0.0001	0.014	2.22	0.001	0.026
	β_2	1.068	0.003	0.002	1.425	0.002	0.004
	β_3	0.787	0.003	0.003	1.051	0.002	0.005
210	α	0.494	0.012	0.001	0.495	0.009	0.001
	C	1.499	0.0004	0.006	2.003	0.001	0.011
	P	1.002	0.002	0.014	1.003	0.003	0.014
	β_1	1.667	0.0003	0.010	2.227	0.002	0.018
	β_2	1.07	0.002	0.001	1.428	0.0002	0.003
	β_3	0.788	0.002	0.002	1.052	0.0004	0.003
300	α	0.496	0.007	0.0008	0.496	0.008	0.0009
	C	1.5	0.0003	0.004	1.999	0.0001	0.007
	P	1.002	0.002	0.009	1.004	0.003	0.009
	β_1	1.667	0.0003	0.007	2.223	0.0003	0.013
	β_2	1.07	0.001	0.001	1.426	0.002	0.002
	β_3	0.788	0.002	0.001	1.05	0.002	0.002

Table (3): The Estimates, RABias and MSE of the parameters under Type-I HCS

n	Parameters	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$					
		$\alpha = .5, C = 1.5, P = 1$			$\alpha = .5, C = 1.5, P = 1.5$		
		Estimators	RABias	MSE	Estimators	RABias	MSE
30	α	0.456	0.087	0.01	0.485	0.084	0.009
	C	1.49	0.005	0.045	1.48	0.009	0.042
	P	1.013	0.013	0.103	1.49	0.004	0.098
	β_1	1.667	0.003	0.077	1.75	0.007	0.079
	β_2	1.059	0.014	0.012	0.89	0.012	0.008
	β_3	0.783	0.007	0.014	0.57	0.004	0.007
90	α	0.483	0.033	0.003	0.485	0.04	0.003
	C	1.49	0.003	0.014	1.49	0.001	0.014
	P	0.993	0.007	0.033	1.5	0.004	0.034
	β_1	1.663	0.002	0.023	1.75	0.0008	0.027
	β_2	1.068	0.002	0.004	0.901	0.005	0.002
	β_3	0.792	0.003	0.005	0.57	0.004	0.002
150	α	0.492	0.016	0.002	0.491	0.019	0.002
	C	1.499	0.0009	0.008	1.498	0.001	0.008
	P	1.004	0.004	0.019	1.5	0.003	0.019
	β_1	1.667	0.0001	0.014	1.757	0.0001	0.016
	β_2	1.068	0.003	0.002	0.902	0.004	0.001
	β_3	0.787	0.003	0.003	0.57	0.004	0.001
210	α	0.494	0.012	0.001	0.495	0.011	0.001
	C	1.499	0.0004	0.006	1.495	0.003	0.005
	P	1.002	0.002	0.014	1.49	0.003	0.013
	β_1	1.667	0.0003	0.010	1.75	0.003	0.011
	β_2	1.07	0.002	0.001	0.904	0.002	0.001
	β_3	0.788	0.002	0.002	0.573	0.0003	0.001
300	α	0.496	0.007	0.0008	0.495	0.009	0.0009
	C	1.5	0.0003	0.004	1.496	0.002	0.004
	P	1.002	0.002	0.009	1.497	0.002	0.01
	β_1	1.667	0.0003	0.007	1.75	0.002	0.007
	β_2	1.07	0.001	0.001	0.904	0.002	0.0009
	β_3	0.788	0.002	0.001	0.573	0.0004	0.0008

Table (4): Confidence bounds of the parameters at confidence level .95 under Type-I HCS

n	Parameters	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$					
		$\alpha = .25, C = 1.5, P = 1$			$\alpha = .5, C = 1.5, P = 1$		
		Estimators	Variance	(Lower,Upper)	Estimators	Variance	(Lower,Upper)
30	α	0.228	0.003	(0.152,0.304)	0.456	0.01	(0.305,0.608)
	C	1.49	0.011	(1.320,1.680)	1.49	0.045	(1.167,1.818)
	P	1.007	0.026	(0.740,1.280)	1.013	0.103	(0.489,1.536)
90	α	0.242	0.0007	(0.193,0.291)	0.483	0.003	(0.385,0.582)
	C	1.49	0.004	(1.39,1.61)	1.496	0.014	(1.279,1.713)
	P	1.001	0.008	(0.83,1.17)	0.993	0.033	(0.657,1.328)
150	α	0.246	0.0004	(0.207,0.285)	0.492	0.002	(0.413,0.57)
	C	1.498	0.002	(1.41,1.59)	1.499	0.008	(1.325,1.67)
	P	1.001	0.005	(0.87,1.14)	1.004	0.019	(0.738,1.27)
210	α	0.247	0.0003	(0.214,0.28)	0.494	0.001	(0.427,0.561)
	C	1.499	0.001	(1.42,1.57)	1.499	0.006	(1.351,1.648)
	P	1.001	0.004	(0.886,1.11)	1.002	0.014	(0.775,1.229)
300	α	0.248	0.0002	(0.22,0.276)	0.496	0.0008	(0.44,0.553)
	C	1.5	0.0002	(1.44,1.563)	1.5	0.004	(1.374,1.625)
	P	1.001	0.0009	(0.904,1.097)	1.002	0.009	(0.811,1.193)

Table (5): Confidence bounds of the parameters at confidence level .95 under Type-I HCS

n	Parameters	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$					
		$\alpha = .5, C = 1.5, P = 1$			$\alpha = .5, C = 2, P = 1$		
		Estimators	Variance	(Lower,Upper)	Estimators	Variance	(Lower,Upper)
30	α	0.456	0.01	(0.305,0.608)	0.453	0.009	(0.301,0.605)
	C	1.49	0.045	(1.167,1.818)	1.975	0.07	(1.541,2.41)
	P	1.013	0.103	(0.489,1.536)	1.009	0.104	(0.486,1.532)
90	α	0.483	0.003	(0.385,0.582)	0.487	0.003	(0.388,0.586)
	C	1.496	0.014	(1.279,1.713)	1.999	0.024	(1.707,2.292)
	P	0.993	0.033	(0.657,1.328)	1.002	0.033	(0.664,1.339)
150	α	0.492	0.002	(0.413,0.57)	0.491	0.002	(0.412,0.569)
	C	1.499	0.008	(1.325,1.67)	1.99	0.015	(1.766,2.227)
	P	1.004	0.019	(0.738,1.27)	1	0.019	(0.734,1.266)
210	α	0.494	0.001	(0.427,0.561)	0.495	0.001	(0.428,0.592)
	C	1.499	0.006	(1.351,1.648)	2.003	0.011	(1.804,2.201)
	P	1.002	0.014	(0.775,1.229)	1.003	0.014	(0.776,1.23)
300	α	0.496	0.0008	(0.44,0.553)	0.496	0.0009	(0.439,0.552)
	C	1.5	0.004	(1.374,1.625)	1.99	0.007	(1.833,2.166)
	P	1.002	0.009	(0.811,1.193)	1.004	0.009	(0.813,1.194)

Table (6): Confidence bounds of the parameters at confidence level .95 under Type-I HCS

n	Parameters	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$					
		$\alpha = .5, C = 1.5, P = 1$			$\alpha = .5, C = 1.5, P = 1.5$		
		Estimators	Variance	(Lower,Upper)	Estimators	Variance	(Lower,Upper)
30	α	0.456	0.01	(0.305,0.608)	0.485	0.009	(0.305,0.611)
	C	1.49	0.045	(1.167,1.818)	1.48	0.042	(1.156,1.814)
	P	1.013	0.103	(0.489,1.536)	1.49	0.098	(0.967,2.022)
90	α	0.483	0.003	(0.385,0.582)	0.485	0.003	(0.386,0.584)
	C	1.496	0.014	(1.279,1.713)	1.48	0.014	(1.28,1.716)
	P	0.993	0.033	(0.657,1.328)	1.5	0.034	(1.171,1.844)
150	α	0.492	0.002	(0.413,0.57)	0.491	0.002	(0.412,0.569)
	C	1.499	0.008	(1.325,1.67)	1.49	0.008	(1.325,1.671)
	P	1.004	0.019	(0.738,1.27)	1.5	0.019	(1.24,1.771)
210	α	0.494	0.001	(0.427,0.561)	0.495	0.001	(0.428,0.562)
	C	1.499	0.006	(1.351,1.648)	1.49	0.005	(1.347,1.643)
	P	1.002	0.014	(0.775,1.229)	1.49	0.013	(1.268,1.722)
300	α	0.496	0.0008	(0.44,0.553)	0.495	0.0009	(0.439,0.551)
	C	1.5	0.004	(1.374,1.625)	1.49	0.004	(1.372,1.621)
	P	1.002	0.009	(0.811,1.193)	1.49	0.01	(1.302,1.688)

Table (7): Estimated shape parameter and the reliability function at $V_u = .5$ with Type-I HCS

	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$									
	$\alpha = .25, C = 1.5, P = 1$					$\alpha = .5, C = 1.5, P = 1$				
	30	90	150	210	300	30	90	150	210	300
$\hat{\beta}_u$	3.034	3.006	3.006	3.004	3.006	3.152	3.019	3.031	3.022	3.017
$\hat{R}_u(t_1)$.99	.99	.99	.99	.99	.89	.91	.92	.92	.92
$\hat{R}_u(t_2)$.96	.97	.97	.97	.98	.82	.83	.84	.85	.85
$\hat{R}_u(t_3)$.89	.92	.92	.92	.92	.74	.76	.76	.76	.77
$\hat{R}_u(t_4)$.78	.80	.81	.81	.81	.65	.67	.67	.67	.68

Table (8): Estimated shape parameter and the reliability function at $V_u = .5$ with Type-I HCS

	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$									
	$\alpha = .5, C = 1.5, P = 1$					$\alpha = .5, C = 2, P = 1$				
	30	90	150	210	300	30	90	150	210	300
$\hat{\beta}_u$	3.152	3.019	3.031	3.022	3.017	4.158	4.059	4.027	4.038	4.026
$\hat{R}_u(t_1)$.89	.91	.92	.92	.92	.96	.97	.97	.97	.97
$\hat{R}_u(t_2)$.82	.83	.84	.85	.85	.92	.94	.94	.95	.95
$\hat{R}_u(t_3)$.74	.76	.76	.76	.77	.87	.89	.90	.90	.90
$\hat{R}_u(t_4)$.65	.67	.67	.67	.68	.81	.83	.84	.84	.84

Table (9): Estimated shape parameter and the reliability function at $V_u = .5$ with Type-I HCS

	$r_j = 60\%n_j, \eta_1 = 5, \eta_2 = 3, \eta_3 = 2, V_1 = .9, V_2 = 1.4, V_3 = 1.9$									
	$\alpha = .5, C = 1.5, P = 1$					$\alpha = .5, C = 1.5, P = 1.5$				
	30	90	150	210	300	30	90	150	210	300
$\hat{\beta}_u$	3.152	3.019	3.031	3.022	3.017	4.367	4.324	4.288	4.238	4.242
$\hat{R}_u(t_1)$.89	.91	.92	.92	.92	.96	.98	.98	.98	.98
$\hat{R}_u(t_2)$.82	.83	.84	.85	.85	.93	.95	.95	.95	.96
$\hat{R}_u(t_3)$.74	.76	.76	.76	.77	.88	.91	.92	.92	.92
$\hat{R}_u(t_4)$.65	.67	.67	.67	.68	.83	.86	.87	.87	.87

ACKNOWLEDGMENTS:

The author is thankful to referee for his very helpful comments.

References

- Abdel-Ghaly, A. A., Attia, A. F. and Aly, H. M. (1998).**“Estimation of the parameters of Pareto distribution and the reliability function using accelerated life testing with censoring”, Communication in Statistics Simulation and Computation, 27(2), 469-484.
- Al-Hussini, E. K. and Abdel-Hamid,A.H.(2004).**“Bayesian estimation of the parameters , reliability and hazard rate functions of mixtures under accelerated life tests”, Communications in Statistics- Simulation and Computation, 33(4),963-982.
- Al-Hussini, E. K. and Abdel-Hamid,A.H.(2006).**“Accelerated life tests under finite mixture models”, Journal of Statistics- Computation and Simulation,76(8), 673-690.
- Aly, H. M. (1997).**“Accelerated life testing under mixture distribution”, Unpublished Ph.D. Thesis, Department of statistics, Faculty of Economics & Political Science, Cairo University, Egypt.
- Attia,A. F.,Shaban,A.S. and Abd El Sattar, M.H. (2013a).**"Estimation in Constant- Stress Accelerated Life Testing for Birnbaum-Saunders Distribution under Censoring", International Journal Of Contemporary Mathematical Sciences,8(4),173-18
- Attia,A. F.,Shaban,A.S. and Abd El Sattar, M.H. (2013b).**"Bayesian and non Bayesian estimationsfor Birnbaum-Saunders Distribution under Accelerated Life Testing Based On Censoring sampling", Applied Mathematical Sciences,7(66),3255-3269 .
- Artur, J. L. and Silvia, L.P. (2011).** "Testing hypothesis in the Birnbaum-Saunders distribution undter type-II censored samples" Computational Statistics & Data Analysis, 55, 2388-2399.
- Birnbaum, Z.W. and Saunders, S.C. (1969).** “A New Family of Life Distribution”, Journal of Applied Probability 6, 319-327.
- Chang, D. S.and Tang, L. C. (1993).**“Reliability bounds and critical time for the Birnbaum-Saunders distribution” IEEE Transactions on Reliability, 42, 464-469.
- Dupuis, D. J.and Milles, J. E. (1998).** “Robust estimation of the Birnbaum-Saunders distribution” IEEE Transactions on Reliability, 47, 88-95.
- Epstein, B. (1954).**"Truncated life test in exponential case", Annals of Mathematical Statistics,25,555-564.

- Hamouda, M.S. (2014).** "Complete and Partially Accelerated Life Testing using Birnbaum-Saunders Distribution", Unpublished Ph.D. thesis, ISSR ,Cairo university , Egypt.
- Hamouda, M.S. (2015).** "Bayesian and Non Bayesian Estimations for Birnbaum-Saunders Distribution under Partially Accelerated Life Testing Based on Censoring Sampling ", International Journal of Science and Research, 4(9), 1805-1813.
- Jeng,S. I. (2003).** "Inference for the fatigue life model based on the Birnbaum-Saunders distribution" Communications in Statistics – Simulation and Computation , 32, 43-60.
- Kevin, S. M. (1999).** "Estimation and predication for the Birnbaum-Saunders distribution using type – II censored samples , with a comparison to the inverse Gaussian distribution", Ph.D. Thesis, Department of statistics, College of arts and science , Kansas state university.
- Mann, N.R.,Schafer, R.E. and Singpurwalla, N.D. (1974).** "Methods for Statistical Analysis of Reliability and Life Data", New York, John Wiley and Sons.
- Miller, R.C., (1981).** "Survival Analysis ", John Wiley and sons, new York.
- Nelson, W. (1990).** "Accelerated Life Testing: Statistical Models, Data Analysis Test Plans", John Wiley and sons, new York.
- Ng, H. K., Kundu, D. and Balakrishnan, N. (2006).** "Point and interval estimation for the two-parameters Birnbaum-Saunders distribution based on Type-II censored samples". Computational Statistics & Data Analysis, 50, 3222-3242.
- Owen, W. J. (1997).** "Accelerated test models using the Birnbaum-Saunders distribution" Ph.D. Thesis, Department of statistics, Faculty of Science, University of South Carolina, U.S.A.
- Rieck, J. R. (1995).** "Parameter estimation for the Birnbaum-Saunders distribution based on symmetrically censored samples." Communication in Statistics, Theory and Methods, 24, 1721-1736.
- Sun, T. and Shi,Y. (2016).** "Estimation for Birnbaum-Saunders Distribution in simple step stress accelerated life test with type – II censoring." Communications in Statistics Simulation and Computation, 45, 880-901.

Wang, Z. ,Desmond,A.F.andLu,X. (2006). “Modified censored moment estimation for the two-parameter Birnbaum-Saunders distribution”, Computational Statistical & Data Analysis 50, 1033-1051.

Watkins,A.J. and John,A.M. (2008). “On constant stress accelerated life tests terminated by type – II censoring at one of the stress levels”, Statistical and Planning Inference , 138, 768-786.

Xu, A. and Tang.Y. (2010).“Reference analysis for Birnbaum-Saunders distribution” Journal of Computational Statistics & Data Analysis, 54, 185-192.

Xu, A. and Tang.Y. (2011).“Bayesian analysis of Birnbaum-Saunders distribution with partial information” Computational Statistics & Data Analysis, 55, 2324-2333.

Cairo University
Institute of Statistical Studies and Research

**The 52nd Annual Conference on Statistics, Computer
Sciences and Operations Research**

Applied Statistics

25-27 Dec. 2017

Index
APPLIED STATISTICS

1	On Quantile Regression for Dynamic Panel Data Models الحسينى عبد البر ، أحمد الشيخ ، حسام الدين محمود أحمد	1-15
2	A note on Markov Models and Chronic Diseases الحسينى عبد البر ، أيمن سعد الدين	16-33
3	Subset Selection in Multiple Linear Regression (Review) أحمد الشيخ ، أسماء وهبه	34-48
4	Estimation of Unbalanced Panel Data الحسينى عبد البر ، أحمد الشيخ ، آيه محمود أحمد	49-56
5	Extracting the Wealth of Knowledge in Composite Index Data أسماء محمد سيد	57-79
6	Review: Conditional and Unconditional AR(1) Model سيد مشعال ، أحمد أمين ، محمد خليفه ، هادية فريد	80-87
7	Estimation of Ranked Set Sampling Methods : Comparative Study نصر إبراهيم رشوان ، نوران مدحت الموان	88-103

On Quantile Regression for Dynamic Panel Data Models

Elhoussainy A. R.¹

Ahmed A.E.²

Hossameldin M. A.³

Abstract

This paper studies a quantile regression dynamic panel model with fixed effects. Panel data fixed effects estimators are typically biased in the presence of lagged dependent variables as regressors. To reduce the dynamic bias, we review the suggestion of the use of the instrumental variables quantile regression method of (Chernozhukov & Hansen, 2008) along with lagged regressors as instruments.

Keywords: Quantile regression, Dynamic panel, fixed effects, Instrumental variables

1. Overview on Quantile Regression (QR)

Quantile regression is a statistical technique able to detect more effects than conventional procedures: it does not restrict attention to the conditional mean and therefore it permits to approximate the whole conditional distribution of a response variable. It mainly aims at estimating conditional quantile of a certain variable of study given a set of covariates. Therefore, in the presence of patterned dataset, skewed distributed dataset or outliers, it would be possible to

figure out the effect of set of independent variables on a certain conditional quantile of a certain outcome with variant estimates (Chetverikov, et al., 2016)

A breakthrough in regression analysis is the quantile regression approach proposed by (Koenker & Bassett, 1978). This approach permits estimating various quantile functions of a conditional distribution, among them the median (0.5th quantile) function is a special case. Each quantile regression characterizes a particular (center or tail) point of the conditional distribution; putting different quantile regressions together thus provides a more complete description of the underlying conditional distribution. This analysis is particularly useful when the conditional distribution is heterogeneous and does not have a “standard” shape, such as an asymmetric, fat-tailed. (Powell, 1986) studied quantile regressions for censored data, and (Koenker , 2000) studies Galton, Edgeworth, Frisch, and prospects for quantile regression in econometrics

Quantile regression has several unique features than other regressions. It can be utilized for characterizing the whole conditional distribution of a dependent variable. The estimation is performed by Linear Programming (LP) which can be considered to be relatively easier than other conventional estimation methods. It provides a robust measure of location for non-symmetrically distributed dataset. Quantile regression is relatively able to grant a more efficient estimator than OLS method when error term normality assumption is violated. *L*-estimators (robust estimator based on linear combination of order statistic to be resistant to outliers) which depend on a linear combination of

quantile estimators are relatively more efficient than Least squares estimators (Portnoy & Lin, 2010)

Quantile regression aims at estimating conditional quantile of a certain variable of study given a set of regressors. Therefore, in the presence of patterned dataset, skewed distributed dataset or outliers, it would be possible to figure out the effect of set of independent variables on a certain conditional quantile of a certain outcome. That means for a certain varying outcome variable, several regression lines can be estimated to be the presentation of every quantile of a dependent variable (Chernozhukov & Fernández, 2011).

It is able to identify the behavioral effect of different regressors on a certain variable of study. That would help recognize if there is any change of their effects at several quantiles or stable effect over the whole distribution (Chetverikov, et al., 2016)

Conceding that, $\theta \in [0, 1]$ that would generate several estimated parameters according to variant values of θ . In other words, at the θ^{th} quantile, the coefficient ($\hat{\beta}_{\theta i}$) can be interpreted as the amount of change in the variable of study when a certain regressor changes by a one unit holding the rest of regressors fixed. That would allow researchers to investigate the effect of a set of covariates on the variable of interest at several quantiles. Studying the effects on a dependent variable at several levels of quantile, can indicate whether the effect is relatively the same cross the quantiles or it is variant which denotes the different influences on variable of interest at the low quantiles than at the high

ones. Detecting these effects cross quantiles is a considerably benefited from utilizing quantile regression.

An estimate for θ^{th} quantile of y can be obtained by minimizing the sample analog of population objective function as follows

$$\hat{\beta}_{\theta} = \min_{\beta} \frac{1}{n} \sum_{i=1}^N \rho_{\theta}(Y_i - X_i' \beta) \quad (1)$$

Where $i = 1, 2, \dots, n$, $\{y_1, \dots, y_n\}$ is a random sample of response variable Y , x_i is the covariate vectore corresponding to the i^{th} observation, ρ_{θ} is an indicator function $I(\cdot)$ defined to be θ for $i: y_i \geq x_i' \beta$ and $(1 - \theta)$ for $i: y_i < x_i' \beta$. $\hat{\beta}_{\theta}$ is the estimated parameters at a particular level of θ . The minimization analog in (1) can be presented in an alternative form to be as following:

$$\hat{\beta}_{\theta} = \min_{\beta} \left[\sum_{i: y_i \geq x_i' \beta}^N \theta |y_i - x_i' \beta| + \sum_{i: y_i < x_i' \beta}^N (1 - \theta) |y_i - x_i' \beta| \right] \quad (2)$$

Where $i = 1, 2, \dots, N$

In view of (2), the θ^{th} quantile regression estimator of β can be obtained by minimizing its sample counterpart, i.e., the average of asymmetrically weighted absolute errors with weight θ on positive errors ($u_i: y_i > x_i' \beta$ and weight $(1 - \theta)$ on negative errors $u_i: y_i < x_i' \beta$. Accordingly, LAD is obtained at $\theta = 0.5$ where the median regression locates. The argument in (2) can also be defined as

$$\hat{\beta}_{\theta} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n \rho_{\theta}(y_i - x'_i \beta) \quad (3)$$

where Median estimator (L_1) can be obtained in a certain special case where $\theta = 0.5$ which refers to the median estimator.

The function in (2) is minimized several linear programming methods, which can guarantee to yield a solution in a finite number of iterations. Quantile regression masters the advantage over OLS. OLS can be inefficient if the errors are highly non-normal, QR is more robust to non-normal errors and outliers. QR also provides a richer characterization of the data, permitting considering the impact of a covariate on the entire distribution of y , not merely its conditional mean. This non-differentiable function is minimized via the simplex method, which is guaranteed to yield a solution in a finite number of iterations **Invalid source specified.** and (Chetverikov, et al., 2016)

Since the fundamental goal of quantile regression is to estimate the conditional quantile of a response variable given a set of covariates at certain quantile of θ , therefore, the model in (1) can be presented in the form of θ^{th} conditional quantile of y on a $K \times 1$ vector of covariates $x = (1, x_1, \dots, x_K)$

$$Q_{\theta}(y|x) = x' \beta_{\theta} \quad (4)$$

That would permit the conditional quantile term in (3) to be re-written as a conventional econometric model as following:

$$y = x'\beta_{\theta} + u_{\theta} \quad (5)$$

This model is considered to follow the traditional OLS assumptions. and the population vector of parameters can be presented as following:

$$\hat{\beta}_{\theta} = \min_{\beta} E[\rho_{\theta}(u_{\theta}|x)]$$

Or alternatively,

$$\hat{\beta}_{\theta} = \min_{\beta} E[\rho_{\theta}(y - x'\beta|x)] \quad (6)$$

Within the frame of sample, the minimization argument at θ^{th} quantile for conditional quantile can be defined in the same sense, given $(y_i, x_i), i = 1, 2, \dots, n$ which is a i.i.d sample drawn from a certain population, the vector of parameters can be defined as following:

$$\hat{\beta}_{\theta} = \min_{\beta} \frac{1}{n} \sum_{i=1}^n \rho_{\theta}(u_{\theta}) \quad (7)$$

Or it can be defined within sample frame in an alternative way as following:

$$\hat{\beta}_{\theta} = \min_{\beta} \frac{1}{n} \left[\sum_{i:y_i \geq x'_i \beta} \theta |y_i - x'_i \beta| + \sum_{i:y_i < x'_i \beta} (1 - \theta) |y_i - x'_i \beta| \right]$$

2. Dynamic Panel Data Model with Fixed Effects (DPDFE)

Dynamic panel data (DPD) has increasingly become a main pillar in several real life applications, especially in Economics field. It provides the privilege of lagged dependent variable inclusion which permits more explanation to time invariant nature of several variables of studies. The conventional

dynamic panel data model would involve a cross section specific effect for a particular cross section, as a result of that, this time-invariant term would provide the differential for a specific cross section.

(Galvao, 2011) considered the following DPD-FE model of cross section units of $i = 1, 2, \dots, N$ and time periods of $t = 1, 2, \dots, T$.

$$Y_{it} = J(Y_{it-1}, X'_{it}, \mu_i, \varepsilon_{it})$$

$$Y_{it} = \delta Y_{it-1} + X'_{it} \beta + \mu_i + \varepsilon_{it} \quad (9)$$

Where Y_{it} is the variable of study or response variable, δ is the parameter of lagged dependent variable, and for stationarity purposes, $|\delta| < 1$ should be satisfied. X'_{it} is a P - dimensional row vector of continuous time-variant exogenous variables, β is a P - dimensional column vector of parameters for exogenous variables. μ_i represents cross sectional individual specific-effect and ε_{it} is innovation classical error term and the endogenous lagged dependent variable Y_{it-1} . The model in (9) is considered –strictly - to be DPD fixed effect model if mainly the correlation between individual effect μ_i and regressors is present. The model is presented with taking the following assumptions into consideration:

Assumption (1): across the panels, the error terms are distributed as $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$

Assumption (2): the error term observations are not correlated with the explanatory variables, so exogeneity is present and $E(x_{it}, \varepsilon_{it}) = 0$.

Assumption (3): the exogenous variables are correlated with the individual effect, so $E(x_{it}, \mu_i) \neq 0$.

Assumption (4): the error terms are correlated with the lagged endogenous variable $E(y_{it-1}, \varepsilon_{it}) = 0$

Assumption (5): across the panels the initial value of Y_{i1} is abtained by

$Y_{i1} = \frac{\mu_i}{1-\delta} + G_{i1}$, where $i = 1, 2, \dots, N$, $G_{i1} = \sum_{m=0}^{\infty} \delta^m \varepsilon_{it-1}$ and it is in no correlation with individual specific effect μ_i .

Least Square Dummy variable (LSDV), within and between groups estimators are considered to be the main estimation methods for DPD-FE models. To eliminate the effect of individual effect, the first difference is the transformation technique that is used in most of early literature. (Nickell, 1981) showed that wiping this term can be performed by taking difference approach follows:

$$\begin{aligned} Y_{it} - Y_{it-1} &= \delta(Y_{it-1} - Y_{it-2}) + (X_{it} - X_{it-1})\beta + (\mu_i - \mu_i) + (\varepsilon_{it} - \varepsilon_{it-1}) \\ \Delta Y_{it} &= \delta \Delta Y_{it-1} + \beta \Delta X_{it} + (\mu_i - \mu_i) + \Delta \varepsilon_{it} \end{aligned} \quad (10)$$

Since the individual effect term is time invariant, therefore it will be removed. As long as the time-invariant individual effect μ_i is a component in the functions of variable of study Y_{it} and Y_{it-1} , therefore, Y_{it-1} is correlated with the idiosyncratic error term U_{it} . As a result of that, LSDV would obtain inconsistent and certainly biased estimators even if error term observations are not in serial correlation. (Belloni & Chernozhukov, 2011) have also showed that, despite the

usage of LSDV (FE method), the existence of Y_{it-1} will maintain the correlation with the error term observations which would produce a bias of $O(1/T)$.

Handling this bias has been extensively studied by (Anderson & Hsiao, 1981). They have provided two main approaches to treat the arising problem of endogeneity from the differenced term of ΔY_{it-1} ; they depend on replacing it by a term which is more correlated with the variable of study and in no correlation with the error term. They proposed the essential approaches of considering lagged difference (ΔY_{it-2}) and lagged value (Y_{it-2}) of Y_{it-1} as these instruments are in no correlation with $\Delta \varepsilon_{it}$. In real life, it is quite hard to find an instrument for the lagged difference value of dependent variable. (Kiviet, 1995) showed that DPD-FE models that utilize instrumental variable estimation would suffer from weak finite sample efficiency and bias of $O(N^{-1} T^{-3/2})$.

3. Quantile Regression for DPD With Fixed Effects (QRDPDFE)

The estimation of quantile regression in dynamic panel data with fixed effects (QRPFE) has a growing literature over the last decade. The presence of individual effects is based on the assumption that there is a correlation between cross section-specific effect and regressors. Individual effects can also manage the heterogeneity across the panels. In addition to that, the inclusion of instrumental variable(s) due to the existence of lagged regressor has become a new approach. Recall the model in (9), it can be put in concise matrix form as follows:

$$Y = Z\mu + \delta Y_{-1} + X\beta + \varepsilon$$

Z is an incidence matrix which represents the parameters of separate individuals (N) in the sample of interest, and is obtained by $I_N \otimes l_T$ is a $T \times 1$ is a ones-vector. $\mu = (\mu_1, \dots, \mu_N)'$ is the $N \times 1$ individual specific effects vector which represents the intercepts.

The model in (9) can be presented in the frame of quantile regression. Considering i th individual and t th time period observations and at a particular level of quantile (θ) the model can be as follows:

$$Q_{y_{it}}(\theta | y_{it-1}, x_{it}) = \delta(\theta) Y_{it-1} + X'_{it} \beta(\theta) + \mu_i \quad (11)$$

Both y_{it-1} and x_{it} are the only regressors that rely on quantile level of (θ). And as long as this is a fixed effect model, therefore it permits the correlation between individual effect and covariates. Some of the variability in the response variable y_{it} is interpreted by the inclusion of μ 's. Involving them would

allow more controlling for unobserved heterogeneity that cannot be sufficiently interpreted by other regressors.

μ 's denote the time-invariant covariate and their parameters represent pure location shift effect on the variable of study y_{it} . Panel data would acquire its important nature from the excess number of cross sections over the length of given modestly smaller length of periods. In order to that, it would be considerably difficult, unrealistic and meaningless to estimate θ th distributional shift $\mu_{i\theta}$ for each individual, as line might not pass adequately due to the relatively small time series length. Hence, to avoid this problem, it has been suggested by (Koenker, 2000) to assume that the estimation of individual effect is independent from the level of quantile (θ). This assumption can be performed by estimating the model in (11) for several quantiles simultaneously. Based on (Koenker, 2004), the model in (11) can be presented as follows:

$$\min_{\mu, \delta, \beta} \sum_{s=1}^S \sum_{i=1}^N \sum_{t=1}^T p_s \rho_{\theta} (Y_{it} - X'_{it} \beta(\theta) - \delta(\theta) Y_{it-1} - \mu_i) \quad (12)$$

Where $\rho_{\theta}(u) := u(\theta - I(u < 0))$ as defined by Koenker and Basset (1978), the weights p_s control the potential relative effect of a given S quantiles $(\theta_1, \theta_2, \dots, \theta_S)$, the selection of the weights and the related quantiles are based on Weighted L-statistic according to (Mosteller & Tukey, 2002).

4. Quantile Regression for Dynamic Panel Data with Fixed Effects - Instrumental Variables Approach (QRDPDFEIV)

In dynamic panel data, the presence of lagged value of dependent variable will lead to biased estimators would generate biased estimators due to the correlation between error term and lagged dependent variable. The biased estimators would result from OLS and QR techniques.

Within the conceptual frame of linear regression estimation, Anderson and (Hsiao, 1986) and (Arellano & Bond, 1991), provided several IV based- frameworks that can help provide consistent estimators in DPD models. Based on that objective, they provided the perspective of utilizing lagged or lagged differences of the endogenous variables of interest. The exploitation of these instrument(s) extract its importance from being correlated with their endogenous variables but not with the innovation term. As a result of that, the bias can be minimized and obtain more consistent estimators.

The involvement of a particular valid instrument (m_{it}) in the model as a part of regressors is the traditional method of estimating a specific parameter(s). By considering the minimization argument in (12), this frame can be presented as follows:

$$\min_{\mu, \delta, \beta, \lambda} \sum_{s=1}^S \sum_{i=1}^N \sum_{t=1}^T p_s \rho_{\theta} (Y_{it} - X'_{it} \beta(\theta) - \delta(\theta)Y_{it-1} - \lambda(\theta)m_{it} - \mu_i) \quad (13)$$

Where, Y_{it-1} is a dim (δ)-vector of endogenous variable, μ_i is the FE term, X'_{it} is a dim (β)-vector of exogenous covariates, m_{it} is a dim (λ)-vector of instrumental variables. And it should be taken into consideration that $\dim(\lambda) \geq \dim(\delta)$. The involvement of Y_{it-1} would generate bias, however, it can be trimmed by involving IV (m_{it}), that can influence the variable of interest Y_{it} , nevertheless it is independent from ε_{it} . In Case of acquiring valid instruments, therefore they will assist in estimating δ and β in a better way.

The lagged or lagged differences values for two time periods or more of (Y_{it}) and (X_{it}) can be probably used to as instruments (m_{it}) to estimate δ and/or β .

In real life, one of two main basic techniques is followed. The first technique is based on running (m_{it}) – which is/are selected based on the theory- along with the other covariates to gain the corresponding effects on Y_{it} .

The second technique depends on estimating (m_{it}) by using the predicted values from OLS projection of lagged Y on m and X , which are considered to be independent from the error term as follows:

$$\tilde{y}_{i-} = g(m_{it}, x_{it}, v_{it}) \quad (14)$$

And then to insert it in the following model:

$$Y_{it} = \delta \tilde{Y}_{i-} + X'_{it} \beta + \mu_i + \varepsilon_{it} \quad (15)$$

The two approaches are useful to minimize the bias effect. And they can be stemmed from the economic theory (inside the model) or can be estimated by one of the mentioned techniques.

5. Conclusion

We presented the evolvement of applying the quantile regression in dynamic panel data. In addition to that we reviewed the recent developments in handling the bias problem in DPDFE models by utilizing IV approach. This approach is intended to minimize the presence of bias in DPDFEIV models.

6. References

- Alizadeh , M., Bagheri, S., & Nadarajah, S. (2015). Efficient Estimation of the PDF and the CDF of the Exponentiated Gumbel Distribution. *Communication in Statistics - Theory and Methods*, 45(1), 339-361.
- Anderson, T., & Hsiao, C. (1981). Estimation of dynamic models with error components. *JASA*, 589-606.
- Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to. *The Review of Economic Studies*, 58(2), 277-297.
- Bassett, G., Koenker, R., & Kordas, G. (2004). Pessimistic Portfolio Allocation and Choquet Expected Utility. *Journal of Financial Econometrics*, 2, 477–492.
- Belloni, A., & Chernozhukov, V. (2011). ℓ_1 -penalized quantile regression in high-dimensional sparse models. *The Annals of Statistics*, 39(1), 82-130.

- Chernozhukov, V., & Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, 142(1), 379-398.
- Chetverikov, D., Larsen, B., & Palmer, C. (2016). IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade. *Journal of Econometric Society*, 84(2), 809-833.
- Galvao, A. (2011). Quantile regression for Dynamic Panel Data with Fixed Effects. *Journal of Econometrics*, 164, 142-157.
- Hsiao, C. (1986). *Analysis of Panel Data*. New York: Cambridge University Press.
- Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics*, 68(1), 53-78.
- Koenker, R. (2000). Galton, Edgeworth, Frisch, and prospects for quantile regression in Econometrics. *Journal of Econometrics*, 95, 347-374.
- Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1), 74-89.
- Koenker, R., & Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1), 33-50.
- Mosteller, F., & Tukey, J. (2002). Data Analysis and Regression. *The Annals of Statistics*, 30(6), 1535–1575.
- Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49(6), 1417-1426.
- Portnoy, S., & Lin, G. (2010). Asymptotics for Censored Regression Quantiles. *Journal of Nonparametric Statistics*, 22(1), 115-130.
- Powell, J. (1986). Censored regression quantiles. *Journal of Econometrics*, 32, 143-155.

A note on Markov Models and Chronic Diseases

El-Houssainy A. Rady*

Ayman S. Anwar**

Abstract

The paper is concerned with introducing the concept of Markov model concentrating on its relationship with chronic diseases. An introduction of Markov models is given in section one, the Hidden Markov model is explained in section 2. Section 3 is concerned with kidney disease and its stages, Support Vector Machine classification technique described in section 4.

Keywords: Markov Model, Hidden Markov Model, Chronic Kidney Disease, Data Mining techniques, Support vector machine.

Introduction

Chronic diseases are often described by stages of severity. Clinical decisions about what to do are influenced by the stage, whether a patient is progressing, and the rate of progression.

The stage of disease and rate of progression or regression are important to deciding whether to treat, how to treat, and how often to monitor a patient. Moreover, knowledge about transition rates between stages helps patients understand what to expect and policymakers what to plan.

One approach for analyzing disease stage data is hidden Markov models (HMMs) (MacDonald and Zucchini 1997, 2009). Unlike ordinary Markov models, HMMs account for the fact that sometimes the observed disease stages are different from the underlying disease stages as a result of measurement error.

Institute of Statistical Studies & Research

* Professor of Applied Statistic, Institute of Statistical Studies & Research.

** Corporate MIS Manager, Magrabi Hospitals & Centers.

Recently, researchers have used continuous-time HMMs to analyze data in a variety of clinical areas, such as hepatocellular cancer (Kay 1986), HIV progression (Satten and Longini 1996), and aortic aneurysms (Jackson 2003). However, a continuous-time model is computationally costly, and may be infeasible if the sample size is large, which is typically the case with electronic health records (EHR) data. Further, for many studies there would be no benefit to having finer information about the timing of a measurement than the calendar date. Discrete-time HMMs are a useful alternative, and have been developed and applied to a variety of health problems (Shirley et al. 2010; Rabiner 1986; Jackson and Sharples 2002; Scott 1999; Scott 2002; Scott et al. 2005; Gentleman et al. 1994; Bureau et al. 2000).

While discrete-time HMMs have many desirable features, the estimation of transition rates typically requires large observational studies with long follow-up times as transitioning usually occurs over years. The resources required for such studies are often costly and time prohibitive. Use of longitudinal EHRs data from large primary care practices offers an alternative means of assembling longitudinal health experience of a population. Such data have the advantage of having both sufficient follow-up time and sample size to reliably and accurately estimate these rare transition rates.

1. Markov Models

Markov model is a stochastic model used to model randomly changing systems where it is assumed that future states depend only on the present state and not on the sequence of events that preceded it (that is, it assumes the Markov property). Generally, this assumption enables reasoning and computation with the model that would otherwise be intractable.

Markov chain

It is a stochastic model described as follows: we have a chain of a set of states, $S = \{s_1, s_2, \dots, s_n\}$. the process starts in one of these states and moves successively from one state to another. Each move is called a step. If the chain is currently in state s_m then its moves to state s_n at the next step with a probability p_{mn} and this probability doesn't depend on which state was the chain in before the current state.

The Markov property suggests that the distribution for this variable depends only on the distribution of the previous state.

Formally, A Markov chain is a sequence of random variables W_1, W_2, W_3, \dots with the Markov property, namely that, given the present state, the future and past states are independent.

The components of Markov Chain

- State space denoted by S which is a set of all possible state of the chain
- Transition probabilities denoted by P which is a set of all possible transition probabilities to move from each state to another state or stay on the same state.

Example-1:

Suppose that a diabetic patient has three stages of disease severity: Normal Stage, Medium Severity Stage and Critical Severity Stage. Let's assume for the moment that the severity of disease doesn't change during the day.

Disease severity prediction is trying to guess what the stage of disease will be next day based on a history of observations of disease severity stages., we can express that by the following probabilities:

$$P(s_n \setminus s_{n-1}, s_{n-2}, s_{n-3}, \dots, s_1) \quad (1)$$

Using formula (1), we can give probabilities of diabetic severity stages for next day using n days of history., but the problem here that for large n more statistics must be collected, if we suppose that $n=10$, then we have to collect statistics for $3^{10}=59,049$ past histories, so we will simplify it by using Markov Assumption:

$$P(s_n \setminus s_{n-1}, s_{n-2}, s_{n-3}, \dots, s_1) \approx P(s_n \setminus s_{n-1}) \quad (2)$$

Also, we can express the joint probability using the Markov Assumption as follow:

$$P(s_1, s_2, s_3, \dots, s_n) = \prod_{i=1}^n P(s_i \setminus s_{i-1}) \quad (3)$$

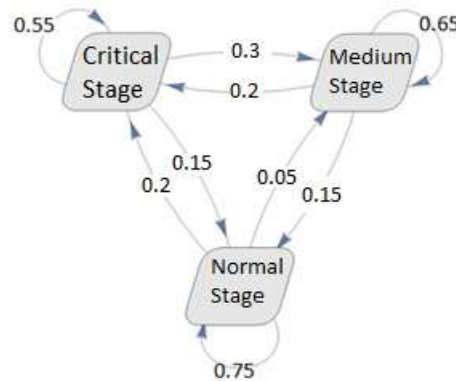
Let's arbitrarily pick some numbers for probabilities of next day diabetic severity stage based on today severity stage (Transition Probabilities P), expressed in table (1)

Table 1. Transition Probabilities for diabetic severity stages

		Next day Diabetic Severity Stage		
		Normal Stage	Medium Severity Stage	Critical Severity Stage
Today's Diabetic	Normal Stage	0.75	0.05	0.2
	Medium Severity Stage	0.15	0.65	0.2
	Critical Severity Stage	0.15	0.3	0.55

$$P = \begin{pmatrix} 0.75 & 0.05 & 0.2 \\ 0.15 & 0.65 & 0.2 \\ 0.15 & 0.3 & 0.55 \end{pmatrix}$$

Figure -1. Diabetic severity stage transition diagram



So, given that today diabetic patient in a normal stage it's easy compute the Probability that tomorrow will be also in normal stage and the day after tomorrow will be in a critical stage as the following: -

$$\begin{aligned}
 P(s_3=critical, s_2=normal \mid s_1= normal) &= P(s_3=critical \mid w_2=normal, w_1= normal)^* \\
 &P(s_2=normal \mid s_1= normal) \\
 &= P(s_3=critical \mid s_2=normal)^* \\
 &P(s_2=normal \mid s_1= normal) \\
 &= 0.2 * 0.75 \\
 &= 0.15
 \end{aligned}$$

Also, the probability that the diabetic patient on the day after tomorrow will be in a medium stage given that today severity stage is critical can be computed as follow: -

$$\begin{aligned}
 P (s_3=medium \mid s_1= critical) &= P (s_3=medium, s_2=critical \mid s_1= critical) + \\
 &P (s_3= medium, s_2= medium \mid s_1= critical) + \\
 &P (s_3= medium, s_2=normal \mid s_1= critical) +
 \end{aligned}$$

$$\begin{aligned}
 &= P(s_3 = \text{medium} \mid s_2 = \text{critical}) P(s_2 = \text{critical} \mid s_1 = \\
 &\text{critical}) + \\
 &P(s_3 = \text{medium} \mid s_2 = \text{medium}) P(s_2 = \text{medium} \mid \\
 &s_1 = \text{critical}) + \\
 &P(s_3 = \text{medium} \mid s_2 = \text{normal}) P(s_2 = \text{normal} \mid s_1 = \\
 &\text{critical}) \\
 &= 0.3 * 0.55 + 0.65 * 0.3 + 0.05 * 0.15 \\
 &= 0.165 + 0.195 + 0.0075 \\
 &= 0.3675
 \end{aligned}$$

It can be also computed directly by calculate $P^{(2)}$

$$\begin{aligned}
 P^{(2)} &= \begin{pmatrix} 0.75 & 0.05 & 0.2 \\ 0.15 & 0.65 & 0.2 \\ 0.15 & 0.3 & 0.55 \end{pmatrix} * \begin{pmatrix} 0.75 & 0.05 & 0.2 \\ 0.15 & 0.65 & 0.2 \\ 0.15 & 0.3 & 0.55 \end{pmatrix} \\
 &= \begin{pmatrix} 0.6 & 0.13 & 0.27 \\ 0.24 & 0.49 & 0.27 \\ 0.24 & 0.3675 & 0.3925 \end{pmatrix}
 \end{aligned}$$

$$P_{\text{critical, medium}}^{(2)} = 0.3675$$

Absorbing Markov Chains

A State s_i of a Markov chain is called absorbing if it is impossible to leave it (i.e., $p_{ii} = 1$)., A Markov Chain is *absorbing* if it has at least one absorbing state, and if from every state it's possible to go to an absorbing state within one or more step., also other non-absorbing states called *transient*.

Example – 2

Once the patient has clinical diabetic kidney disease, he can either remain in that health state or progress to higher stage kidney disease. Once the patient has contracted preliminary diabetic kidney disease, the model assumes he cannot be cured of it. Therefore, the patient either progresses to the clinical stage or remains in the preliminary stage. The following is a sample data of diabetic kidney disease transition probability matrix

Table 2. Transition Probabilities matrix for CKD stages

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Stage 1	0.93	0.06	0	0	0.01
Stage 2	0	0.85	0.12	0	0.03
Stage 3	0	0	0.75	0.15	0.10
Stage 4	0	0	0	0.65	0.35
Stage 5	0	0	0	0	1

$$P = \begin{pmatrix} 0.93 & 0.06 & 0 & 0 & 0.01 \\ 0 & 0.85 & 0.12 & 0 & 0.03 \\ 0 & 0 & 0.75 & 0.15 & 0.10 \\ 0 & 0 & 0 & 0.65 & 0.35 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

on that Case Stage 5 is *absorbing* State., Stage 2,...,Stage 4 are *transient*.

On such case (absorbing chains) it's interesting to determine the following: -

- What is the probability that the process will reach an absorbing state?
- What is the probability that the process will end up in a certain absorbing state?
- How long the process will take in average to be absorbed?

- How many times on average the process will be in each transient state?

Canonical Form

By reordering the transition probability matrix so that transient states come first and absorbing states on the next will get the following canonical form

$$P = \begin{matrix} \text{TRN} \\ \text{ABS} \end{matrix} \left(\begin{array}{c|c} \mathbf{Q} & \mathbf{R} \\ \hline \mathbf{0} & \mathbf{I} \end{array} \right)$$

Where: -

Q: is an $t \times t$ matrix, t is no of transient states.

R: is a none zero $t \times r$ matrix, r is no of absorbing states.

0: is an $r \times t$ zero matrix.

I: is an $r \times r$ identity matrix.

Fundamental Matrix

Definition (1) For an absorbing Markov chain P , the matrix $F = (I - Q)^{-1}$ is called the fundamental matrix for P and the entry n_{ij} of F gives the expected number of times that the process is in the transient state s_j if it is started in the transient state s_i .

$$P \text{ (Canonical form)} = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \left(\begin{array}{cccc|c} 0.93 & 0.06 & 0 & 0 & 0.01 \\ 0 & 0.85 & 0.12 & 0 & 0.03 \\ 0 & 0 & .75 & 0.15 & 0.10 \\ 0 & 0 & 0 & 0.65 & 0.35 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

$$\begin{aligned}
 F &= \left[\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.93 & 0.06 & 0 & 0 \\ 0 & 0.85 & 0.12 & 0 \\ 0 & 0 & 0.75 & 0.15 \\ 0 & 0 & 0 & 0.65 \end{pmatrix} \right]^{-1} \\
 &= \begin{matrix} 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{pmatrix} 14.3 & 5.7 & 2.7 & 1.2 \\ 0 & 6.7 & 3.2 & 1.3 \\ 0 & 0 & 4 & 1.7 \\ 0 & 0 & 0 & 2.9 \end{pmatrix}
 \end{aligned}$$

Time to Absorption

The Sum for all entries in the i^{th} row of F gives total expected number of times in any of transient states for a given starting state s_i which mean that, the expected time required before the chain being absorbed., Generally it can be computed as follow:

$$t = Fc$$

where: -

t : column vector whose entries represent the expected time before the chain is absorbed given all possible starting states.

F : Represent fundamental matrix.

c : is a column vector all of whose rows are 1

$$t = \begin{matrix} 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{pmatrix} 14.3 & 5.7 & 2.7 & 1.2 \\ 0 & 6.7 & 3.2 & 1.3 \\ 0 & 0 & 4 & 1.7 \\ 0 & 0 & 0 & 2.9 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 23.9 \\ 11.2 \\ 5.7 \\ 2.8 \end{pmatrix}$$

2. Hidden Markov Model

A hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. The mathematics behind the HMM was developed by L. E. Baum and coworkers 1966. It is closely related to an earlier work on the optimal nonlinear filtering problem by Ruslan L. Stratonovich 1960, who was the first to describe the forward-backward procedure.

In simple Markov models (like a Markov chain), the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore, the sequence of tokens generated by an HMM gives some information about the sequence of states. Note that the adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; the model is still referred to as a 'hidden' Markov model even if these parameters are known exactly.

Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics.

A hidden Markov model can be considered a generalization of a mixture model where the hidden variables, which control the mixture component to be selected for each observation, are related through a Markov process rather than independent of each other.

In its discrete form, a hidden Markov process can be visualized as a generalization of the Urn problem with replacement (where each item from the urn is returned to the original urn before the next step).

Example-3

Recall Example-1 and suppose that physician who follow-up diabetic patients is just come back from vacation, and he were asked about the stage of diabetic severity for examined patient. The only piece of evidence he has is the examined patient face status is it pale or not.

Let’s assume the following probabilities: -

Table 3. Emission probabilities for seeing a pale face

	Probability of pale face
Normal Stage	0.1
Medium Stage	0.3
Critical Stage	0.8

Now we have to factor in the fact that the actual severity stage for the examined patient is *hidden* from him. We do that by using Bayes’ Rule:

$$P(s_1, s_2, \dots, s_n \mid u_1, u_2, \dots, u_n) = \frac{P(u_1, u_2, \dots, u_n \mid s_1, s_2, \dots, s_n)P(s_1, s_2, \dots, s_n)}{P(u_1, u_2, \dots, u_n)}$$

Where u_i is true if the examined patient has a pale face on a day i , and false if the examined patient hasn’t. the probability $P(s_1, s_2, \dots, s_n)$ is same as the Markov model from the last section, and the probability $P(u_1, u_2, \dots, u_n)$ is the prior probability of seeing a particular sequence of pale face events e.g. {true, false, true}. The probability $P(u_1, u_2, \dots, u_n \mid s_1, s_2, \dots, s_n)$ can be estimated as $\prod_{i=1}^n P(u_i \mid s_i)$, if all i , given s_i, u_i is independent of all u_i and s_j , for all $j \neq i$.

Suppose that the patient severity stage was in Normal stage; the examined patient had a pale face on day 2, but not on day 3. Assuming that the prior probability that the

examined patient has pale face is 0.5 then the probability that the disease severity stage is critical in day 3 can be computed as follow: -

$$\begin{aligned}
 P(s_3 = C \setminus s_1 = N, u_2 = T, u_3 = F) &= P(s_2 = \text{critical}, s_3 = \text{critical} \setminus s_1 = \\
 &\text{normal}, u_2 = \text{True}, u_3 = \text{False}) \\
 &+ P(s_2 = \text{medium}, s_3 = \text{critical} \setminus s_1 = \text{normal}, u_2 = \\
 &\text{True}, u_3 = \text{False}) \\
 &+ P(s_2 = \text{normal}, s_3 = \text{critical} \setminus s_1 = \text{normal}, u_2 = \\
 &\text{True}, u_3 = \text{False}) \\
 &= \frac{P(u_3=F \setminus s_3=C)P(u_2=T \setminus s_2=C)P(s_3=C \setminus s_2=C)P(s_2=C \setminus s_1=N)P(s_1=N)}{P(u_3=F)P(u_2=T)P(s_1=N)} + \\
 &\frac{P(u_3=F \setminus s_3=C)P(u_2=T \setminus s_2=M)P(s_3=C \setminus s_2=M)P(s_2=C \setminus s_1=N)P(s_1=N)}{P(u_3=F)P(u_2=T)P(s_1=N)} + \\
 &\frac{P(u_3=F \setminus s_3=C)P(u_2=T \setminus s_2=N)P(s_3=C \setminus s_2=N)P(s_2=N \setminus s_1=N)P(s_1=N)}{P(u_3=F)P(u_2=T)P(s_1=N)} + \\
 &= \frac{(0.7)(0.3)(0.55)(0.2)}{(0.5)(0.5)} + \frac{(0.7)(0.8)(0.2)(0.05)}{(0.5)(0.5)} + \frac{(0.7)(0.1)(0.2)(0.75)}{(0.5)(0.5)} \\
 &= 0.1568
 \end{aligned}$$

3. Chronic Kidney Disease (CKD)

Chronic kidney disease (CKD) is one of a growing health problem worldwide. It is a progressive disease associated with a high risk of cardiovascular diseases, mortality and morbidity rates, and high health care costs. Therefore, early detection of the disease to control and manage the consequences is highly significance., Because of the dynamic nature of CKD, its covert nature in the early stages, and heterogeneity of patients, predicting the CKD progression with reasonable accuracy is necessary.

CKD progression can be considered as a function of various parameters including underlying renal diseases, blood pressure, hypertension, proteinuria, and age., Early diagnosis of the CKD are takes great attention among physicians, especially on

determining the appropriate time to apply medical treatments and to control identified risk factors that reflect on the disease progression to End Stage Renal Disease (ESRD) like hypertension, proteinuria, and hyperphosphatemia.

Stages of Chronic Kidney Disease

The stages of CKD (Chronic Kidney Disease) are mainly based on measured or estimated GFR (Glomerular Filtration Rate), there are five stages, but kidney function is normal in Stage 1, and minimally reduced in Stage 2.

The KDOQI (Kidney Disease Outcomes Quality Initiative) stages of kidney disease are:

Table 4. CKD Stages according to GFR measurement value

Stage	GFR	Description	Treatment stage
1	90+	Normal kidney function but urine findings or structural abnormalities or genetic trait point to kidney disease	Observation, control of blood pressure.
2	60-89	Mildly reduced kidney function, and other findings (as for stage 1) point to kidney disease	Observation, control of blood pressure and risk factors.
3A 3B	45-59 30-44	Moderately reduced kidney function	Observation, control of blood pressure and risk factors.
4	15-29	Severely reduced kidney function	Planning for end stage renal failure.
5	<15 or on dialysis	Very severe, or end stage kidney failure (sometimes call established renal failure)	Treatment choices.

Definition of chronic: Labelling someone as having CKD requires two samples at least 90 days apart. Historical values can be used.

eGFR is estimated Glomerular Filtration Rate, usually based on serum Creatinine level, age, sex, and race. The most widely used method for this is the abbreviated MDRD (Modification of Diet in Renal Disease) equation, as it has proved the most robust and accurate. Normal GFR is approximately 100mls/min/1.73m².

$$\text{eGFR} = 186 \times (\text{Creatinine} / 88.4)^{-1.154} \times (\text{Age})^{-0.203} \times (0.742 \text{ if female}) \times (1.210 \text{ if black})$$

4. Support Vector Machine (SVM)

Data mining is the process of extracting hidden information from the large dataset. Data mining techniques are used in various applications like fault diagnosis, anomaly detection, medical diagnosis, e-mail filtering, face recognition and oil spill detection. Data mining techniques such as classification, clustering and association rule etc. plays a great role in extracting unknown knowledge from the databases. Classification is a data mining technique used to predict group membership for data instances., the algorithm processes a training set containing a set of attributes and the respective outcome, usually called goal or prediction attribute., one of data mining Techniques is Support Vector Machine.,

SVM is a method for the classification of both linear and nonlinear data., SVM is an algorithm that works as follows. It uses a nonlinear mapping to renovate the unique training data into a higher dimension. Surrounded by this new dimension, it examines for the linear optimal separating hyperplane i.e., a “decision boundary” sorting out the tuples of one class from another. With a suitable nonlinear mapping to a necessarily high dimension, data from two classes can always be separated by a hyperplane. The SVM finds the hyperplane using support vectors and margins. Although the training time of even the fastest SVMs can be exceedingly slow, they are extremely accurate, outstanding

to their ability to model complex nonlinear decision boundaries. They are much less prone to over fitting than other methods. The Support vectors initiate also provide a compact description of the learned model. SVMs can be used for prediction along with classification. They have been applied to a several areas, including handwritten digit recognition, object recognition, and speaker identification and benchmark time-series prediction tests.

5. Conclusion

More work and research still in progress to formalize a new approach to estimate and predict chronic kidney disease stage transition rates using the Hidden Markov Model and Support Vector Machine Data Mining technique which will support physicians & policymaker in their medical & economical decisions

References

1. MacDonald, Iain L. and Zucchini, Walter (1997). Hidden Markov and Other Models for Discrete-valued Time Series. Chapman Hall.
2. MacDonald, Iain L. and Zucchini, Walter (2009). Hidden Markov Models for Time Series: An Introduction Using R. Chapman Hall.
3. Kay, Richard (1986). A Markov Model for Analysing Cancer Markers and Disease States in Survival Studies. Biometrics, Vol.42, No.4, 855-865.
4. Satten, Glen A., Longini Ira M., Jr. (1996). Markov Chains with Measurement Error: Estimating the “True” Course of a Marker of the Progression of Human Immunodeficiency Virus Disease. Applied Statistics, Vol. 45, No.3, 275-309.

5. Jackson, C., Sharples, L., Thompson, S., Duffy, S., and Couto, E. (2003). Multistate Markov Models for Disease Progression with Classification Error. *Journal of the Royal Statistical Society. Series D*, Vol. 52, No.2,193-209.
6. Shirley, K., Small, D., Lynch, K., Maisto, S., and Oslin, D. (2010). Hidden Markov Models for Alcoholism Treatment Trial Data. *Annals of Applied Statistics*, Vol. 4, No. 1, 366-395.
7. Rabiner, L. R. (1986). An Introduction to Hidden Markov Models. *IEEE Acoustics, Speech, and Signal Processing* January, 4-16.
8. Jackson, C. and Sharples, L. (2002). Hidden Markov models for the onset and progression of bronchiolitis obliterans syndrome in lung transplant. *Statist. Med.*, 21,113-128.
9. Scott, Steven L. (1999). Bayesian Analysis of a Two-State Markov Modulated Poisson Process. *Journal of Computational and Graphical Statistics*, Vol.8, No. 3, 662-670.
10. Scott, Steven L. (2002). Bayesian Methods for Hidden Markov Models: Recursive Computing in the 21st Century. *Journal of the American Statistical Association*, Vol. 97, No. 457, 337-351.
11. Gentleman R.C, Lawless, J.F., Lindsey, J.C., and Yan, P. (1994). Multi-State Markov Models for Analysing Incomplete Disease History Data with Illustrations for HIV Disease. *Statistics in Medicine*, Vol. 13, 805-821.
12. Bureau, A., Hughes, J., Shiboski, S. (2000). An S-Plus Implementation of Hidden Markov Models in Continuous Time. *Journal of Computational and Graphical Statistics*, Vol. 9, No.4, 621-632.

13. Levey, AS., Bosch, JP., Lewis, JB., Greene, T., Rogers, N., Roth, D. (1999). A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. *Annals of Internal Medicine*,130, 461-470
14. Scott, S., James, G. and Sugar, C. (2005). Hidden Markov Models for Longitudinal Comparisons. *Journal of the American Statistical Association*, Vol 100, Issue 470.
15. Jackson, C. (2007). Multi-state modelling with R: the msm package. Medical Research Council Biostatistics Unit: Cambridge, UK.
16. Altman, R.M. (2007), Mixed Hidden Markov Models: An Extension of the Hidden Markov Model to the Longitudinal Data Setting, *Journal of the American Statistical Association*, 102, 201-210.
17. S. Vijayarani, S.Dhayanand ,2015 "KIDNEY DISEASE PREDICTION USING SVM AND ANN ALGORITHMS" ,*International Journal of Computing and Business Research (IJCBR)*, Vol 6, Issue 2
18. Ishani Mandli, Mahesh Panchal, (2014) "Selection of Most Relevant Features from High Dimensional Data using IG-GA Hybrid Approach", *International Journal of Computer Science and Mobile Computing*, Vol.3 Issue 2, 827-830
19. Sunitha Devi1. P, Sowjanya. CH, Sunitha. K.V.N, (2014) "A Review of Supervised Learning Based Classification for Text to Speech System", *International Journal of Application or Innovation in Engineering & Management*, Vol 3, Issue 6, 79-86
20. Sunitha Devi1. P, Sowjanya. CH, Sunitha. K.V.N, (2014) "A Review of Supervised Learning Based Classification for Text to Speech System", *International Journal of Application or Innovation in Engineering & Management*, Vol 3, Issue 6, 79-86
21. Sonali Sonavane, Alka Khade, Gaikwad. V.B, (2013) "Novel Approach for Localization of Indian Car Number Plate Recognition System using Support Vector

Machine”, International Journal of Advanced Research in Computer Science and Software Engineering, Vol 3, Issue 8, 179-183

22. Shaohua Teng, Hongle Du, Naiqi Wu, Wei Zhang, Jiangyi Su, (2010) A Cooperative Network Intrusion Detection Based on Fuzzy SVMs, JOURNAL OF NETWORKS, VOL. 5, NO. 4, APRIL, 475-483

23. Chronic Kidney Disease, Clinical Practice Recommendations for Primary Care Physicians and Healthcare Providers 2011, HENRY FORD HEALTH SYSTEM.

Subset Selection in Multiple Linear Regression (Review)

Ahmed Amin El-Sheikh*, Asmaa Mohamad Wahba**

Abstract:

This paper provides a brief review of subset selection criteria, in particular stepwise criterion, expected mean squared prediction error criterion, distance between predicted values based on the full model and predicted values based on the sub model criterion, Kullback's discrepancy criterion, The Loss Rank Principle (LoRP) criterion, and other Criteria.

Keywords: EMSE, Multiple Linear Regression, Stepwise, Subset Selection.

1. Introduction:

Subset selection is one of the most fundamental problems in statistical applications, it appears with the need to model the relationship between variable of interest and a set of explanatory variables, and then there is an urgent need to answer the following question; which variables will be selected to enter the model among competing explanatory variables.

Suppose Y is a variable of interest, and X_1, \dots, X_k is a set of prospective explanatory variables, the problem of variable selection, or subset selection as it is often called, arises when one wants to model the relationship between Y and a subset of X_1, \dots, X_k , but there is uncertainty about which subset to use. Such a situation is particularly of interest when k is large and X_1, \dots, X_k is thought to contain many redundant or irrelevant variables.

The problem of subset selection lies in how to choose the variables which construct the "best" model among the alternative models. Over the decades, scientists have investigated an assortment of statistical approaches to select among alternative models (the number of possible subsets of one or more variables out of k is (2^k-1)). **Miller (1990)**

*Professor of Statistic, Department of Applied Statistic and Econometrics, Institute of Statistical Studies and Research, Cairo University, Egypt.

E-Mail: Aham103@yahoo.com

**Assistant Lecturer, Department of Mathematics, Insurance, and Statistic, Faculty of Commerce, University of Sadat city, Egypt.

E-Mail: asmaa_wahba@com.usc.edu.eg

2. Stepwise criterion:

Some of the most popular methods are forward selection, stepwise regression and backward elimination. Forward selection method starts with a model that has just a constant in it. Next, choose the variable that satisfies the criteria (e.g. R^2 , p value, actual F change, etc) that has been set for entry. The following step are to choose from remaining variable the variable that meets the entry criteria, keeping doing this until there are no variables that meet the entry criteria. On the contrary, backward elimination starts with all variables, including a constant if there is one. The variable which meets the elimination criteria is chosen for deletion. In each step, all the variable in the models have been evaluated, then remove the one that meet the elimination criteria. This process will be stopped when there are no more variables that meet the criteria for removal. **Norusis (2006)**

Stepwise regression is used to mean Efroymson's algorithm which proposed by Efroymson in 1960. This method differs from forward selection in that after each added variable to the set of selected variables, a test is made to see if any of the previously selected variables can be deleted. **Miller (1990)**.

In spite of the fact that the forward selection, stepwise regression and backward elimination methods are remaining popular in many computer software packages, there are many critiques to these methods. Of which, there is no guarantee that the final model is the optimal one. In addition, the procedure reaches to a single final model, although in some cases there are several good models.

3. Expected mean squared prediction error criterion:

Some popular subset selection methods are based on estimating the predictive ability of each model. The predictive ability of the model could be measured by the expected mean squared prediction error i.e.,

$$EMSE(M_j) = E(y_{new} - \hat{y}_{new})^2$$

Where y_{new} is randomly selected new data from population and \hat{y}_{new} is the prediction for y_{new} generated by model M_j .

By Estimating $EMSE$ for all possible models and choosing the model that has the smallest $EMSE$. Although it is difficult to Estimate $EMSE$ because the data about the all possible samples is not available, many studies tried to present an unbiased estimate of $EMSE$. **John, et al. (2005)**

Mallows (1973) investigated an unbiased estimate of $EMSE$ and recommended the use of C_p plots to help gauge subset selection for p explanatory variables, this statistic is defined as:

$$C_p = \left(\frac{RSS_p}{\hat{\sigma}^2} \right) - n + 2p$$

Where RSS_p is the residual sum of squares for the subset model containing p explanatory variables including the intercept (i.e., the number of parameters in the subset model), $\hat{\sigma}^2$ is an estimate of the error variance σ^2 . The mean-square error (MSE) from the full model is often used as the estimate of σ^2 . So $\hat{\sigma}^2 = MSE$ for the full model (full model means the model containing all explanatory variables of interest counting the intercept), n is the number of observations. Usually c_p is plotted against p for the collection of subset models. Acceptable models in the sense of minimizing the total bias of the predicted values are those models for which C_p approaches the value p . (i.e., those subset models that fall near the line $C_p = p$ in the plot).

Kobayashi and Sakata (1990) illustrated that subset selection using Mallows C_p is biased. This is because in some cases the probability of selecting one from two linear models is not equal. As Mallows C_p criterion gives greater selecting probability to the simpler model from two nested models equal in mean square predicted error. Kobayashi and Sakata presented a bias correction of the Mallows C_p by suggesting a critical value for statistical test. This test is

requested only in the case if the two models equals in Mallows C_p or the difference of Mallows C_p statistics is less than one in absolute value.

Ronchetti and Staudte (1994) modified a robust version of Mallows's C_p for regression models. It is defined by

$$RC_p = \frac{W_p}{\hat{\sigma}^2} - (U_p - V_p)$$

Where W_p is a weighted residual sum of squares computed from a robust fit of model with p independents, $\hat{\sigma}^2$ is a robust and consistent estimator of σ^2 in the full model, and U_p and V_p are constants depending on the weight function and the number of parameters in model p . The robust model selection procedure based on RC_p allows choosing the models that fit the majority of the data by taking into account the presence of outliers and possible departures from the normality assumption on the error distribution.

Gilmour (1996) indicated that if the MSE from the full model is used to estimate σ^2 , the distribution of $\left(\frac{RSS_p}{\hat{\sigma}^2}\right)$ can be obtained and gives an expected value of C_p which is not p . Moreover, Gilmour (1996) cited the words of Myers (1992) who said “Since the residual mean square for the complete model need not be the smallest estimate of σ^2 among those for the candidate models, it is quite possible that the equation will yield a $C_p < p$ for a few of the candidate models”. So Gilmour suggested an adjustment to C_p . The motivated criterion is called \bar{C}_p , an adjusted C_p has the property that if the estimate of σ^2 comes from the full mode $E(\bar{C}_p) = p$.

$$\bar{C}_p = C_p - \frac{2(k - p + 1)}{n - k - 3}$$

Where k is a total number of candidate regressors.

Sommer and Huggins (1996) suggested a criterion based on the Wald test statistic and is defined by T_p . The selected submodels is the model that have T_p -values close to or smaller than p , and, as with Mallows's C_p , they will be selected by graphical rather than stepwise methods.

$$T_p = W_p - k + 2p$$

where W_p is the Wald statistic for testing whether the coefficients of the variables not in the submodel are = 0, k and p are the numbers of parameters in the full and submodel respectively.

As T_p is equivalent to C_p in least squares linear regression, it may not be reliable in the presence of outliers. Sommer and Huggins suggested a robust version of T_p (RT_p), based on generalized M-estimator. RT_p need less computational time than RC_p because the weights are determined from the full model only. Furthermore, RT_p and T_p do not need to define residuals like C_p .

4. Distance between predicted values based on the full model \hat{Y}_k and predicted values based on the sub model \hat{Y}_p criterion:

Kashid and Kulkarni (2002) investigated a more general criterion called S_p criterion, for subset selection in the multiple linear regression models. S_p -criterion is unlike the subset selection methods which are based on the Least Squares (LS) Estimator. If the data contain an influential (outlier) observation or the distribution of the error variable is not normality, the LS estimator performs 'poorly' and hence a method based on this estimator (for example, Mallows' C_p -criterion) tends to select a 'wrong' subset. S_p -criterion overcomes this problem and it can be used with any type of estimator (either the LS estimator or any robust estimator).

$$S_p = \frac{\sum_{i=1}^n (\hat{Y}_{ik} - \hat{Y}_{ip})^2}{\sigma^2} - (k - 2p),$$

\hat{Y}_{ik} the vector of predicted values based on the full model and \hat{Y}_{ip} the vector of predicted values based on submodel, σ^2 the variance of the error.

Sp statistic is calculated for every model (for all possible values of p) and select the subset for which the corresponding value of Sp is approximately equal to p. the graph of the values of Sp against p will help in selecting a better subset.

Kashid and Kulkarni (2003) suggested a subset selection criterion in a multiple linear regression model in case if the distribution of error variable is heavy tailed. The motivated criterion called Np-criterion, and based on the nonparametric estimator of regression coefficients. It is explained that the performance of Np-criterion is better than RCp and Cp irrespective of the type of underlying distribution. Moreover Np is robust to outliers.

$$N_p = \frac{\sum_{i=1}^n (\hat{Y}_{ik} - \hat{Y}_{ip})^2}{\tau^2} - (k - 2p),$$

\hat{Y}_{ik} the vector of predicted values based on the full model and \hat{Y}_{ip} the vector of predicted values based on sub model, The scale parameter τ is similar to σ in the least square regression. τ measures the variability in the distribution of errors in non-parametric regression.

Dorugade and Kashid (2010 a) motivated a method called Rp-criterion. The authors used the method of ridge regression to estimate the model parameter. The study assumed that there is a problem of multicollinearity in the data. Rp- criterion is defined as:

$$R_p = \frac{\sum_{i=1}^n (\hat{Y}_{ik} - \hat{Y}_{ip})^2}{\sigma^2} - tr(H_R H_R) + tr(H_{Rp} H_{Rp}) + P$$

Where H_R and H_{RA} are the hat matrices for the the full model and the sub model respectively, and defined as follows: $H_R = X(X'X + rI)^{-1} X'$, $H_{Rp} = X_p(X'_p X_p + rI)^{-1} X'_p$, and X_p is an $n \times p$ matrix of the observations on p-1 predictors.

The procedure of R_p - criterion is beginning with computing the value of statistic for all possible subset models. Then select a subset of minimum size, for which the value of the statistic is close to 'p' or plot the values of statistic R_p vs. 'p' for all possible subset models and select the subset which is closer to the line i.e. R_p statistic = P .

Dorugade and Kashid (2010 b) proposed a new subset selection method called GR_p - criterion. They have used a various bias constant determination methods proposed by the researchers for the generalized ridge regression. GR_p - criterion is defined as:

$$GR_p = \frac{\sum_{i=1}^n (\hat{Y}_{ik} - \hat{Y}_{ip})^2}{\sigma^2} - tr(\hat{H}_{GR}H_{GR}) + tr(\hat{H}_{RA}H_{RA}) + P$$

Where, H_{GR} and H_{RA} denote the prediction matrix for the full model and subset model respectively based on the generalized ridge estimator.

One of an important advantage of the proposed method is that it can be used with least squares estimator or generalized ridge estimator of β without any modification in the proposed statistic.

Jadhav et al. (2014) developed a generalized version of S_p statistic for subset selection in the presence of outlier and multicollinearity. The performance of the proposed method is considerably better as compare to some alternative methods when the outlier observations and multicollinearity occur simultaneously in the data.

$$GS_p = \frac{\sum_{i=1}^n (\hat{Y}_{ik} - \hat{Y}_{ip})^2}{\sigma^2} - tr[(H - H_1)'(H - H_1)] + p,$$

Where, H and H_1 denote the prediction matrix for the full model and subset model respectively based on the Jackknifed Ridge M-estimator.

GS_p statistic is calculated for every model (for all possible values of p), then select the subset for which the corresponding value of GS_p is approximately equal to p .

5. Kullback's discrepancy criterion:

Some subset selection techniques depended on “Kullback's discrepancy” to measure the discrepancy between the unknown true likelihood function and a model- specific likelihood function. Kullback–Leibler (K-L) information or distance, The K-L distance has also been called the K-L discrepancy, divergence, information, and number (Kullback,1987 preferred the term discrimination information).

In terms of a probability distribution the full reality or truth model which we want to make inferences and an approximating model is noted f and g respectively. Kullback–Leibler information or distance between full reality or truth model and an approximating model is defined for continuous functions as the (usually multi-dimensional) integral.

$$I(f, g) = \int f(x) \log \left(\frac{f(x)}{g(x|\theta)} \right) dx,$$

We use the notation $\int f(x)$ to denote that integration is over the variable x , and use θ to represent generally a parameter or vector of parameters. Thus the notation $g(x|\theta)$ is refers to the estimated models have specific parameter values from ML or LS estimation based on the given data and model. \log denotes the natural logarithm. The notation $I(f, g)$ denotes the “Information lost when g is used to approximate f .” Also, $I(f, g)$ is the distance from g to f . **Burnham and Anderson (2002)**

Akaike (1974) demonstrated a procedure called Akaike information criterion AIC . AIC is an estimate of relative, expected K-L information between a fitted model and the true model. (AIC) and defined by:

$$AIC = -2 \log(L) + 2k$$

Where, L is the maximum likelihood estimates of the parameters, and k is the number of parameters within the model. AIC is calculated for a set of candidate models, and then choose the mode that has the least AIC value.

It should be noted that even if all candidate models are poor, *AIC* will select one to be the best. Thus choosing the model using *AIC* doesn't mean that the chosen model is will defined but it is the best among candidate models. Moreover *AIC* should not used to compare models from different data sets, because the inference is conditional on the given data set. *AIC* may perform poorly in small sample (when the sample size is small with respect to the number of estimated parameters). **Burnham and Anderson (2002)**

Schwarz (1978) introduced a procedure depended on Bayesian context Bayesian information criteria (BIC) (sometimes called Schwarz's information criterion (SIC)). From a Bayesian perspective, BIC is designed to find the most probable model given the data.

$$BIC = -2 \log(L) + k \log(n)$$

It is noted that BIC differs from AIC only in the second term which depends here on sample size n. Models that minimize the Bayesian Information Criteria are selected. [see Acquah (2010)]

Sugiura (1978) developed a version of *AIC* by modified the bias and obtained the criterion denoted by correct *AIC* (C- *AIC*).

$$C- AIC = (-2) \log(L) + \frac{2n(m+1)}{n-m-2}$$

Where *m* is the rank of the designed matrix.

Hurvich and Tsai (1989) developed a version of *AIC* that is applicable to the small sample size relation to the number of estimated parameters. They presented a small-sample bias adjustment, and introduced a modification that led to *AIC_c* criterion.

$$AIC_c = (-2) \log(L) + 2k \frac{n}{n - k - 1}$$

Konishi and Kitagawa (1996) proposed a subset selection criteria called Generalised information criteria (GIC). GIC is derived from the information-theoretic in the context of functional statistics without the assumption that the specified family of probability distributions contains the true structure generating the data. The proposed criteria is not restricted to maximum likelihood estimation but may be applied to the evaluation of models constructed by a different of methods, unlike AIC dealing models estimated by the maximum likelihood methods.

$$GIC(X_n; \hat{G}) := (-2n) \left\{ \frac{1}{n} \sum_{i=1}^n \log f(X_i | \hat{\theta}) - \frac{1}{n} b_1(\hat{G}) \right\}$$

Where X_n is a random sample of size n from an unknown distribution $G(x)$. θ is a p -dimensional vector of unknown parameters. $b_1(\hat{G})$ is a bias estimate obtained by replacing the unknown distribution G by the empirical distribution \hat{G}

Bengtsson and Cavanaugh (2006) modified the Akaike information criterion and introduced AIC_i that can be used in the broad framework of the linear state-space model. The subscript i denoted to “improved”. The results indicated that AIC_i estimates is less bias than traditional AIC or corrected AIC .

$$AIC_i = -2 \log(L) + \hat{\beta}_T(k, \theta_s)$$

$\hat{\beta}_T(k, \theta_s)$ is an approximation for bias and obtained by Monte Carlo simulation using a conveniently chosen simulation parameter θ_s

Jayakumar and Sulthan (2014) explained how the multicollinearity affected estimation using least square method and how it affected the model selection decision. The authors presented multicollinearity corrected version of generalized information criteria called modified generalized information criteria ($MGIC$) to select a best model among the various candidate models. The correction was by exchange the inflated error variance by un-inflated error variance of a model. The generalized information criterion is given as:

$$MGIC = n \log(\hat{\sigma}_{INF(e)}^2) + f(n, k)$$

Where $\hat{\sigma}_{INF(e)}^2$ is the inflated estimate of population error variance which is inflated by the variance inflation factor (VIF). $f(n, k)$ is the penalty function.

6. The Loss Rank Principle (LoRP) criterion:

Hutter (2007) proposed a modern criterion called The Loss Rank Principle (LoRP). Supset selection according to LoRP is through applying the following rule:

Given a class of linear models the best model is the one with the smallest loss rank i.e. the smallest M^{best} .

$$M^{best} = \arg \min \left\{ \frac{n}{2} \log(y S_{\alpha} y) - \frac{1}{2} \log \det S_{\alpha} \right\},$$

Where symbol refers to both model and regression matrices, I is the identity matrix, and $S_{\alpha} = (I_n - M)^T (I_n - M) + \alpha I$, where M is a projection matrix and $\alpha > 0$ is a small number to be determined. (Hutter and Tran, 2010) discussed the optimality Properties of LoRP for Variable Selection.

Tran (2011) proposed a LR criterion, for selecting shrinkage parameters for variable selection purposes. a LR criterion is derived from the loss rank principle (LoRP), that was introduced by Hutter (2007) and Hutter & Tran (2010). The model that has the lowest LR criterion is chosen as a best model.

$$LR_{\lambda} = \frac{n}{2} \log \|y\|^2 - \frac{n}{2} \text{KL} \left(\left(\frac{df_{\lambda}}{n} \mid 1 - \rho_{\lambda} \right) \right)$$

Where λ is a certain shrinkage parameter, df_{λ} equal the number of non-zero coefficients, and KL is the Kullback–Leibler divergence between the Bernoulli distributions with parameters $p, q \in (0, 1)$. $\rho_{\lambda} := \|y - \hat{y}_{s_{\lambda}}\|^2 / \|y\|^2$ and $\hat{y}_{s_{\lambda}}$ The fitted vector under the sub model. When λ increases from 0 to infinity, the number of non-zero coefficients will be a non-increasing step function i.e. the covariates are in turn removed from the models.

7 Other Criteria:

Greenberg and Parks (1997) provided a subset selection method based on predictive approach. The least square method is used to estimate the linear regression coefficients.

The effect of including X on the predictive covariance matrix is measured by generalized variance ratio (GVR). GVR is defined as the ratio of determinants of covariance matrices when a subset of explanatory variables is included or excluded from a regression.

$$GVR = \frac{|cov(y_0|X_{01})|}{|cov(y_0|X_0)|}$$

Where $cov(y_0|X_{01})$ is the covariance matrix from the partial model. While, $cov(y_0|X_0)$ is the covariance matrix from the full model.

Andre et al. (2000) modified the robust coefficient of determination for minimum sum of absolute errors MSAE regression.

$$R = RSR / (RSR + (n - p)\hat{\sigma}/2),$$

where RSR is the reduction in the sum of absolute residuals when P-1 independent variables included in the model i.e

$$RSR = \sum_{i=1}^n |y_i - \hat{\beta}_0| - \sum_{i=1}^n |y_i - \hat{y}_i|$$

and $\hat{\sigma} = SAR/n$, where SAR is the sum of absolute residuals.

Muller and Welsh (2005) investigated a method for subset selection based on combining a robust penalized criterion and a robust conditional expected prediction loss function that is estimated using a stratified bootstrap. The model selection method is consistent under some conditions.

Sakate and Kashid (2014) proposed a model selection criterion based on deviance called the deviance-based criterion (DBC). DBC is gotten by penalizing the difference between the fitted model and the full model. Under some conditions, DBC is appeared to be a consistent model selection criterion. The selected model asymptotically equals the optimal model relating dependant and independent variables.

8. References:

1. Acqua H. (2010), "*Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship*", Journal of Development and Agricultural Economics, Vol. 2, No.1, pp.1-6.
2. Akaike H. (1974), "*A new look at the statistical model identification*", IEEE Transactions on Automatic Control, Vol. AC-19, No. 6, pp. 716-723.
3. Andre C., Elian S., Narula S. and Tavares R. (2000), "*Coefficients of determinations for variable selection in the msae regression*", Communications in Statistics - Theory and Methods, Vol.29, No.3, pp. 623-642.
4. Bengtsson T. and Cavanaugh J. (2006), "*An Improved Akaike information criterion for state - space model selection*", Computational Statistics & Data Analysis, Vol. 50, Issue 10, pp. 2635–2654.
5. Burnham K. And Anderson D. (2002), "*Model selection and multimodel inference: A practical information-theoretic approach*", 2nd edition, Springer-Verlag, New York.
6. Dorugade A. And Kashid D. (2010 a), "*Variable selection in linear regression based on ridge estimator*", Journal of Statistical Computation and Simulation, Vol. 80, No. 11, pp. 1211-1224.
7. Dorugade A. And Kashid D. (2010 b), "*Subset selection in linear regression using generalized ridge estimator*", Journal of Statistical Theory and Practice, Volume 4, No. 3, pp. 357-389.
8. Gilmour S. (1996), "*The interpretation of Mallows's Cp-statistic*", Journal of the Royal Statistical Society. Series D (The Statistician), Vol. 45, No. 1, pp. 49-56.
9. Greenberg E. and Parks R. (1997), "*A predictive approach to model selection and multicollinearity*", Journal of Applied Econometrics, Vol. 12, No. , pp. 67-75.

10. Hurvich C. and Tsai C. (1989), "*Regression and time series model selection in small samples*", *Biometrika*, Vol. 76, No. 2, pp. 297-307.
11. Hutter M. (2007), "*The loss rank principle for model selection*", In Proceedings of the 20th Annual Conf. on Learning Theory (COLT'07), Vol. 4539 of LNAI (eds N. Bshouty & C. Gentile), pp.589-603, Springer, Berlin and San Diego.
12. Hutter M. and Tran M. (2010), "*Model selection with the loss rank principle*". *Comput. Statist. Data Anal*, Vol. 54, 1288–1306.
13. Jadhav N., Kashid D., and Kulkarni Subhash R. (2014), "*Subset selection in multiple linear regression in the presence of outlier and multicollinearity*", *Statistical Methodology*, Vol. 19, pp. 44-59.
14. Jayakumar D. and Sulthan A. (2014), "*Identification of multicollinearity and its effect in model selection*" *Electronic Journal of Applied Statistical Analysis*, Vol. 7, Issue 1, pp. 179-153.
15. John Dziak, Runze Li and Linda Collins (2005), "*Critical review and comparison of variable selection procedures for linear regression : Technical report*", Methodology Center, Penn State University.
16. Kashid D. and Kulkarni S. (2002), "*A more general criterion for subset selection in multiple linear regression*", *Commun. Statist.-Theory Meth.*, Vol. 31, No. 5, pp. 795–811.
17. Kashid D. and Kulkarni S. (2003), "*Subset selection in multiple linear regression with heavy tailed error distribution*", *Journal of Statistical Computation and Simulation*, Vol. 73, No. 11, pp. 791-805.
18. Kobayashi M. and Sakata S. (1990), "*Mallows 'Cp criterion and unbiasedness of model selection*", *Journal of Econometrics*, Volume 45, Issue 3, Pages 385-395.
19. Konishi S. and Kitagawa G. (1996), "*Generalised information criteria in model selection*", *Biometrika*, Vol. 83, No. 4, pp. 875-890.
20. Mallows C. (1973), "*Some comments on C_p* ", *Technometrics*, Vol. 15 , Iss. 4, pp. 661-675.

21. Miller A. (1990), "*Subset selection in regression*", Chapman & Hall/CRC Monographs on Statistics & Applied Probability, United States of America.
22. Muller S. and Welsh A. (2005), "*Outlier robust model selection in linear regression*", Journal of the American Statistical Association, Vol. 100, No.472, pp. 1297-1310.
23. Norusis M. (2006), "*SPSS 15.0 statistical procedures companion*", Prentice Hall.
24. Ronchetti E. and Staudte R. (1994), " *A robust version of Mallows's C_p* ", Journal of the American Statistical Association, Vol. 89, No. 426, pp.550-559.
25. Sakate D. and Kashid D. (2014), "*A deviance-based criterion for model selection in GLM*", A Journal of Theoretical and Applied Statistics, Vol. 48, No. 1, pp. 34-48.
26. Sommer S. and Huggins R. (1996), "*Variables selection using the wald test and a robust C_p* " Journal of the Royal Statistical Society. Series C (Applied Statistics), Vol. 45, No. 1, pp. 15-29.
27. Sugiura N. (1978), "*Further analysts of the data by Akaike' s information criterion and the finite corrections*", Communications in Statistics - Theory and Methods, vol.7, No.1, pp. 13-26.
28. Tran M. (2011), "*The Loss rank criterion for variable selection in linear regression analysis*", Scandinavian Journal of Statistics, Vol. 38, pp. 466–479.

Estimation of Unbalanced Panel Data

Elhoussainy A. R.¹

Ahmed A.E.²

Aya. M. A.³

Abstract Panel data set is one of the most important kinds of data sets now days because its help the researcher to look at dynamic relationships and that would not be happened with a single cross section and time series. In this paper the unbalanced panel data approach will be presented and literature review about unbalanced Panel data models will be presented.

Key words Fixed Effect, Random Effect, and Unbalanced Panel data.

1. INTRODUCTION

Panel data includes both diminutions of time and cross section. The time dimension in panel data is represented in the observations collected over time periods on each individual. So this kind of data provid rich sources of information about the economy. Missing values in Panel data sets is very common problem facing the researchers in analyzing Panel data, this problem leads to unbalanced panels. While the nature of these unbalanced panels doesn't affect applicability of many commonly used estimators (such as Fixed Effect or Random Effect estimator), but this is not true for all estimators.

2. MODEL

$$y_{it} = x_{it}\beta + \epsilon_{it}, \quad t = 1, 2, \dots, T, \quad i = 1, 2, \dots, N$$

Where $\epsilon_{it} = \mu_i + u_{it}$

- y_{it} Denotes the dependent variables
- β Denotes regression coefficient
- x_{it} Denotes the independent variables
- μ_i denotes the individual error for the cross section
- u_{it} Denotes the classical error term.

The difference between Random and Fixed Effect models is:

- $E(\mu_i x_{is}) = E(u_{it} x_{is}) = 0$ (Random Effect Model)
- $E(\mu_i x_{is}) \neq 0, E(u_{it} x_{is}) = 0$ (Fixed Effect Model)

3. Review of Unbalanced Panel Data Model

In this section literature review for unbalanced panel data will be presented

3.1 Fixed Effect Models

Alev, et al (2011) developed a semi-parametric panel model to explain the trend in UK temperatures and other weather outcomes over the last century. They used the monthly averaged maximum and minimum temperatures observed at the twenty-six Meteorological Office stations. The data is an unbalanced panel. They allowed the trend to evolve in a nonparametric way so that they obtained a fuller picture of the evolution of common temperature in the medium timescale.

Jason (2012) proposed the Chamberlain projection approach, it is a powerful tool for the analysis of linear fixed-effects models, was introduced within the context of balanced panels. He proposed the Chamberlain projection approach to unbalanced panels. The extension is especially useful for models with sequential exogeneity, where existing control-variable approaches are not applicable. A generalized method of moments (GMM) estimation framework is considered, and hypothesis tests (testing strict exogeneity, testing random effects, etc.) are discussed within the GMM context.

Jianhua, et al (2012) proposed missing data in longitudinal studies can create enormous challenges in data analysis when coupled with the positive-definiteness constraint on a covariance matrix. For complete balanced data, the Cholesky decomposition of a covariance matrix makes it possible to remove the positive-definiteness constraint and use a generalized linear model setup to jointly model the mean and covariance using covariates. **Kristiaan and Ignace (2014)** explored the effect of balancing unbalanced panel data when estimating primal productivity indices using non-parametric frontier estimators. First, they listed a series of pseudo-solutions aimed at making an unbalanced panel balanced. Then, they discussed some intermediate solutions (e.g., balancing 2-years by 2 years). Furthermore, they linked this problem with a variety of literatures on infeasibilities, statistical inference of non-parametric frontier estimators, and the index theory literature focusing on the dynamics of entry and exit in industries.

Wei, et al (2014) introduced “Planned missing designs” they became increasingly popular, but because there is no consensus on how to implement them in panel data research, they simulated panel data to distinguish between strategies of assigning items to forms and of

assigning forms to participants across measurement occasions. Using relative efficiency as the criterion, results indicate that balanced item assignment coupled with assigning different forms over time most often yields the optimal assignment method, but only if variables are reliable.

3.2 Random Effect Models

Jinhong and Xian (2008) proposed the estimation of the unbalanced panel data partially linear models with a one-way error components structure. A weighted semiparametric least squares estimator (WSLSE) is developed using polynomial spline approximation and least squares. They showed that the WSLSE is asymptotically more efficient than the corresponding unweighted estimator for both parametric and nonparametric components of the model.

Boris (2009) extended the (Least Squares Dummy Variables) LSDV bias-corrected estimator in (Bun and Carree, 2005) to unbalanced panels and discusses the analytic method of obtaining the solution. Using a Monte Carlo approach, he compared the performance of this estimator with three other available techniques for dynamic panel data models.

Jeffrey (2009) proposed some strategies for allowing unobserved heterogeneity to be correlated with observed covariates and sample selection for unbalanced panels. The methods are extensions of the Chamberlain-Mundlak approach for balanced panels. Even for nonlinear models, in many cases the estimators can be implemented using standard software. The framework suggests straightforward tests of correlation between heterogeneity and the covariates, as well as sample selection that is correlated with unobserved shocks while

allowing selection to be correlated with the observed covariates and unobserved heterogeneity.

David (2010) proposed a simple analytical approach for constructing one-sided β -content, γ confidence tolerance limits is proposed for general random effects models with normal data in both balanced and unbalanced data scenarios. The approach is based on an approximation to the noncentral t distribution and modified large sample methods for constructing confidence bounds on functions of variance components. An alternative bootstrap-adjusted limit is also proposed.

Jun, et al (2010) introduced panel data with covariate measurement error appear frequently in various studies. Due to the sampling design and/or missing data, panel data are often unbalanced in the sense that panels have different sizes. For balanced panel data (i.e., panels having the same size), there exists a generalized method of moments (GMM) approach for adjusting covariate measurement error, which does not require additional validation data. They extended the GMM approach of adjusting covariate measurement error to unbalanced panel data. Two health related longitudinal surveys are used to illustrate the implementation of the proposed method.

Erik (2013) proposed a system of regression equations for analyzing panel data with random heterogeneity in intercepts and coefficients, and unbalanced panel data is considered. A maximum likelihood (ML) procedure for joint estimation of all parameters is described. Since, its implementation for numerical computation is complicated, simplified procedures are presented. The simplifications essentially concern the estimation of the covariance matrices of the random coefficients. The application and 'anatomy' of the proposed algorithm for modified ML

estimation are illustrated by using panel data for output, inputs and costs for 111 manufacturing firms observed up to 22 years.

Wei and Lung (2013) suggested and compared different methods for estimating spatial autoregressive panel models with randomly missing data in the dependent variable. They started with a random effects model and then generalize the model by introducing the spatial Mundlak approach. A nonlinear least squares method is suggested and a generalized method of moments estimation is developed for the model. A two-stage least squares estimation with imputation is proposed as well. They analytically compared these estimation methods and find that the generalized nonlinear least squares, best generalized two-stage least squares with imputation, and best method of moments estimators have identical asymptotic variances.

Liwen and Dengkui (2016) presented parametric bootstrap (PB) approaches for hypothesis testing and interval estimation for the regression coefficients of panel data regression models with incomplete panels. Some simulation results are presented to compare the performance of the PB approaches with the approximate inferences. They showed that the PB approaches perform satisfactorily for various sample sizes and parameter configurations, and the performance of PB approaches is mostly better than the approximate methods with respect to the coverage probabilities and the type I error rates.

References

1. Alev, A., Oliver, L. and Zhijie X. (2011). A semiparametric panel model for unbalanced data with application to climate change in the United Kingdom. *Journal of Econometrics*. 164, 92 – 115.
2. Boris, L. (2009). Further results on bias in dynamic unbalanced panel data models with an application to firm R&D investment. *Applied Economics Letters*.1227-1233.
3. Erik, B. (2013). Estimating SUR system with random coefficients: the unbalanced panel data case. *Empir Econ*. 47, 451–468.
4. Jason, A. (2012). The projection approach for unbalanced panel data. *Econometrics Journal*. 16,161–178.
5. Jeffrey, M. (2009). Correlated Random Effects Models With Unbalanced Panels. Department of Economics, Michigan State University. Presented at the 15th Conference on Panel Data, Bonn, Germany, July 3-5, 2009.
6. Jianhua, Z., Min,C., Mehdi,M. and Mohsen,P. (2012). A cautionary note on generalized linear models for covariance of unbalanced longitudinal data. *Journal of Statistical Planning and Inference*.142, 734-751.
7. Jinhong, Y. and Xian, Z. (2008). Partially linear models and polynomial spline approximations for the analysis of unbalanced panel data. *Journal of Statistical Planning and Inference*139, 679-695.
8. Jun,S., Zhiguo, X. and Ruifeng ,X. (2010). Estimation with unbalanced panel data having covariate measurement error. *Journal of Statistical Planning and Inference*. 141, 800-808
9. Kristiaan, K. and Ignace, V. (2014). Comparing Malmquist and Hicks–Moorsteen productivity indices: Exploring the impact of unbalanced

- vs. balanced panel data. *European Journal of Operational Research*.233, 749–758.
10. Wei, W. and Lung, F. (2013). Estimation of spatial panel data models with randomly missing data in the dependent variable. *Regional Science and Urban Economics*. 43, 521-538.
 11. Wei, W., Terrence, D., Mijke, R., Alexander, S., Brent, M., and Todd, D. (2014). Optimal assignment methods in three-form planned missing data designs for longitudinal panel studies. *International Journal of Behavioral Development*. 38(5), 397–410.

Websites

12. David, H. (2010). One-Sided Tolerance Limits for Balanced and Unbalanced Random Effects Models. 303-312. Taylor and Francis Group.
(http://amstat.tandfonline.com/doi/abs/10.1198/TECH.2010.09171#.WinA_FWGPV)
13. Liwen, X. and Dengue, W. (2016). Parametric bootstrap inferences for unbalanced panel data models. *Communications in Statistics - Simulation and Computation*. Published online: 24 Oct 2016
(<http://www.tandfonline.com/doi/abs/10.1080/03610918.2016.1248567?journalCode=lssp20>)

Extracting the Wealth of Knowledge in Composite Index Data

Asmaa Mohamed Sayed¹

Abstract:

Composite indices data are high-dimensional data, where the number of variables sometimes exceeds the number of observations. A composite index is a single number summary for each observation in the data. No matter how informative a single number is, it can not capture all the features of such highly-dimensional data. One purpose of this paper is to look at these data from various angles using different techniques in an attempt to discover and extract the wealth of information or knowledge that these data contain. We particularly focus our attention on three challenging issues: (a) The choice of weights for indicators, (b) How to avoid measuring the standard error of the index, and (c) High dimensionality of the data. We look at the data from a multivariate point of view and use existing techniques (such as the Principal Components Analysis, Multidimensional Scaling) to develop new alternative indices for the same data.

We illustrate these methods using the Ibrahim Index of African Governance (IIAG) data for 2016, the most recent year for which the data are available at the time of writing. These data are provided by 34 well known and trusted providers. The data are collected on 54 African countries including Egypt. The data consists of 154 variables, some of which are grouped into indicators. The total number of indicators is 95. These are grouped into 4 main categories, and 14 subcategories. The main purpose is to compute an index number for each of the African countries that can be used to assess the overall performance of 54 African countries in terms of their abilities to deliver public goods and services.

Key Words:

Bootstrap; Cluster Analysis; Composite Indices; Hierarchical Clustering; Indicator Weights; Imputation; K-means Algorithm; Margin of Error; Missing Observations; Multidimensional Scaling; Principal Components Index.

¹M.Sc. Candidate, Department of Applied Statistics and Econometrics, Institute of Statistical Studies and Research, Cairo University.

1 Introduction

A composite index is a combination of various sources of information known as variables and indicators, measured in or of a system in order to provide a summary of the system that is itself not directly measurable. Composite indices have venerable history. Budro (2008) provides a survey of the current composite indices around the world. At that time Budro (2008) found 187 indices. Examples of composite indices are: the corruption perception index (CPI), Lambsdorff (2012) Modified corruption Index (MCPI), The Human Development Index (HDI), De Muro, Mazziotta, and Pareto (2011) proposal of modification of Human Development Index (MHDI), the Ibrahim Index of African Governance (IIAG), the Gender Inequality Index (GII), the Climate Change performance Index (CCPI), and The consumer Price Index (CPI). to mention only a few.

Most recently, the International Knowledge Index (IKI) was computed for the first time and published in November 2017 by the United Nations Development Program (UNDP). The IKI extended the Arab Knowledge Index (AKI) which was computed for the first time in 2015 by the Al Maktoum Foundation (<http://www.mbrf.ae/>) to measure knowledge in the Arab countries, the Happiness Index (among youth 18–29 years) by Abo Elazm (2017), and Child Wellbeing Index by Rabea (2017).

The construction of a composite index usually starts with defining the concept of interest that needs to be measured. For example, corruption as in the CPI, gender inequality as in the GII, governance as in the IIAG, or Knowledge as in the AKI and IKI. These concepts are difficult to measure directly. So experts in the domain field select several variables to be used as proxy measures of the concept of interest. Some of these variables are grouped into indicators. An indicator is usually a weighted sum (e.g., the mean) of the variables included in it. Thus an indicator can be thought of as an index of the variables involved. For simplicity we will use the terms variables and indicators synonymously.

The variables and/or indicators are then partitioned into groups (categories), subgroups, and sometimes subsubgroups. Although these groupings are obtained after extensive discussion among various experts in the field, experts do not always agree on these groupings. Even when they agree, the question is what weights should be given to indicators within subgroups, to subgroups within a group, and to groups

within the overall index. In some indices, the weights are given by experts or derived at by other means. In other indices the weights of indicators within a subgroup are given equal weight, the subgroups within a group are given equal weight, and finally the groups are given equal weight to obtain the final composite index.

For example, the IIAG consists of measurements on 54 African countries. For each country, 154 constituent variables, taken from 34 data providers, are then grouped into 95 indicators. These 95 indicators are divided into four overarching categories. Each of the four categories consists of subcategories. The total number of subcategories is 14. All indicators within a subcategory are given equal weights, all subcategories within a category are given equal weights, and finally all categories are given equal weights to reach the final IIAG.

There are several drawbacks of such construction of indices. The indicator weights depend on how one divides the indicators into groups and subgroups. The indicator weights will not be equal unless all groups have equal number of subgroups and all subgroups have equal number of indicators. Second, although the division of indicators into groups and subgroups are usually arrived at after careful consideration by various experts in the field, other partitioning of the indicators can be reasonably obtained, yielding possibly different indices of the same data. Third, it seems that this choice of weighting scheme is based on simplicity but not on any optimality criteria. We propose other complementary ways of constructing indices (scores) derived from composite index data that do not depend partitioning the indicators or a priori choice of the weights.

For the index to be useful and meaningful, its construction requires careful consideration of several important aspects of the potentially disparate and multiple indicators that help convey its meaning, Dobbie and Dail (2013).

The construction of composite indices are conceptually easy, but in practice several challenges are encountered. Perhaps, the most important one of which is data quality because the quality of an index cannot exceed the quality of the data that are used in its construction. Some of these challenges are discussed in the following sections.

1.1 Data Availability

It goes without saying that a good index should use good and well-known data providers. The data sources can provide raw or modified data. Variables should be selected based on their relevance, accessibility, and timeliness. Proxy measures can be used if the desired variables are not available. The accuracy of proxy measures can be investigated through sensitivity analysis and correlation. Their consistency and reliability can be investigated, for example, by the use of Cronbach's α (Cronbach, 1951). Sometimes practitioners tend to use whatever is available, but poor data quality will produce a poor index. A good composite index should allow improvement and development of its method of construction and renew its data over time.

1.2 Indicator Scales

Because the data come from different sources and in different units of measurements, normalization is required before any aggregation. Three of the common scaling methods mentioned in the OECD 2008 are:

1. Standardization: This method subtracts the mean of each indicator x from each score x_i then divides by the standard deviation of the scores of each indicator, that is,

$$z_i = \frac{x_i - \bar{x}}{s_x}, \quad i = 1, \dots, n, \quad (1)$$

where x_i is the i th value of the indicator x , \bar{x} and s_x are the mean and standard deviation of the indicator x , and n is the number of observations. The resultant standardized indicator z will have a mean of 0 and standard deviation 1.

2. Min-Max Normalization: This method converts the score of an indicator X by subtracting the minimum from each score then dividing by the range of the scores, that is,

$$y_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}, \quad i = 1, \dots, n. \quad (2)$$

The values of the resultant normalized indicator y will have values from 0 to 1.

The main disadvantage of these methods, however, is that they are not robust, that is, they are affected by the presence of outliers. Of course, this disadvantage can be avoided by replacing the mean and standard deviation but more robust measures of location and scale.

3. Ranking: The scores in an indicator are replaced by their ranks. Ranking is perhaps the simplest method of normalization and it is not affected by outliers. This method allows performances of countries to be compared to each other or a country to be compared to itself based on its position or rank over time.

1.3 The Presence of Missing Values

Composite index data often contain missing values. The presence of missing values threatens the validity of the statistical inference. Missing values can be categorized as follows (Donders et al. (2006)):

1. Missing at random. Missing values don't depend on the variable itself but they are conditional on other variables in the data set.
2. Missing completely at random. Missing values are not related to the variable or any other variables in the data. Like losing a questionnaire of a study accidentally. Statistical inference in that case will not be biased but of course it will be less efficient.
3. Not missing at random. Missing values depend on the values themselves. Like people of high income tends not to report their income.

There are various methods of handling missing data. The selection of a method is related to the type of the missing observations. The methods of handling missing values include:

1. Case Deletion: Exclude any incomplete observation from the analysis. In this case we end up with less information and the standard error of the index will increase in the reduced sample.
2. Single Imputation: Replace each missing value by newly imputed value. For example, a missing value can be replaced by the mean or the median, or by other methods such as regression imputation and expectation-maximization imputation. Single imputation underestimates the variance, because it partially reflects the uncertainty in the imputation.

3. Multiple Imputations: Replace each missing value by several newly imputed values. It can more effectively represent the uncertainty due to the imputation. Most of the statistical packages nowadays offer the multiple imputation technique like Monte Carlo simulation.

1.4 The Presence of Outliers

Outliers are a minority of observations that are inconsistent with the pattern suggested by the majority of observations in a data. Outliers usually lie an abnormal distances from the other observations in the data. There are several methods for the identification of outliers in composite index data. Outliers can be identified for each individual indicator or by considering indicators jointly as a multivariate data.

An example of the identification of outliers in an individual indicator is described as follows: The $\alpha\%$ trimmed mean and trimmed standard deviation are computed based on the central $(100 - \alpha)\%$ of the values. Commonly used choices of α are 5% and 10%. Then *all* observations more than 3 trimmed standard deviations away from the trimmed mean are then replaced by:

$$\textit{Trimmed mean} + 3.1 \times \textit{Trimmed standard deviation}, \quad (3)$$

if they are in the right tail, or by

$$\textit{Trimmed mean} - 3.1 \times \textit{Trimmed standard deviation}, \quad (4)$$

if they are in the left tail.

Because this rule is simple and it strikes a balance between efficiency and robustness, it is adopted by well-established indices, such as the Ibrahim Index of African Governance (IIAG) (see the MIF foundation Web site at: mo.ibrahim.foundation).

Identifying outliers in individual indicators may not be able to detect multivariate outliers, that is, outliers in the multivariate space. There are several outlier identification methods for multivariate data. See, for example, Rocke and Woodruff (1996), Rousseuw and Van Driessse (1999), and Billor et al. (2000).

1.5 The Choice of Indicator Weights

Weights can have a significant effect on the overall results of a composite index. A number of weighting schemes that exist in the literature include the following (see, e.g., Dobbie and Dail, 2013):

1. Equal Weights. This approach is commonly used because of its simplicity. It is based on the hypothesis that all indicators are equally important.
2. Data-Driven Weights. This approach is adaptive, that is, the weights are chosen depending on the data. Some of the methods of adaptive weights include (see, e.g., Decancq and Lugo, 2013):
 - (a) Frequency based weights. The choice of weights is based on the proportional of population that suffering deprivation in an indicator. For example the smaller proportion of individuals that suffer from a certain deprivation the higher weight. Or on the other hand giving less weight to those variables where problems exist.
 - (b) Expert Opinions. Experts in the domain field may be asked about their assessments of the relative weights for the indicators. A summary measure of their weights (e.g., the mean) can be used as the weights of indicators in the construction of the index.

1.6 The Margins of Errors

Composite index numbers are subject to various types of errors such as the use of proxies due to the difficulties in obtaining direct measurements, the variability in the methodology, completeness of data sources, the use of imputation of missing values, etc. Therefore, index numbers should be accompanied by margins of errors or confidence bands that reflect their imprecision. This is important because the index values can be used to rank the countries. But is a given country's rank really different from other countries with neighboring ranks?

To answer this question, we need to compute the standard error of the index values that can be used to construct confidence intervals for the index values. From these intervals we can infer whether the rank of a country is significantly different from

the ranks of the other countries. The problem here is that the standard errors of the index values are not theoretically available. A way out of this is to use the Bootstrap method (Efron and Tibshirani, 1993) to estimate the standard errors. This is used, for example, by the Ibrahim Index of African Governance (IIAG).

The Bootstrap as used in the IIAG is described as follows:

1. Draw a large number M (e.g., 1,000) bootstrap samples from the IIAG data (95 indicators in 2016) within their respective subcategories (14 in 2016).
2. For each bootstrap sample compute the IIAG for the subcategories, the categories, and the over all index.
3. Compute the standard error, $SE(M_{ti})$, of the M scores, where M_{ti} is the IIAG score for a subcategory, a category, or the overall index.
4. Compute the $(1 - \alpha)100\%$ confidence bands for the expected value $\mathbb{E}(M_{ti})$ by

$$M_{ti} \pm z_{\alpha/2} SE(M_{ti}), \quad (5)$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ quantile of the standard normal distribution.

The overall Ibrahim Index of African Governance (IIAG), M_{ti} , for 2016 together with the MIF bootstrap standard errors SE are shown in Table 1.

The confidence intervals in Table 1 can be used to determine if there are significance differences among country scores and/or among their ranks. For example, for the five highest ranked countries (1 = Mauritius, 2 = Botswana, 3 = Cabo Verde, 4 = Seychelles, and 5 = Namibia), one can see that Mauritius has a significantly higher score and rank than the other four countries, whereas the other four countries are not significantly different from each other.

One advantage of the bootstrap methodology for constructing the error bands is that it is nonparametric, that is, it does not make assumptions about the probability distribution of the indicators. However, this methodology has its own disadvantages.

1. It is computationally intensive because it requires a huge numbers of bootstrap samples for the resulting error bands to be reliable.
2. The resulting error bands are random, that is, it gives different error bands every time the method is applied to the same data.

Table 1: The 2016 IIAG and the Bootstrap Standard Errors and 95% Conf. Limits.

No.	ISO	Country	IIAG	Rank	SE	Low Limit	High Limit
1	DZ	Algeria	53.8	20	2.3	49.3	58.3
2	AO	Angola	39.2	45	2.5	34.4	44.0
3	BJ	Benin	57.5	16	1.9	53.8	61.2
4	BW	Botswana	73.7	2	2.2	69.4	78.0
5	BF	Burkina Faso	51.8	23	2.3	47.3	56.3
6	BI	Burundi	41.9	43	2.2	37.6	46.2
7	CV	Cabo Verde	73.0	3	2.1	68.8	77.2
8	CM	Cameroon	45.7	38	2.1	41.6	49.8
9	CF	Central African Republic	25.7	52	2.4	21.0	30.4
10	TD	Chad	34.8	48	2.4	30.2	39.4
11	KM	Comoros	50.3	26	2.6	45.1	55.5
12	CD	Congo	43.0	42	2.5	38.1	47.9
13	CI	Cote d'Ivoire	52.3	21	1.6	49.1	55.5
14	CG	Democratic Republic of Congo	35.8	46	2.5	30.8	40.8
15	DJ	Djibouti	46.5	37	2.7	41.1	51.9
16	EG	Egypt	51.0	24	2.5	46.1	55.9
17	GQ	Equatorial Guinea	35.4	47	2.9	29.7	41.1
18	ER	Eritrea	30.0	50	2.5	25.1	34.9
19	ET	Ethiopia	49.1	31	2.2	44.7	53.5
20	GA	Gabon	48.8	32	2.3	44.2	53.4
21	GM	Gambia	46.6	35	2.4	41.8	51.4
22	GH	Ghana	63.9	8	1.9	60.2	67.6
23	GN	Guinea	43.3	41	2.0	39.4	47.2
24	GW	Guinea-Bissau	41.3	44	2.6	36.2	46.4
25	KE	Kenya	58.9	12	2.0	55.0	62.8
26	LS	Lesotho	57.8	15	2.0	53.9	61.7
27	LR	Liberia	50.0	28	2.1	45.8	54.2
28	LY	Libya	29.0	51	2.9	23.3	34.7
29	MG	Madagascar	48.5	34	1.9	44.8	52.2
30	MW	Malawi	56.6	17	1.8	53.2	60.0
31	ML	Mali	50.6	25	1.8	47.0	54.2
32	MR	Mauritania	43.5	40	2.5	38.7	48.3
33	MU	Mauritius	79.9	1	1.9	76.3	83.5
34	MA	Morocco	58.3	14	2.3	53.8	62.8
35	MZ	Mozambique	52.3	22	1.8	48.9	55.7
36	NA	Namibia	69.8	5	1.8	66.2	73.4
37	NE	Niger	50.2	27	2.2	45.8	54.6
38	NG	Nigeria	46.5	36	2.1	42.3	50.7
39	RW	Rwanda	62.3	9	2.5	57.5	67.1
40	ST	Sao Tome and Principe	60.5	11	2.4	55.8	65.2
41	SN	Senegal	60.8	10	2.0	56.9	64.7
42	SC	Seychelles	72.6	4	2.2	68.3	76.9
43	SL	Sierra Leone	49.4	30	1.9	45.6	53.2
44	SO	Somalia	10.6	54	2.1	6.60	14.6
45	ZA	South Africa	69.4	6	2.1	65.4	73.4
46	SS	South Sudan	18.6	53	2.2	14.3	22.9
47	SD	Sudan	30.4	49	2.2	26.0	34.8
48	SZ	Swaziland	49.7	29	2.0	45.9	53.5
49	TZ	Tanzania	56.5	18	2.0	52.6	60.4
50	TG	Togo	48.5	33	2.1	44.4	52.6
51	TN	Tunisia	65.4	7	2.3	60.9	69.9
52	UG	Uganda	56.2	19	1.9	52.5	59.9
53	ZM	Zambia	58.8	13	1.7	55.5	62.1
54	ZW	Zimbabwe	44.3	39	2.1	40.2	48.4

3. It is done on the subcategory level, where the number of values to bootstrap is as small as 5 or even less.
4. The bootstrap assumes that the indicators are interchangeable within their respective subcategories.

These disadvantages cast doubt as to the appropriateness and reliability of using the bootstrap for estimating the margins of errors. One objective of this paper is to propose methods that do not depend on the computation of the margins of errors (Section 3).

In this paper, we focus our attention on three challenging issues: (a) The high dimensionality of the data, (b) The choice of weights for the indicators, and (c) Measuring the margins of errors of the index.

The rest of this paper is organized as follows: Section 2 examines alternative ways of constructing composite indices that do not depend on partitioning of the indicators. In particular, Section 2.1 considers an index which gives equal weights to all indicators. Section 2.2 proposes an index based on the Principal Components. An important advantage of these indices is that they do not depend on, nor do they require any particular partitioning of indicators into groups or subgroups.

Section 3 proposes ways of clustering the multivariate composite indices data. We can avoid both the issues of grouping variables/indicators and of the need to compute the standard errors by using clustering as explained in Section 3. We illustrate these methods using the Ibrahim Index of African Governance and the HDS2014 Children Well-Being data. This Demographic and Health Survey for Egypt is used to create measure the deprivations of the Egyptian children. The data contain three groups: Children under five years old, children aged 5 to 17 years old, and all children aged 0 to 17. The first group, contains seven deprivation indicators these indicators are: Education, Shelter, Health, Water, Sanitation, Knowledge and Nutrition. for the other two groups, the Nutrition deprivation indicator is dropped from the dominance (for details see, Abd ElAziz, 2017). Finally, summary and concluding remarks are given in Section 4.

2 Alternative Ways of Constructing Composite Indices

As alternative to the composite indices which are based on partitioning the indicators into groups and subgroups, we consider the equal-weight index (Section 2.1), where all indicators are given equal weights without grouping them. We also propose another index based on the Principal Components (Section 2.2). We then compare these indices with the classical way of index construction using the African Governance data.

2.1 Equal-Weights Index (EWI)

Composite indices are constructed by first, grouping the variables/indicators into groups, subgroups, and sometimes subsubgroups. Even within a subgroup, one may find indicators each of which is a composite index of several variables or indicators. Although these groupings are obtained after extensive discussion among the experts that are involved in the construction of the index, experts do not always agree on these groupings. Even when they agree, the question is what weights should be given to indicators within subgroups and to subgroups within a group. In some indices, the weights are given by experts or derived at by other means. In other indices the weights of indicators within a subgroup are given equal weight, the subgroups within a group are given equal weight, and finally the groups are given equal weight to obtain the final composite index.

To see the impact of different grouping and weighting schemes on the computation of the index, we consider two ways of computing the IIAG. The first is the way used by the Ibrahim Foundation, where all indicators within each sub-category are given equal weights, all sub-categories within each category are given equal weights, and all categories are given equal weights to arrive at the overall IIAG. As pointed out earlier, one advantage of this choice of “equal” weighting seems to be its simplicity, although in actuality, the final indicator weights are not equal but very different. This is because the number of indicators are not equal across subcategories and the number of subcategories are not equal across categories.

The second way is simply to ignore the groupings and subgrouping of variables and indicators and treat the variable generically by giving them equal weights. Specif-

ically, let \mathbf{X} the data matrix containing all indicators as columns and all observations as rows. The Equal-Weights index (EWI) assigns equal weights to all indicators (without grouping) and the index in this case will simply be the arithmetic mean of the rows of the data matrix \mathbf{X} . In the case where the i th observation contain missing values, they can be ignored or replaced by the average of that row (or any other desired imputed value). Note that the EWI here is different from the indices based on equal weights which are assigned after grouping the indicators into groups and subgroups. The EWI does not partition the indicators.

The EWI is simple to compute and explain. Another important advantage of EWI is that it does not depend on, nor does it require, any particular partitioning of indicators into groups or subgroups as is the case for many composite indices. One obvious drawback of the EWI is that the choice of truly equal weights seem to be selected for simplicity.

To see the impact on grouping and weights on the IIAG index, Table 2 shows the IIAG rankings of the 54 countries in 2016, the rankings based on equal weights, and the differences between the two rankings. ranking versus the IIAG ranking of the 54 countries in 2016. Figure 1 also shows the scatter plot of equal weight ranking versus the IIAG ranking of the 54 countries in 2016. One can see from Figure 1 that the two methods are strongly related but there are substantial differences among the two rankings of the countries. For example, as can be seen in Table 2, the IIAG ranking of Swaziland is 29 but the equal weights ranking is 22, a difference of 7 positions.

2.2 Index Based on the Principal Components (PCI)

One way to deal with the high dimensionality of the data is to use multivariate dimension reduction techniques. A common technique for dimension reduction is the principal components analysis (PCA). In PCA we find an $n \times p$ matrix $\mathbf{Y} = \mathbf{XV}$, where \mathbf{V} is chosen such that the variables in \mathbf{Y} are orthogonal to each other and have decreasing variances. The variables in \mathbf{Y} , which are linear functions of the variables in \mathbf{X} , are called the principal components (PCs) of \mathbf{X} .

To find the PCs, let \mathbf{S} be the variance-covariance matrix of \mathbf{X} and $\lambda_1 \geq \dots \geq \lambda_p$ be the ordered eigenvalues of \mathbf{S} and $\mathbf{V} = (V_1, \dots, V_p)$ be the matrix containing the corresponding eigenvectors. It can be shown that the PCs of \mathbf{X} are given by

Table 2: The IIAG and Equal Weight Rankings of the 54 Countries in 2016.

No.	ISO	Country	IIAG	EWI	Difference
1	DZ	Algeria	20	20	0
2	AO	Angola	45	44	1
3	BJ	Benin	16	16	0
4	BW	Botswana	2	2	0
5	BF	Burkina Faso	23	24	1
6	BI	Burundi	43	43	0
7	CV	Cabo Verde	3	3	0
8	CM	Cameroon	38	37	1
9	CF	Central African Republic	52	52	0
10	TD	Chad	48	48	0
11	KM	Comoros	26	29	3
12	CD	Congo	42	41	1
13	CI	Cote d'Ivoire	21	25	4
14	CG	Democratic Republic of Congo	46	47	-1
15	DJ	Djibouti	37	35	2
16	EG	Egypt	24	21	3
17	GQ	Equatorial Guinea	47	46	1
18	ER	Eritrea	50	49	1
19	ET	Ethiopia	31	26	5
20	GA	Gabon	32	30	2
21	GM	Gambia	35	34	1
22	GH	Ghana	8	9	1
23	GN	Guinea	41	42	1
24	GW	Guinea-Bissau	44	45	1
25	KE	Kenya	12	13	1
26	LS	Lesotho	15	15	0
27	LR	Liberia	28	31	3
28	LY	Libya	51	51	0
29	MG	Madagascar	34	36	2
30	MW	Malawi	17	19	2
31	ML	Mali	25	27	2
32	MR	Mauritania	40	40	0
33	MU	Mauritius	1	1	0
34	MA	Morocco	14	10	4
35	MZ	Mozambique	22	23	1
36	NA	Namibia	5	6	1
37	NE	Niger	27	28	1
38	NG	Nigeria	36	38	2
39	RW	Rwanda	9	7	2
40	ST	Sao Tome and Principe	11	12	1
41	SN	Senegal	10	11	1
42	SC	Seychelles	4	4	0
43	SL	Sierra Leone	30	33	3
44	SO	Somalia	54	54	0
45	ZA	South Africa	6	5	1
46	SS	South Sudan	53	53	0
47	SD	Sudan	49	50	1
48	SZ	Swaziland	29	22	7
49	TZ	Tanzania	18	17	1
50	TG	Togo	33	32	1
51	TN	Tunisia	7	8	1
52	UG	Uganda	19	18	1
53	ZM	Zambia	13	14	1
54	ZW	Zimbabwe	39	39	0

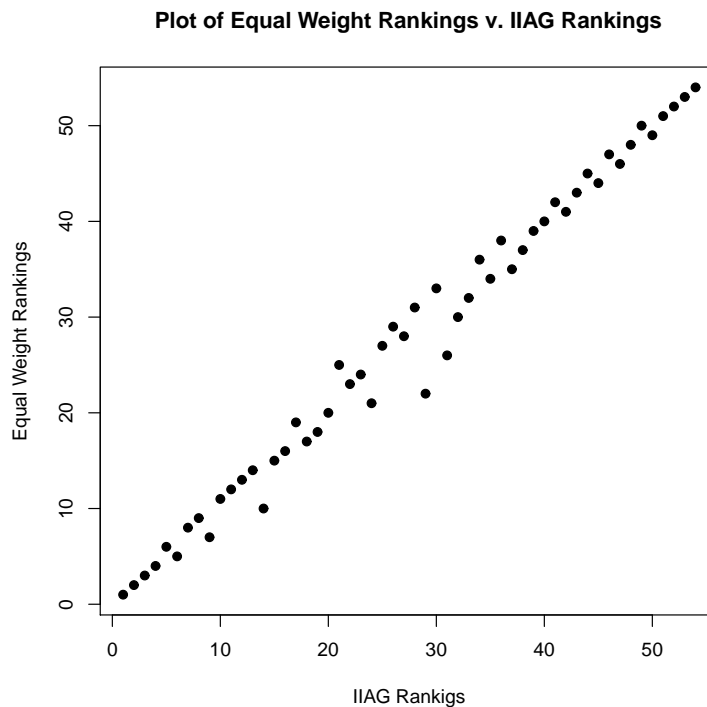


Figure 1: Plot of Equal Weight Rankings v. IIAG Rankings.

$\mathbf{Y} = \mathbf{XV}$. Furthermore, the variance of the j -th column of \mathbf{Y} is the j -th eigenvalue $\lambda_j, j = 1, \dots, p$. If $p > n$, as is the case here, it can also be shown that the last $p - n = 95 - 54 = 41$ eigenvalues are zeros, which means that each of the last 41 PCs have zero variance. Accordingly, the total variance of the variables in \mathbf{X} is equal to the total variance of the first 54 variables in \mathbf{Y} . Thus we can use the first 54 variables in \mathbf{Y} instead of \mathbf{X} without loss in the variance. Further reduction is possible if the variables in \mathbf{X} are highly correlated but with some loss of the variance. For example if we decide to keep only $k \ll 54$ PCs, then the total variance of the first k variables in \mathbf{Y} accounts for

$$\frac{\sum_{j=1}^k \lambda_j}{\sum_{j=1}^{54} \lambda_j} \times 100\% \quad (6)$$

of the total variance in \mathbf{X} . For example for the 2016 data, the first $k = 2$ components account for 44% of the total variance, whereas the first $k = 5$ components account for 66%, the first $k = 20$ components account for 98%, and the first $k = 24$ components account for 100% (this is because $\lambda_{25} = 0$).

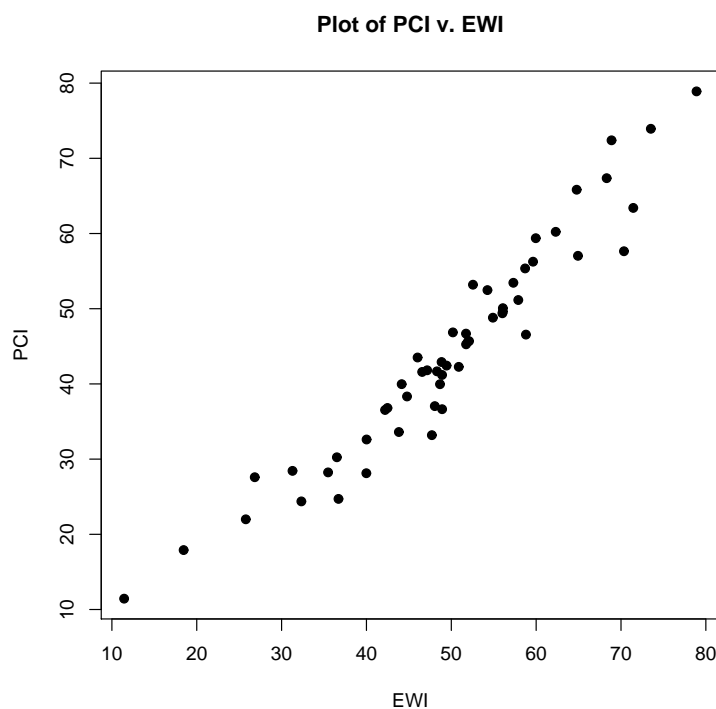


Figure 2: Scatter Plot of CPI v. EWI of the 2016 African Governance Data.

The steps for computing an index based on the Principal Components (PCI) are:

1. Compute the ordered eigenvalues $(\lambda_1, \dots, \lambda_p)$ and corresponding eigenvectors $\mathbf{V} = (V_1, \dots, V_p)$ of the matrix $\mathbf{X}^T \mathbf{X}$.
2. Set $\boldsymbol{\alpha} = \mathbf{V}_1$ and compute $\mathbf{U} = \mathbf{X}\boldsymbol{\alpha}$. If the correlation between \mathbf{U} and EWI is negative, then replace \mathbf{U} by $-\mathbf{U}$.
3. The PCI scores can then be transformed to an index like measure by

$$\text{CPI}_i = \min(\text{EWI}) + \frac{U_i - \min(\mathbf{U})}{\text{Range}(\mathbf{U})} \times \text{Range}(\text{EWI}), \quad (7)$$

so that CPI will have the same range as the EWI.

4. Rank countries according to CPI.

Figure 2 shows the strong linear relationship between PCI and EWI.

3 Clustering Multivariate Composite Indices Data

Instead of computing a score and a rank of each country, we can group countries into K groups such that the countries in one group are similar to each other but different from the countries in other groups. Clustering methods can be used to achieve this objective. There are numerous methods for clustering in the literature. Here we consider two of the commonly used ones in practice: the K-Means Clustering and the Hierarchical Clustering.

3.0.1 K-Means Clustering

One of the most commonly used clustering algorithm is the K-Means algorithm. It can be described briefly as follows (Johnson and Wichern, 1992):

1. Choose the number of clusters, K .
2. Generate K random points and use them as cluster centers.
3. Assign each point to the nearest cluster center.
4. Compute the new cluster centers.
5. Repeat Steps 3 and 4 until some convergence criterion is met (usually that the assignment hasn't changed).

The K -Means algorithm is popular due to its simplicity and speed (it is suitable for large data sets). But it has serious disadvantage. One needs to specify the desired number of clusters K . As a rule of thumb K is approximately $\sqrt{n/2}$. More seriously, from Step 2, the output is random, that is, it does not yield the same results with each run. Also, it tries to minimize intra-cluster variance, but does not ensure a global minimum. For these reasons we will consider here the hierarchical clustering algorithm.

3.0.2 Hierarchical Clustering

There are several algorithms that fall under the class of hierarchical clustering algorithms. One commonly used hierarchical clustering algorithms is the Agglomerative method, which is briefly summarized as follows (Rousseeuw, 1987):

1. Compute an appropriate distance (e.g., Euclidean distance) between every pair of observations.
2. Assign each of the n observations to its own cluster.
3. Compute an appropriate distance (e.g., using Ward Method or Complete Linkage distance) between every pair of clusters and merge the two closest clusters into one cluster.
4. Repeat Step 3 until all n observations belong to one cluster.
5. Draw the Dendrogram (cluster tree) and from the tree form the K clusters.

Using the Agglomerative Hierarchical Clustering method on the 2016 African Governance data without grouping, we obtain the dendrogram shown in Figure 3. From the dendrogram, one can conclude that the appropriate number of clusters (groups) is 5. The five clusters (groups) are shown within the five rectangular boxes. The countries in each of the 5 groups are given in Table 3.

Figure 4 shows the scatter plot of the Hierarchical Cluster grouping versus the IIAG rankings of the 54 countries in 2016. The plot shows some similarities but also some substantially different rankings of the countries. The top seven countries according to the IIAG are all in the Group 1 (the highest ranked group). Similarly, the four lowest ranked countries according to the IIAG rankings are all in the lowest ranked Group 5. However, Sudan (SD) with rank 49 was placed in Group 5 whereas Eritrea (ER) with rank 50 was placed in Group 4. Also, Egypt (EG) with rank 24 was placed in Group 2 surpassing many countries with better IIAG rankings. Perhaps the most surprising result is that Sao Tome and Principe (ST), which is the eleventh top ranked country by IIAG, is placed in the second lowest Group 4.

Figure 5 compares the Equal Weight rankings (see Table 2) with the Hierarchical Cluster rankings. We see here also that there are some similarities and some differences. The six top ranked countries are placed in the highest Hierarchical Cluster Group 1. Similarly, the five lowest ranked countries are placed in the lowest Hierarchical Cluster Group 5. But there are many differences. For example, Tunisia (TN), which is ranked 8th, is placed in Group 1 ahead of Rowanda (RW), which is ranked 7th.

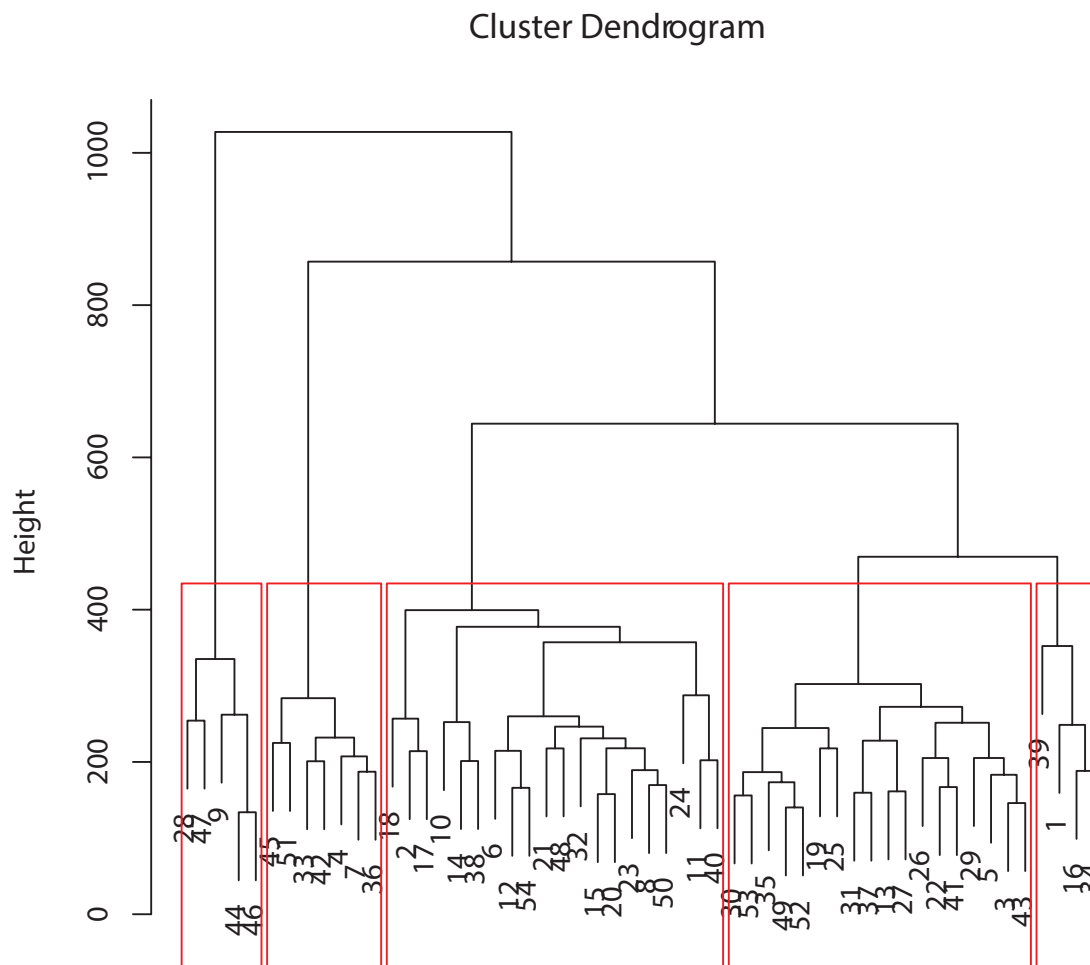


Figure 3: Dendrogram for the 2016 African data without grouping.

Table 3: Results of Hierarchical Clustering for the 54 Countries in 2016.

Country	IIAG Ranks	HC Groups	Country	IIAG Ranks	HC Groups
Botswana	2	1	Uganda	19	3
Cabo Verde	3	1	Zambia	13	3
Mauritius	1	1	Angola	45	4
Namibia	5	1	Burundi	43	4
Seychelles	4	1	Cameroon	38	4
South Africa	6	1	Chad	48	4
Tunisia	7	1	Comoros	26	4
Algeria	20	2	Congo	42	4
Egypt	24	2	D. R. of Congo	46	4
Morocco	14	2	Djibouti	37	4
Rwanda	9	2	Equatorial Guinea	47	4
Benin	16	3	Eritrea	50	4
Burkina Faso	23	3	Gabon	32	4
Cote d'Ivoire	21	3	Gambia	35	4
Ethiopia	31	3	Guinea	41	4
Ghana	8	3	Guinea-Bissau	44	4
Kenya	12	3	Mauritania	40	4
Lesotho	15	3	Nigeria	36	4
Liberia	28	3	Sao Tome & Principe	11	4
Madagascar	34	3	Swaziland	29	4
Malawi	17	3	Togo	33	4
Mali	25	3	Zimbabwe	39	4
Mozambique	22	3	Cent. African Republic	52	5
Niger	27	3	Libya	51	5
Senegal	10	3	Somalia	54	5
Sierra Leone	30	3	South Sudan	53	5
Tanzania	18	3	Sudan	49	5

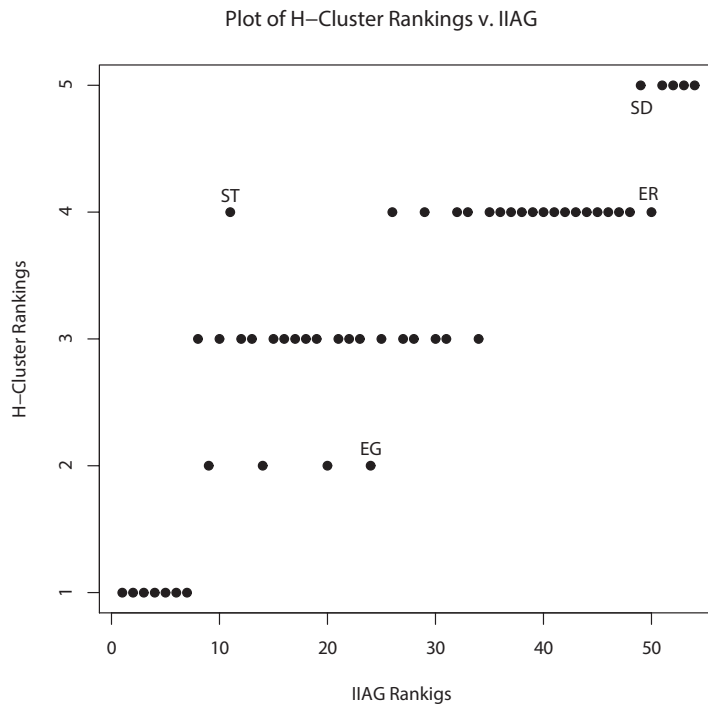


Figure 4: Plot of Hierarchical Cluster Rankings v. IIAG Rankings.

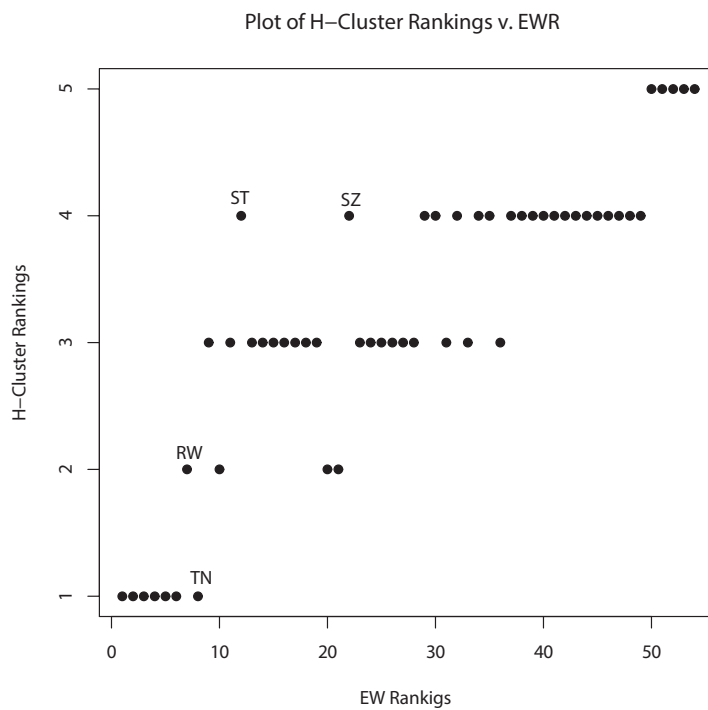


Figure 5: Plot of H-Cluster Rankings v. Equal Weight Rankings.

4 Summary and Conclusions

The contributions of this paper can be summarized as follows:

1. As an alternative methodology to computing composite indices, we propose the use of clustering methods to group objects or countries into K groups such that the countries in one group are similar to each other but different from the countries in the other groups. The use of clustering avoids two issues with composite indices, that is, the need to group variables/indicators by experts and the need to compute the standard errors. Countries in the same group are taken to be similar (having approximately the same rank) and countries that are classified into different groups are interpreted as having different ranks.
2. Different ways of ranking the countries are compared: The IIAG rankings, the Equal Weight rankings, the Principal Components Index rankings, and the Hierarchical Cluster rankings. These methods provide different rankings. Experts in the relevant domains can use these results and decide on which method gives appropriate rankings based on their theoretical understanding of the domain of interest (e.g., African governance for the African governance data).

References

1. Abd ElAziz, T. E. R. (2017), “Generalizing Almost First Order Dominance to the Multidimensional Case with Application on Poverty in Egypt,” Ph.D. Thesis, Faculty of Economics and Political Science, Cairo University.
2. Abo Elazm, H. F., (2017), “Measuring Happiness Among Youth (18–29 Years) in Egypt in 2009,” M. Sc. Thesis, Faculty of Economics and Political Science, Cairo University.
3. Bandura R. (2008), “A Survey of Composite Indices Measuring Country Performance,” United Nations Development Program, Office of Development Studies.
4. Billor, N., Hadi, A. S., and Velleman, P. F. (2000), “Blocked Adaptive, Computationally-Efficient outlier Numerator,” *Computational statistics and data analysis*, 34, 279–298.

5. Cronbach, L. J. (1951), “Coefficient alpha and the internal structure of tests,” *Psychometrika*, 16, 297–333.
6. Decanco, K., and Lugo, A. M. (2013) “Weights in multidimensional indices of well-being (An overview),” *Econometric Reviews*, 32, 7–34.
7. Dobbie, M. J., and Dail, D. (2013), “Robustness and sensitivity of weighting and aggregation in constructing composite indices,” *Ecological Indicators*, 29, 270–277.
8. Donders, A. R. T., Van der Hijden, G. J. M. G., Stijen, T., and Moons, K. O. M., (2006), “ Review: A gentle introduction to imputation of missing values,” *Journal of Clinical Epidemiology*, 59, 1087–1091.
9. Efron, B. and Tibshirani, R. (1993), “An Introduction to Bootstrap,” Chapman and Hall: New York.
10. Johnson, R. A., and Wichern, D. W. (1992), “Applied Multivariate Statistical Analysis,” New Jersey: Prentice Hall.
11. Jolliffe, I. T., (2002), “Principle Components Analysis,” Springer: New York.
12. Organization for Economic Co-Operation and Development (OECD), (2008), “Handbook on Constructing Composite Indicators - Methodology and User Guide,” OECD.
13. Rabea, S. A. R. (2017), “Measuring Child Wellbeing in Egypt Using Micro-Data,” M.Sc. Thesis, Faculty of Economics and Political Science, Cairo University.
14. Rocke, D.M., and Woodruff, D. L., (1996), “Identification of Outliers in Multivariate Data,” *Journal of the American Statistical Association*, 91, 1047–1061.
15. Rousseeuw, P. J. (1987), “Silhouettes: A Graphical Aid to the Interpretation and validation of cluster analysis,” *Journal of Computational and Applied Mathematics*, 20, 53–65.

16. Rousseuw, P. J., and Van Driessse, K., (1999), “A Fast Algorithm for Minimum Covariance Determinant Estimator,” *Technometrics*, 41, 212–223.

List of Sites:

17. Al Maktoum Foundation <http://WWW.mbrf.ae>, Last access 5/12/2017.
18. http://old.usc.ac.ir/IPPWebV1C035/TemplateFileFolder/10-8-2013/OriginalFolder/0a6e29c4-1fa6-4496-a362-000f0851d4fc_indices_2008.bandura.pdf.
Last access 5/12/2017.
19. IIAG: <http://www.moibrahimfoundation.org>, Last access 5/12/2017.

Review: Conditional and Unconditional AR (1) Model

Sayed M. El-Sayed¹, Ahmed A. El-Sheikh²,
Mohamed K. A. Issa³ and Hadia F.M.A. Azmy⁴

Abstract

In this paper, a literature of conditional and unconditional autoregressive model AR (1) will be discussed and explained which estimator is the best for application.

Keywords: First order autoregressive, Ordinary least squares, Maximum likelihood, Weighted symmetric.

1. Introduction

There are some similar aspects between autoregressive model and linear regression model, where y_t is the dependent variable, while $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ are the independent variables, and ε_t is considered to be the error term in autoregressive model. On the other hand, there is a difference between the two models; where the independent variables in auto-regressive model are random and correlated and there is another difference that the initial value (y_0) in AR model, this problem summarized in the first term of error (ε_1), because the value of ε_1 depends on the initial value y_0 and it cannot be estimated from the available observations y_1, y_2, \dots, y_n , but if $n \rightarrow \infty$ the effect of ε_1 will be negligible and so it will tend to zero. So there are different methods of estimation applied with the autoregressive models when y_0 is fixed, this method of estimation is called conditional estimation and when y_0 is stochastic

¹Prof of Applied Statistics, ISSR, Cairo University.

²Prof of Applied Statistics, ISSR, Cairo University.

³Assist. Prof of Applied Statistics, the Higher Institute of Cooperative and Managerial Studies.

⁴PhD. Student, ISSR, Cairo University.

and $N\left(0, \frac{\sigma^2}{1-\rho^2}\right)$, this method of estimation is called unconditional estimation [Sharawy (2005)].

The review is divided into two parts: the first part introduces a literature review of conditional AR (1) model, the second part introduces a literature review of unconditional AR (1) model.

2. Conditional AR (1) Model

Dickey and Fuller (1979) introduced the first order autoregressive model AR(1) which is defined as

$$y_t = \rho y_{t-1} + \varepsilon_t \quad (1)$$

Where $t = 1, 2, \dots, n$, y_0 is fixed constant, and $\varepsilon_t \sim iid N(0, \sigma^2)$. They investigated the properties of the estimator ρ using the ordinary least squares method (OLS) under the assumption that $\rho = \pm 1$. They derived the limit distribution of the estimator of ρ . The results of Monte Carlo simulation indicated that tests based on the estimated ρ were more powerful for tests against stationarity than the \hat{t} statistics.

Fuller and Hasza (1979) investigated an estimator for AR (1) model

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t \quad (2)$$

Where $t = 1, 2, \dots, n$, y_0 is fixed constant, and $\varepsilon_t \sim iid N(0, \sigma^2)$. They studied alternative predictors for stationary and non-stationary models using the Monte Carlo method. They showed that the OLS predictor is unbiased for symmetric error distributions for AR (1) model.

Fujikoshi and Ochi (1984) introduced the AR (1) model which is defined in (1) Where $t = \dots, -1, 0, 1, \dots$ and ρ is an unknown parameter, $|\rho| < 1$ and $\varepsilon_t \sim iid N(0, \sigma^2)$. They derived an asymptotic expansion of the distribution of

Maximum likelihood (ML) estimator based on T observations for the AR (1) normal model up to the term of order T^{-1} .

Dickey, et al. (1984) proposed the simple weighted symmetric (SWS) estimator for AR (1) model without constant as defined in (1). Where $t=1,2,\dots,n$ and y_0 is fixed constant, $\varepsilon_t \sim iid N(0, \sigma^2)$. The SWS estimator of ρ is given by

$$\hat{\rho}_{SWS} = - \left[0.5(y_1^2 + y_n^2) + \sum_{t=2}^{n-1} y_t^2 \right]^{-1} \sum_{t=2}^n y_t y_{t-1}$$

Breton and Pham (1989) introduced an estimator for the AR (1) model without constant as defined in (1). Where $t \geq 1$, $y_0 = 0$, ρ is a real parameter and $\varepsilon_t \sim iid N(0, \sigma^2)$. They obtained the exact formula for the bias of estimator and derived the asymptotic bias using OLS and ML.

Park and Fuller (1995) introduced a weighted symmetric (WS) estimator for AR(1) model without constant as defined in (1). Where $t = 2, 3, \dots, n$, $\varepsilon_t \sim iid N(0, \sigma^2)$ and y_0 is fixed constant. Under the stationary condition, that is, $|\rho| = 1$ could have a backward representation of the autoregressive time series. That is,

$$y_t = \rho y_{t+1} + \varepsilon_t \quad t = 2, 3, \dots, n \quad (3)$$

The WS estimator of ρ denoted by $\hat{\rho}_{WS}$, is derived by minimizing the following function

$$Q_{WS}(\rho) = \sum_{t=2}^n w_t (y_t - \rho y_{t-1})^2 + \sum_{t=1}^{n-1} (1 - w_{t+1}) (y_t - \rho y_{t+1})^2$$

The WS estimator for AR (1) as follows:

$$\hat{\rho}_{WS} = \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^{n-1} y_t^2 + n^{-1} \sum_{t=1}^n y_t^2}$$

Shenton and Vinod (1995) conducted closed forms for asymptotic bias and variance for AR (1) model without constant as defined in (1) with unit roots and with

zero initial value. They provided the closed- form for the asymptotic bias and variance of ML estimator as follows;

$$\hat{\rho} = \frac{\sum_1^n y_t y_{t-1}}{\sum_1^{n-1} y_t^2}$$

Where $\varepsilon_t \sim iid N(0, \sigma^2)$ and $|\rho| = 1$. Numerical accuracy up to six significant digits was achieved for sample sizes when $n > 100$.

Ali (2002) investigated the finite sample distribution of the OLS estimator of the autoregressive parameter in AR (1) model without constant as defined in (1), obtained a uniform asymptotic expansion for the distribution applicable to both stationary and non-stationary cases, used three cases of initial values y_0 , where $y_0 = 0$ is fixed, $y_0 \neq 0$ is constant and y_0 is stochastic.

Provost and Sanjel (2005) investigated the estimator of the coefficient of AR (1) model as defined in (1). Where $t = 0, \pm 1, \pm 2, \dots, \pm n$, $\varepsilon_t \sim iid N(0, \sigma^2)$ and $|\rho| = 1$. They compared estimator of the parameter of AR (1) model using the OLS, modified least squares (MLS), the Yule-Walker, and Burg's estimators. They showed that the MLS estimator is the least biased and the OLS and Burg's estimators have similar means and variances for any $n > 10$.

Youssef (2008) introduced an estimator for the AR (1) model with unit root as defined in (2). Where $t = 2, 3 \dots, n$, $\varepsilon_t \sim iid N(0, \sigma^2)$. He compared the OLS, MLS, WS, modified weighted symmetric (MWS), and ML estimators for AR (1) in the case of unit root and he studied it using Monte Carlo method, to see how well the estimators and the predictors performed on different samples sizes. The Monte Carlo simulation results suggested that the MLS estimator performed well, in the sense of less biased and MSE, in small and large sample size. Therefore, he recommend to use MLS to estimate the unit roots.

Also Youssef, et al (2012) introduced a comparison between some methods of estimation, such as (OLS), (MLS), (WS), (MWS) and unconditional maximum likelihood (UML) for AR (1) model as defined in (2). Where $t = 2, 3 \dots, n$,

$\varepsilon_t \sim iid N(0, \sigma^2)$. Also, the properties of OLS estimator were discussed. A comparative study between these methods was considered in the case of unit root and explosive time series, using Monte Carlo simulation. They concluded that the estimator of MLS method was the best one among other estimators for small and medium sample sizes, while UML was the best one for large sample size.

El-Sayed et al (2015) proved the linearization and the unbiasedness of the estimated parameter of AR (1) model without constant as defined in (1) and (3) using WS method, where $t = 2, 3, \dots, n$, $\varepsilon_t \sim iid N(0, \sigma^2)$, and $|\rho| = 1$.

3. Unconditional AR(1) Model

Box and Jenkins (1976) investigated ML estimator of the parameter for AR(1) model as defined in (1). They showed that there is no direct analytical formula in terms of series data when derive the logarithmic of unconditional maximum likelihood function (4), so the following function has to be solving by one of numerical methods.

$$\ln L = -\frac{n}{2} \ln(2\pi\sigma^2) + \frac{1}{2} \ln(1 - \rho^2) - \frac{1}{2\sigma^2} \left[y_1^2(1 - \rho^2) + \sum_{t=2}^n (y_t - \rho y_{t-1})^2 \right] \quad (4)$$

They showed that the fundamental reason in the absence of such formula due to the presence of quantity $\ln(1 - \rho^2)$ in the logarithmic function, so assumed that the effect of $\ln(1 - \rho^2)$ will be negligible and so it will tend to zero if $n \rightarrow \infty$. Then the logarithm of likelihood function can be written in approximate way as the follow.

$$\ln(L) \approx -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \left[y_1^2(1 - \rho^2) + \sum_{t=2}^n (y_t - \rho y_{t-1})^2 \right]$$

And they obtain the UML estimator as follows;

$$\hat{\rho}_{UML} = \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2}$$

They concluded that the estimator $\hat{\rho}_{UML}$ is the same as the unconditional ordinary least squares (UOLS) estimator $\hat{\rho}_{UOLS}$.

Hasza (1980) derived approximate estimator for AR (1) for $\rho \in (-\infty, \infty)$ and showed that there is a root in each of interval $(-\infty, -1)$, $(-1, 1)$, and $(1, \infty)$ by using ML. If y_t is stationary, then the estimator for AR(1) for ρ is a root of the following cubic equation

$$f(\rho) = a_3\rho^3 + a_2\rho^2 + a_1\rho + a_0$$

Where

$$a_3 = \left[\frac{n-1}{n} \right] \sum_{t=2}^{n-1} y_t^2, \quad a_2 = - \left[\frac{n-2}{n} \right] \sum_{t=2}^n y_t y_{t-1}, \quad a_1 = \sum_{t=2}^{n-1} y_t^2 - \frac{1}{n} \sum_{t=1}^n y_t^2, \quad a_0 = \sum_{t=2}^n y_t y_{t-1}$$

Shin and Fuller (1998) proposed a Likelihood ratio test based on UML estimation for autoregressive moving average (ARMA) that may contained an AR unit root, and this was considered as an extension of the work of Hasza (1980). They showed that the limiting distribution of normalized ML estimator of the unit root was the same as the estimator for AR (1) model.

El-Sayed, et al. (2017) investigated a closed form of the bias of the parameter using UML of AR(1) model as defined in (1) Where $t = 1, 2, 3, \dots, n$, $\varepsilon_t \sim iidN(0, \sigma^2)$, assumed that y_0 is $N(0, \frac{\sigma^2}{1-\rho^2})$ under the stationarity condition $|\rho| < 1$, which is considered as an extension of Box and Jenkins (1976). Moreover, the consistency of the estimated parameter has been proved.

4. Conclusion

Based on the previous review, the following points can be presented, as;

- For AR (1) model with unit root problem, the MLS method was the best one among other estimators for small and medium sample sizes, while UML method was the best one for large sample size [Youssef (2008)] and [Youssef, et al (2012)].
- For AR (1) model with explosive problem, the UML method which considered the only method of estimation that can be used in this case [Youssef, et al (2012)].
- For stationary AR (1) model a closed form of the bias using UML as the method estimation has been derived [El-Sayed, et al. (2017)].

References

1. Ali, M.M. (2002). Distribution of The Least Squares Estimation in a First Order Autoregressive Model. *Econometric Reviews*, 21(1), 89-119.
2. Box, G. E. P. and Jenkins, G. M. (1976). *Time series analysis forecasting and control*. San Francisco, Holden Day
3. Breton, A. L. and Pham, D. T. (1989). On the bias of the least squares estimator for the first order autoregressive process. *Ann. Inst. Statist. Math.* 41(3), 555-563.
4. Dickey, D.A. and Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *JASA*, 74(366), 427-431.
5. Dickey, D. A., Haszaand, D. P. and Fuller, W. A. (1984). Testing for unit roots in seasonal time series, *JASA*, 79, 355-367.
6. El-Sayed, S.M., El-Sheikh, A. A. and Issa, M. K. A. (2015). Weighted Symmetric Estimators of Autoregressive Models. *Far East Journal of Advances and Applications in Statistics*, 47(2), 145-152.
7. El-Sayed, S.M., El-Sheikh, A. A., Issa, M. K. A. and Azmy, H.F.M.A (2017). A Closed Form of Biased AR (1) Model. *Far East Journal of Advances and Applications in Statistics*, 50(3), 191-199.

8. Fujikoshi, Y. and Ochi, Y. (1984). Asymptotic Properties of the Maximum Likelihood Estimate in the First Order Autoregressive Process. *Ann. Inst. Statist. Math. P.A*, 119-128.
9. Fuller, W.A. and Hasza, D.P. (1979). Predictors for the First-Order Autoregressive Process. *Journal of Econometrics*, 13, 139-157.
10. Hasza, D.P. (1980). A Note on Maximum Likelihood Estimation for the First Order Autoregressive Process. *Communication in Statistics, Theory and Methods*, 9 (13), 1411-1415.
11. Park, H. and Fuller, W.A. (1995). Alternative Estimators and Unit Root Tests for the Autoregressive Process. *Journal of Time Series Analysis*. 16, 415-429.
12. Provost, S.B. and Sanjel, S. (2005). Inference About the First Order Autoregressive Coefficient. *Communication in Statistics-Theory and Method*, 34(5), 1183-1201.
13. Sharawy, S.M. (2005). *An Introduction of Modern Time Series Analysis*. Scientific Research Centre, Faculty of Science, King Abd-Elaziz University, Jeddah, Kingdom of Saudi Arabia.
14. Shenton. L.R. and Vinod, H.D. (1995). Closed Forms for Asymptotic Bias and Variance in Autoregressive Models with Unit Roots, *Journal of Computational and Applied Mathematics*. 61, 231-243.
15. Shin, D.W. and Fuller, W.A. (1998). A unit Root Tests Based on Unconditional Maximum Likelihood Estimation for the Autoregressive and Moving Average. *Journal of Time Series Analysis*, 19(5), 591-599.
16. Youssef, A. H. (2008). A New Estimator for the Unit Root. *Journal of Statistical Computation and Simulation*, 78, 515 – 526.
17. Youssef, A. H., El-Sheikh, A. A and Issa, M. K. A. (2012). A Comparison of Estimators for the First Order Autoregressive Process with a Unit Root. *Far East Journal of Theoretical Statistics*, 41(2), 133-147.

Estimation of Ranked Set Sampling Methods : Comparative Study

Nasr I. Rashwan¹ and Nuran M. Al-Mawan^{2,3}

Abstract :

In environmental monitoring and assessment, the main focus is to achieve observational economy and to collect data with unbiased, efficient and cost effective sampling methods. Ranked set sampling (RSS) is one traditional method that is mostly used for accomplishing observational economy. In this article, quartile double ranked set sampling (QDRSS) is compared to different ranked set sampling methods which are useful to estimate the population mean based on simple random sampling (SRS) method. When the underlying distributions are symmetric and asymmetric, it turns out that quartile double RSS (QDRSS) produce unbiased estimator is more efficient than SRS, paired RSS (PPRSS), extreme RSS (ERSS), median RSS (MRSS), double RSS (DRSS), quartile RSS (QRSS), percentile RSS (PRSS), paired double RSS (PPDRSS) and RSS based on the sample size. It is shown that, the variance of the mean estimator under QDRSS is always less than the variance of the mean estimator based on SRS and the other methods.

Keywords:

Ranked set sampling; Paired ranked set sampling; Extreme ranked set sampling; Median ranked set sampling; Double ranked set sampling; Quartile ranked set sampling; Percentile ranked set sampling; Quartile double ranked set sampling; Paired double ranked set sampling; Relative precision.

1. Introduction

There are many sampling methods that can be used in surveys of natural resources in agriculture, biology, environmental management, forestry, etc. The main objective of whatever sampling method is used is to obtain precise estimates of population parameters with minimum cost and expenditure. One method is ranked set sampling (RSS). RSS becomes an efficient alternative to simple random sampling (SRS) when taking exact measurements of selected units is very costly whereas ranking a small set of selected units is cheap. Ranking may be visually with respect to the study variable or by any inexpensive method. McIntyre^[10] was the first to suggest the RSS method for estimation of pasture and forage yields. McIntyre's goal was to maintain the unbiasedness of SRS while effectively incorporating into the estimates the outside information provided by visual inspection. Takahasi and Wakimoto^[21] developed the theory of RSS procedure under the assumption of perfect ranking. Dell and Clutter^[5] showed that the mean estimator under imperfect RSS remains an unbiased estimator

Department of Statistics & Mathematics , Faculty of Commerce , Tanta University , Tanta , Egypt
Department of Applied Statistics and Econometric , Institute of Statistical Studies and Research ,Cairo University ,
Cairo , Egypt

Department of Basic Science , Faculty of Engineering ,Modern Academy , Cairo , Egypt

of population mean. Patil et al. ^[16] compared the mean estimator based on perfect and imperfect rankings with the SRS based regression estimator of population mean. Muttlak ^[11] and Muttlak ^[12] introduced paired ranked set sampling (PPRSS) and median ranked set sampling (MRSS) schemes for estimation of population mean, respectively. PPRS can be used when it is difficult to select a large number of units from the population of interest. Samawi et al. ^[19] suggested using extreme ranked set sampling (ERSS) to estimate the population mean if the underlying distribution is symmetric and it is more efficient than the SRS estimator. Muttlak ^[13] extended the work of Patil et al. ^[16] and showed that the mean estimator under MRSS is more efficient than the mean estimators with SRS and RSS. Al-Saleh and Al-Kadiri^[1] suggested double ranked set sampling method (DRSS) for estimating the population mean. They showed that the sample mean based on DRSS is more efficient than the sample mean with RSS. Muttlak ^[14,15] proposed quartile ranked set sampling (QRSS) and percentile ranked set sampling (PRSS) for estimating the population mean. Al-Saleh and Al-Omari ^[2] introduced quartile double ranked set sampling (QDRSS) procedure for estimating the population mean. Haq et al. ^[7] suggested paired double ranked set sampling (PPDRSS) for estimating the population mean. The main advantage of using PPDRSS over DRSS is that it requires less number of identified units as compared with DRSS. This helps in reducing the time and cost that is involved in the ranking process.

In this paper, quartile double ranked set sampling (QDRSS) method is compared with the other ranked set sampling methods. It is shown that for the probability distributions considered in this study, QDRSS unbiased estimator has less variance than all the other methods. Computer simulation results are given to compare the efficiency for the estimators based on SRS with its counterparts RSS, PPRS, ERSS, MRSS, QRSS, PRSS, DRSS, QDRSS and PPDRSS.

2. Ranked set sampling methods

This section is concerned with methods of ranked set sampling given a definition of each method and defined the mean and the variance of each case in the methods.

Let X_1, X_2, \dots, X_n be a random sample with probability density function $f(x)$ with mean μ and variance σ^2 . Let $X_{1(1:n)}, X_{2(1:n)}, \dots, X_{n(1:n)}$; $X_{1(2:n)}, X_{2(2:n)}, \dots, X_{n(2:n)}$; ... ; $X_{1(n:n)}, X_{2(n:n)}, \dots, X_{n(n:n)}$ be independent random variables all with the same cumulative distribution function $F(x)$. Let $X_{i(i:n)}$ denotes the i^{th} order statistic from the i^{th} sample of size n ($i=1, 2, \dots, n$). The cycle may be repeated m times to get mn units. Let $X_{(i:n)j}$ denotes the j^{th} cycle of size m ($j=1, 2, \dots, m$). To understand the calculation in the study assume the cycle is repeated once ; i.e. $m=1$. The variance or the mean square error of the sample means for the RSS, PPRS, ERSS, MRSS, QRSS, PRSS, DRSS, QDRSS and PPDRSS were calculated using the moments of the order statistics for the distributions considered in the study (see Harter and Balakrishnan ^[8] and

Balakrishnan and Chen ^[3]). The relative efficiency (RE) of estimating the population mean using any of the RSS based methods with respect to the usual estimator using SRS is defined as ^[14]

$$RE(\bar{X}_{SRS}, \bar{X}_{RSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{RSS})}$$

Here \bar{X}_{SRS} and \bar{X}_{RSS} are both unbiased estimators. In case of biased estimator, the variance is replaced by the *MSE* of the estimator where ^[17]

$$MSE(\bar{X}_{RSS}) = Var(\bar{X}_{RSS}) + [E(\bar{X}_{RSS}) - \mu]^2.$$

2.1. Ranked set sampling

The original idea of ranked set sampling (RSS) was introduced by McIntyre's ^[10]. Ranked set sampling is a methodology which can improve the efficiency of techniques such as estimation and confidence intervals without increasing the number of substantial observations. It uses additional information from the population to provide more structure to the data collection process and decreases the likelihood of an unrepresentative sample. In addition, it is designed to minimize the number of measured observations required to achieve the desired precision in making inferences. The RSS scheme can be described as follows : In order to draw a sample of size n , we identify n^2 units from the target population. The n units in each sample are ranked visually or by any inexpensive method with respect to the variable of interest. From the first set of n units, the smallest ranked unit is measured. From the second set of n units, the second smallest ranked unit is measured. The process is continued until from the n th set of n units the largest ranked unit is measured. The estimator of the population mean for RSS of size n is given by

$$\bar{X}_{RSS} = \frac{1}{n} \sum_{i=1}^n X_{(i:n)},$$

with variance

$$Var(\bar{X}_{RSS}) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2_{(i:n)}, \text{ where } \sigma^2 = E(X_{(i:n)} - \mu_{(i:n)})^2 = \sigma^2 - \sum_{i=1}^n (\mu_{(i:n)} - \mu)^2.$$

The RE for the symmetric and asymmetric distributions are

$$RE(\bar{X}_{SRS}, \bar{X}_{RSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{RSS})}.$$

2.2. Paired ranked set sampling

Muttlak ^[11] introduced paired ranked set sampling (PPRSS) procedure for estimating the population mean. PPRSS procedure is explained as follows: for even sample size n , identify n^2 sets each of size n from the target population. Each sample is ranked in itself as in ranked set sampling design. Select the smallest and largest ranked units from the first set. Similarly, select the second smallest and second largest ranked units from the second set. The procedure continues, then $\frac{n}{2}$ th and $\frac{n+2}{2}$ th ranked units are selected from the last set. Similarly, for odd sample size n , identify $\frac{n+1}{2}$ sets each of size n from the target population. Each sample is ranked in itself as in ranked set sampling design. Select the smallest and largest ranked units

from the first set. Similarly, select the second smallest and second largest ranked units from the second set. The procedure continues, then the $\frac{n+1}{2}$ th ranked units are selected from the last set. The sample means under PPRSS for even and odd n are given below:^[11]

$$\bar{X}_{PPRSSE} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n}{2}} X_{i:(n)} + \sum_{i=1}^{\frac{n}{2}} X_{i(n-i+1:n)} \right)$$

$$\text{and } \bar{X}_{PPRSSO} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n+1}{2}} X_{i:(n)} + \sum_{i=1}^{\frac{n-1}{2}} X_{i(n-i+1:n)} \right),$$

Both \bar{X}_{PPRSSE} and \bar{X}_{PPRSSO} are unbiased estimators. The variances of $Var(\bar{X}_{PPRSSE})$ and $Var(\bar{X}_{PPRSSO})$ respectively, are

$$Var(\bar{X}_{PPRSSE}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^2_{(i:n)} + \sum_{i=1}^{\frac{n}{2}} \sigma^2_{(n-i+1:n)} \right),$$

$$Var(\bar{X}_{PPRSSO}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n+1}{2}} \sigma^2_{(i:n)} + \sum_{i=1}^{\frac{n-1}{2}} \sigma^2_{(n-i+1:n)} \right).$$

The RE for the symmetric distribution is

$$RE(\bar{X}_{SRS}, \bar{X}_{PPRSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{PPRSS})},$$

$$\text{and } RE(\bar{X}_{SRS}, \bar{X}_{PPRSS}) = \frac{Var(\bar{X}_{SRS})}{MSE(\bar{X}_{PPRSS})} \text{ for asymmetric distribution.}$$

2.3. Extreme ranked set sampling

Samawi et al. ^[19] used extreme ranked set sample (ERSS) in case of even sample size with is easier to use than the usual RSS procedure to estimate the population mean. ERSS procedure is explained as follows: for even sample size n , identify n^2 units from the target population. Each sample is ranked in itself as in ranked set sampling design. Select the smallest ranked unit $X_{(1)}$ from first $\frac{n}{2}$ sets and the largest ranked unit $X_{(n)}$ from the other $\frac{n}{2}$ sets. Similarly, for odd sample size n , select the smallest ranked unit $X_{(1)}$ from first $\frac{n-1}{2}$ sets. Select the largest ranked unit $X_{(n)}$ from second $\frac{n-1}{2}$ sets and the median with rank $\frac{n+1}{2}$ from the remaining set. The sample means under ERSS depending on even (E) and odd (O) sample sizes are, respectively, given by ^[19]

$$\bar{X}_{ERSSE} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n}{2}} X_{i(1:n)} + \sum_{i=\frac{n}{2}+1}^n X_{i(n:n)} \right)$$

$$\text{and } \bar{X}_{ERSO} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n-1}{2}} X_{i(1:n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} X_{i(n:n)} + X_{i(\frac{n+1}{2}:n)} \right),$$

Both \bar{X}_{ERSSE} and \bar{X}_{ERSO} are unbiased estimators. The variances of $Var(\bar{X}_{ERSSE})$ and $Var(\bar{X}_{ERSO})$ respectively, are

$$Var(\bar{X}_{ERSSE}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^2_{(1:n)} + \sum_{i=\frac{n}{2}+1}^n \sigma^2_{(n:n)} \right)$$

$$Var(\bar{X}_{ERSO}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n-1}{2}} \sigma^2_{(1:n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} \sigma^2_{(n:n)} + \sigma^2_{(\frac{n+1}{2}:n)} \right), \quad (2.1)$$

The RE for the symmetric distribution is

$$RE(\bar{X}_{SRS}, \bar{X}_{ERSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{ERSS})},$$

$$\text{and } RE(\bar{X}_{SRS}, \bar{X}_{ERSS}) = \frac{Var(\bar{X}_{SRS})}{MSE(\bar{X}_{ERSS})} \text{ for asymmetric distribution.}$$

2.4. Median ranked set sampling

Muttalak ^[12] introduced Median ranked set sampling (MRSS) procedure for estimating the population mean. MRSS procedure is explained as follows: for even sample size n , identify n^2 units from the target population. Each sample is ranked in itself as in ranked set sampling design. Select the $\frac{n}{2}th$ smallest rank from the first $\frac{n}{2}$ sets and select the $\frac{n+2}{2}th$ smallest rank from the other $\frac{n}{2}$ sets. Similarly, for odd sample size n , select the $\frac{n+1}{2}th$ smallest rank from all sets. The sample means under MRSS for even and odd n are given below:^[7]

$$\bar{X}_{MRSSe} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n}{2}} X_{i(\frac{n}{2};n)} + \sum_{i=\frac{n}{2}+1}^n X_{i(\frac{n+2}{2};n)} \right)$$

$$\text{and } \bar{X}_{MRSSo} = \frac{1}{n} \left(\sum_{i=1}^n X_{i(\frac{n+1}{2};n)} \right),$$

Both \bar{X}_{MRSSe} and \bar{X}_{MRSSo} are unbiased estimators. The variances of $Var(\bar{X}_{MRSSe})$ and $Var(\bar{X}_{MRSSo})$ respectively, are

$$Var(\bar{X}_{MRSSe}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^2_{(\frac{n}{2};n)} + \sum_{i=\frac{n}{2}+1}^n \sigma^2_{(\frac{n+2}{2};n)} \right),$$

$$Var(\bar{X}_{MRSSo}) = \frac{1}{n^2} \left(\sum_{i=1}^n \sigma^2_{(\frac{n+1}{2};n)} \right) \tag{2.2}$$

The RE for the symmetric distribution is

$$RE(\bar{X}_{SRS}, \bar{X}_{MRSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{MRSS})},$$

$$\text{and } RE(\bar{X}_{SRS}, \bar{X}_{MRSS}) = \frac{Var(\bar{X}_{SRS})}{MSE(\bar{X}_{MRSS})} \text{ for asymmetric distribution.}$$

2.5. Double ranked set sampling

Al-Saleh and Al-Kadiri ^[11] introduced double ranked set sampling (DRSS) procedure for estimating the population mean. In order to select a double ranked set sample of size n , the experimenter needs to identify n^3 units from the target population. This may be difficult when an epidemic breaks out in some area or in queuing problems when data arrive in batches of varying sizes. Moreover, there may be a shortage of experimental units or ranking is difficult, time consuming and costly, see Samawi ^[18] and Haq et al. ^[6]. In all such situations, the balanced DRSS cannot be conducted with full confidence or it is costly. The DRSS scheme can be described as follows: In order to draw a sample of size n , we identify n^3 units from the target population. Randomly allocate these units to n set each of size n^2 . Apply RSS procedure to n sets in order to obtain n ranked set samples each of size n . Again apply the RSS procedure to these n ranked sets, each of size n , to obtain a double ranked set sample of size n . The estimator of the population mean for DRSS of size n is given by ^[7]

$$\bar{X}_{DRSS} = \frac{1}{n} \sum_{i=1}^n X_{i}^*$$

with variance

$$Var(\bar{X}_{DRSS}) = \frac{1}{n^2} \sum_{i=1}^n \sigma^{*2}_{(i;n)}, \text{ where } \sigma^{*2}_{(i;n)} = E(X_{(i;n)} - \mu_{(i;n)})^2.$$

The RE for the symmetric distribution is

$$RE(\bar{X}_{SRS}, \bar{X}_{DRSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{DRSS})},$$

$$\text{and } RE(\bar{X}_{SRS}, \bar{X}_{DRSS}) = \frac{Var(\bar{X}_{SRS})}{MSE(\bar{X}_{DRSS})} \text{ for asymmetric distribution.}$$

2.6. Quartile ranked set sampling

Muttlak ^[14] introduced quartile ranked set sampling (QRSS) procedure for estimating the population mean. QRSS procedure is explained as follows: for even sample size n , identify n^2 units from the target population. Each sample is ranked in itself as in ranked set sampling design. Select the $(q_1(n+1))$ th smallest ranked unit from the first $\frac{n}{2}$ sets and select the $(q_3(n+1))$ th smallest ranked unit from the other $\frac{n}{2}$ sets where $q_1=0.25$ and $q_3=0.75$. Similarly, for odd sample size n , select the $(q_1(n+1))$ th smallest ranked unit from first $\frac{n-1}{2}$ sets. Select the $(q_3(n+1))$ th smallest ranked unit from second $\frac{n-1}{2}$ sets and the median with rank $\frac{n+1}{2}$ from the remaining set. The sample means under QRSS for even and odd n are given below:^[14]

$$\bar{X}_{QRSSSE} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n}{2}} X_{i(q_1(n+1):n)} + \sum_{i=\frac{n}{2}+1}^n X_{i(q_3(n+1):n)} \right)$$

$$\text{and } \bar{X}_{QRSSO} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n-1}{2}} X_{i(q_1(n+1):n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} X_{i(q_3(n+1):n)} + X_{i(\frac{n+1}{2}:n)} \right),$$

Both \bar{X}_{QRSSSE} and \bar{X}_{QRSSO} are unbiased estimators. The variances of $Var(\bar{X}_{QRSSSE})$ and $Var(\bar{X}_{QRSSO})$ respectively, are

$$Var(\bar{X}_{QRSSSE}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^2_{(q_1(n+1):n)} + \sum_{i=\frac{n}{2}+1}^n \sigma^2_{(q_3(n+1):n)} \right) \quad (2.3)$$

$$Var(\bar{X}_{QRSSO}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n-1}{2}} \sigma^2_{(q_1(n+1):n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} \sigma^2_{(q_3(n+1):n)} + \sigma^2_{(\frac{n+1}{2}:n)} \right), \quad (2.4)$$

The RE for the symmetric distribution is

$$RE(\bar{X}_{SRS}, \bar{X}_{QRSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{QRSS})},$$

$$\text{and } RE(\bar{X}_{SRS}, \bar{X}_{QRSS}) = \frac{Var(\bar{X}_{SRS})}{MSE(\bar{X}_{QRSS})} \text{ for asymmetric distribution.}$$

2.7. Percentile ranked set sampling

Muttlak ^[15] introduced percentile ranked set sampling (PRSS) procedure with different value of $0 \leq p \leq 1$ for estimating the population mean. PRSS procedure is explained as follows: for even sample size n , identify n^2 units from the target population. Each sample is ranked in itself as in ranked set sampling design. Select the $(p(n+1))$ th smallest ranked unit from the first $\frac{n}{2}$ sets and select the $(q(n+1))$ th smallest ranked unit from the other $\frac{n}{2}$ sets. Similarly, for odd sample size n , select the $(p(n+1))$ th smallest ranked unit from first $\frac{n-1}{2}$ sets. Select the $(q(n+1))$ th smallest ranked unit from second $\frac{n-1}{2}$ sets and the median with rank $\frac{n+1}{2}$ from the remaining set. The sample means under PRSS for even and odd n are given below:^[15]

$$\bar{X}_{PRSSSE} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n}{2}} X_{i(p(n+1):n)} + \sum_{i=\frac{n}{2}+1}^n X_{i(q(n+1):n)} \right)$$

$$\text{and } \bar{X}_{PRSSO} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n-1}{2}} X_{i(p(n+1):n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} X_{i(q(n+1):n)} + X_{i(\frac{n+1}{2}:n)} \right),$$

Both \bar{X}_{PRSSSE} and \bar{X}_{PRSSO} are unbiased estimators. The variances of $Var(\bar{X}_{PRSSSE})$ and $Var(\bar{X}_{PRSSO})$ respectively, are

$$Var(\bar{X}_{PRSE}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^2_{(p(n+1):n)} + \sum_{i=\frac{n}{2}+1}^n \sigma^2_{(q(n+1):n)} \right) \quad (2.5)$$

$$Var(\bar{X}_{PRSSO}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n-1}{2}} \sigma^2_{(p(n+1):n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} \sigma^2_{(q(n+1):n)} + \sigma^2_{(\frac{n+1}{2}:n)} \right), \quad (2.6)$$

The RE for the symmetric distribution is

$$RE(\bar{X}_{SRS}, \bar{X}_{PRSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{PRSS})},$$

and $RE(\bar{X}_{SRS}, \bar{X}_{PRSS}) = \frac{Var(\bar{X}_{SRS})}{MSE(\bar{X}_{PRSS})}$ for asymmetric distribution.

2.8. Quartile double ranked set sampling

Al-Saleh and Al- Omari ^[2] introduced quartile double ranked set sampling (QDRSS) procedure for estimating the population mean. The QDRSS scheme can be described as follows : In order to draw a sample of size n , we identify n^3 units from the target population. Randomly allocate these units to two sets each of size n^2 . Apply RSS procedure to n sets in order to obtain n ranked set samples each of size n . Apply the QRSS procedure to these n ranked sets, each of size n , to obtain a quartile double ranked set sample of size n . The sample means under QDRSS for even and odd n are given below:^[2]

$$\bar{X}_{QDRSSE} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n}{2}} X^*_{i(q_1(n+1):n)} + \sum_{i=\frac{n}{2}+1}^n X^*_{i(q_3(n+1):n)} \right)$$

$$\text{and } \bar{X}_{QDRSSO} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n-1}{2}} X^*_{i(q_1(n+1):n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} X^*_{i(q_3(n+1):n)} + X^*_{i(\frac{n+1}{2}:n)} \right),$$

Both \bar{X}_{QDRSSE} and \bar{X}_{QDRSSO} are unbiased estimators. The variances of $Var(\bar{X}_{QDRSSE})$ and $Var(\bar{X}_{QDRSSO})$ respectively, are

$$Var(\bar{X}_{QDRSSE}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^{*2}_{(q_1(n+1):n)} + \sum_{i=\frac{n}{2}+1}^n \sigma^{*2}_{(q_3(n+1):n)} \right)$$

$$Var(\bar{X}_{QDRSSO}) = \frac{1}{n} \left(\sum_{i=1}^{\frac{n-1}{2}} \sigma^{*2}_{(q_1(n+1):n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} \sigma^{*2}_{(q_3(n+1):n)} + \sigma^2_{(\frac{n+1}{2}:n)} \right),$$

The RE for the symmetric distribution is

$$RE(\bar{X}_{SRS}, \bar{X}_{QDRSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{QDRSS})},$$

and $RE(\bar{X}_{SRS}, \bar{X}_{QDRSS}) = \frac{Var(\bar{X}_{SRS})}{MSE(\bar{X}_{QDRSS})}$ for asymmetric distribution.

2.9. Paired double ranked set sampling

Haq et al. ^[7] introduced paired double ranked set sampling (PPDRSS) procedure for estimating the population mean. This method can be used as an alternative to DRSS when it's difficult or time consuming or costly to identify n^3 units from the target population, especially when there is a shortage of experimental units. The main steps involved in selecting PPDRSS of size n are as follows: for even sample size n , identify $\frac{n^3}{2}$ units from the target population. Each sample is ranked in itself as in ranked set sampling design. Randomly allocate these units to $\frac{n}{2}$ set of size n^2 . Apply RSS procedure on each set in order to obtain $\frac{n}{2}$ ranked set sample of size n . Now, apply PPRS procedure on $\frac{n}{2}$ set to get an even PPDRSS of size n . Similarly, for odd sample size n , identify $\frac{n^2(n+1)}{2}$ units from the target population. Randomly

allocate these units to $\frac{n+1}{2}$ sets each of size n^2 . Apply RSS procedure on each set in order to obtain $\frac{n+1}{2}$ ranked set samples each of size n . Now, apply PPRS procedure on $\frac{n+1}{2}$ sets to get an odd PPDRSS of size n . The sample means under PPDRSS for even and odd n are given below:^[7]

$$\bar{X}_{PPDRSSE} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n}{2}} X^*_{i(i:n)} + \sum_{i=1}^{\frac{n}{2}} X^*_{i(n-i+1:n)} \right)$$

$$\text{and } \bar{X}_{PPDRSSO} = \frac{1}{n} \left(\sum_{i=1}^{\frac{n+1}{2}} X^*_{i(i:n)} + \sum_{i=1}^{\frac{n-1}{2}} X^*_{i(n-i+1:n)} \right),$$

Both $\bar{X}_{PPDRSSE}$ and $\bar{X}_{PPDRSSO}$ are unbiased estimators. The variances of $Var(\bar{X}_{PPDRSSE})$ and $Var(\bar{X}_{PPDRSSO})$ respectively, are

$$Var(\bar{X}_{PPDRSSE}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^{*2}_{(i:n)} + \sum_{i=1}^{\frac{n}{2}} \sigma^{*2}_{(n-i+1:n)} \right),$$

$$Var(\bar{X}_{PPDRSSO}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n+1}{2}} \sigma^{*2}_{(i:n)} + \sum_{i=1}^{\frac{n-1}{2}} \sigma^{*2}_{(n-i+1:n)} \right).$$

The RE for the symmetric distribution is

$$RE(\bar{X}_{SRS}, \bar{X}_{PPDRSS}) = \frac{Var(\bar{X}_{SRS})}{Var(\bar{X}_{PPDRSS})},$$

and $RE(\bar{X}_{SRS}, \bar{X}_{PPDRSS}) = \frac{Var(\bar{X}_{SRS})}{MSE(\bar{X}_{PPDRSS})}$ for asymmetric distribution.

Lemma (i) $Var(\bar{X}_{MRSSO}) \leq Var(\bar{X}_{ERSO})$ for odd sample size (n).

(ii) $Var(\bar{X}_{MRSSO}) \leq Var(\bar{X}_{QRSSO})$ for odd sample size (n).

(iii) $Var(\bar{X}_{MRSSO}) \leq Var(\bar{X}_{PRSSO})$ for odd sample size (n).

(iv) In PRSS if the value of $p = 0.25$, then $Var(\bar{X}_{PRSS}) = Var(\bar{X}_{QRSS})$ for each sample size, whether even or odd.

Proof. To prove **(i)**, using (2.1) and (2.2)

$$Var(\bar{X}_{ERSO}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n-1}{2}} \sigma^2_{(1:n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} \sigma^2_{(n:n)} + \sigma^2_{(\frac{n+1}{2}:n)} \right)$$

$$= \frac{1}{n^2} \left(\left(\frac{n-1}{2} \right) \sigma^2_{(1)} + \left(\frac{n-1}{2} \right) \sigma^2_{(n)} + \sigma^2_{(\frac{n+1}{n})} \right)$$

$$= \frac{1}{n^2} \left(\frac{n-1}{2} \right) (\sigma^2_{(1)} + \sigma^2_{(n)}) + \frac{1}{n^2} \sigma^2_{(\frac{n+1}{n})}$$

$$= \frac{1}{n^2} \left(\frac{n-1}{2} \right) (\sigma^2_{(1)} + \sigma^2_{(n)}) + Var(\bar{X}_{MRSSO})$$

Remark that $\frac{1}{n^2} \left(\frac{n-1}{2} \right) (\sigma^2_{(1)} + \sigma^2_{(n)}) \geq 0$, then

$$Var(\bar{X}_{MRSSO}) \leq Var(\bar{X}_{ERSO}).$$

(ii) using (2.4) and (2.2),

$$Var(\bar{X}_{QRSSO}) = \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n-1}{2}} \sigma^2_{(q_1(n+1):n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} \sigma^2_{(q_3(n+1):n)} + \sigma^2_{(\frac{n+1}{2}:n)} \right)$$

$$= \frac{1}{n^2} \left(\left(\frac{n-1}{2} \right) \sigma^2_{(q_1(n+1))} + \left(\frac{n-1}{2} \right) \sigma^2_{(q_3(n+1))} + \sigma^2_{(\frac{n+1}{n})} \right)$$

$$= \frac{1}{n^2} \left(\frac{n-1}{2} \right) (\sigma^2_{(q_1(n+1))} + \sigma^2_{(q_3(n+1))}) + \frac{1}{n^2} \sigma^2_{(\frac{n+1}{n})}$$

$$= \frac{1}{n^2} \left(\frac{n-1}{2} \right) (\sigma^2_{(q_1(n+1))} + \sigma^2_{(q_3(n+1))}) + Var(\bar{X}_{MRSSO}) \tag{2.7}$$

Remark that $\frac{1}{n^2} \left(\frac{n-1}{2} \right) (\sigma^2_{(p(n+1))} + \sigma^2_{(q(n+1))}) \geq 0$, then

$$Var(\bar{X}_{MRSSO}) \leq Var(\bar{X}_{QRSSO}).$$

(iii) using (2.6) and (2.2),

$$\begin{aligned}
 Var(\bar{X}_{PRSSO}) &= \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n-1}{2}} \sigma^2_{(p(n+1):n)} + \sum_{i=\frac{n-1}{2}+1}^{n-1} \sigma^2_{(q(n+1):n)} + \sigma^2_{\left(\frac{n+1}{2}:n\right)} \right) \\
 &= \frac{1}{n^2} \left(\left(\frac{n-1}{2}\right) \sigma^2_{(p(n+1))} + \left(\frac{n-1}{2}\right) \sigma^2_{(q(n+1))} + \sigma^2_{\left(\frac{n+1}{n}\right)} \right) \\
 &= \frac{1}{n^2} \left(\frac{n-1}{2}\right) (\sigma^2_{(p(n+1))} + \sigma^2_{(q(n+1))}) + \frac{1}{n^2} \sigma^2_{\left(\frac{n+1}{n}\right)} \\
 &= \frac{1}{n^2} \left(\frac{n-1}{2}\right) (\sigma^2_{(p(n+1))} + \sigma^2_{(q(n+1))}) + Var(\bar{X}_{MRSSO}) \tag{2.8}
 \end{aligned}$$

Remark that $\frac{1}{n^2} \left(\frac{n-1}{2}\right) (\sigma^2_{(q_1(n+1))} + \sigma^2_{(q_3(n+1))}) \geq 0$, then

$$Var(\bar{X}_{MRSSO}) \leq Var(\bar{X}_{PRSSO}).$$

(iv) If $p = q_1 = 0.25$ and $q = q_3 = 0.75$ and the sample size (n) is odd,

from (2.7) and (2.8),

$$\begin{aligned}
 Var(\bar{X}_{QRSSO}) - \frac{1}{n^2} \left(\frac{n-1}{2}\right) (\sigma^2_{(q_1(n+1))} + \sigma^2_{(q_3(n+1))}) &= \\
 Var(\bar{X}_{PRSSO}) - \frac{1}{n^2} \left(\frac{n-1}{2}\right) (\sigma^2_{(p(n+1))} + \sigma^2_{(q(n+1))}) &
 \end{aligned}$$

then

$$Var(\bar{X}_{QRSSO}) = Var(\bar{X}_{PRSSO}),$$

If $p = q_1 = 0.25$ and $q = q_3 = 0.75$ and the sample size (n) is even,

from (2.3) and (2.5),

$$\begin{aligned}
 Var(\bar{X}_{QRSSO}) - \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^2_{(q_1(n+1):n)} + \sum_{i=\frac{n}{2}+1}^n \sigma^2_{(q_3(n+1):n)}\right) &= \\
 Var(\bar{X}_{PRSSO}) - \frac{1}{n^2} \left(\sum_{i=1}^{\frac{n}{2}} \sigma^2_{(p(n+1):n)} + \sum_{i=\frac{n}{2}+1}^n \sigma^2_{(q(n+1):n)}\right) &
 \end{aligned}$$

then

$$Var(\bar{X}_{QRSSO}) = Var(\bar{X}_{PRSSO}).$$

3. Simulation Study

Based on 30000 replications, a computer simulation is conducted to study the behavior of the efficiency of the sample mean using RSS, PPRSS, ERSS, MRSS, QRSS, PRSS with $p = 0.33$ and 0.8 , DRSS, QDRSS, PPDRSS with respect to SRS. The sample size n should be kept very small so that visual ranking errors will not destroy the gain in efficiency. Random observations are generated from (1) Standard normal distribution with ($\mu = 0, \sigma^2 = 1$) (2) Beta distribution with ($\alpha = 3, \beta = 3$) (3) Logistic distribution with ($\alpha = -1, \beta = 1$) (4) Exponential distribution with ($\lambda = 1$) (5) Gamma distribution with ($\alpha = 1, \beta = \frac{1}{2}$) (6) Weibull distribution with ($\lambda = 1, k = 2$), were generated by the (MATHEMATICA) package. Samawi and Muttlak ^[20] used 5000 replications in their paper because of that the replications were extended to be more than 5000.

The simulation results are summarized in list of figures from Figure 1 to Figure 6 for the performance of the samples means for the sample sizes $n = 6, 9, 11$ for one cycle ($m = 1$). The results of these figures show the relative efficiency (the ratio of the variances) of the estimators of the mean relative to SRS.

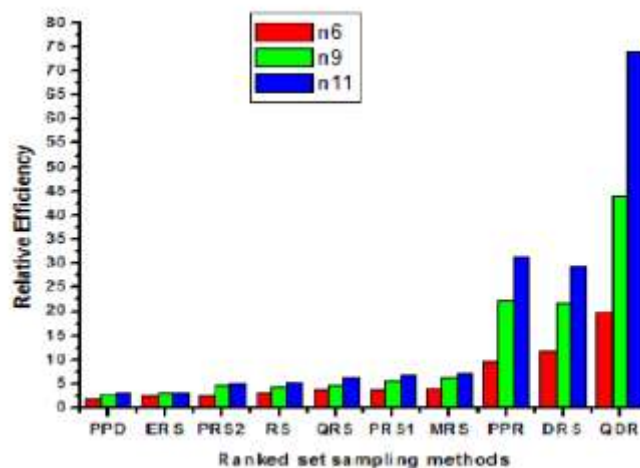


Figure 1 RE for standard normal distribution with $(\mu = 0, \sigma^2 = 1)$

For Standard Normal Distribution:

- Regarding the sample size n , the efficiency of the mean estimators using the ranked set sampling methods increases as n increases for almost results.
- Figure 1 shows that estimating the population mean using QDRSS is substantially more efficient than the other methods and ERSS is less efficient than the other methods.
- For almost sample sizes, MRSS is more efficient than RSS, ERSS, QRSS and PRSS with $(p = 0.33$ and $0.8)$. While DRSS is more efficient than RSS because it's a two stage of the RSS and give more accurate result.
- The results for RSS, ERSS, MRSS, QRSS, PRSS with $(p = 0.33$ and $0.8)$ and PPDRSS are very close more than the results from PPRSS, DRSS and QDRSS.
- The population mean estimator using DRSS is more efficient than the population mean estimator using PPDRSS.

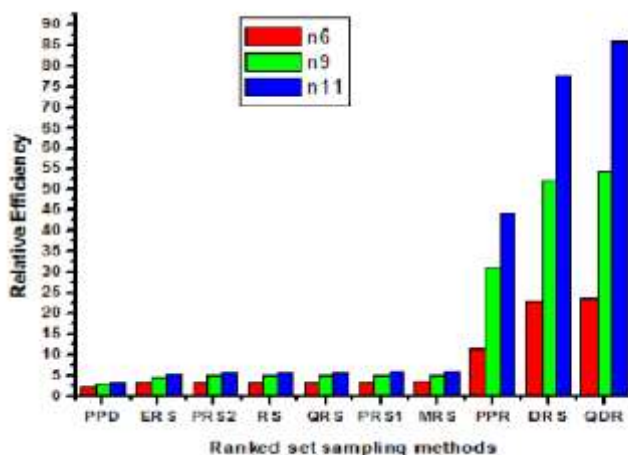


Figure 2 RE for beta distributions with $(\alpha = 3, \beta = 3)$

For Beta Distribution:

- Regarding the sample size n , the efficiency of the mean estimators using the ranked methods increases as n increases for almost results.
- Figure 2 shows that estimating the population mean using QDRSS is more efficient than the other methods and ERSS is less efficient than the other methods.
- For all the sample sizes, the results for RSS, ERSS, MRSS, QRSS and PRSS for $(p = 0.33 \text{ and } 0.8)$ are look constantly.
- The population mean estimator using DRSS and QDRSS are more efficient than the population mean estimator using RSS and QRSS respectively.

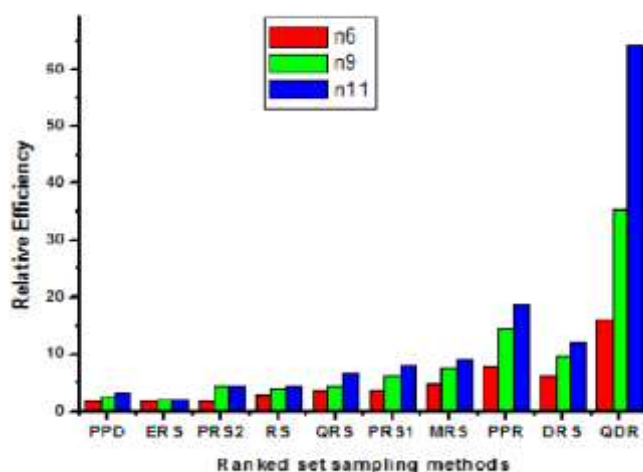


Figure 3 RE for logistic distribution with $(\alpha = -1, \beta = 1)$

For Logistic Distribution:

- Regarding the sample size n , the efficiency of the mean estimators using the ranked methods increases as n increases for almost results.
- Figure 3 shows that estimating the population mean using QDRSS is substantially more efficient than the other methods and ERSS is less efficient than the other methods.
- For almost sample sizes, MRSS is more efficient than RSS, ERSS, QRSS, PRSS with ($p = 0.33$ and 0.8) and PPDRSS. While DRSS is more efficient than RSS because it is a two stage of the RSS and give more accurate result.
- The results for RSS, ERSS, MRSS, QRSS, PRSS with ($p = 0.33$ and 0.8) and PPDRSS are very close more than the results from PPRSS, DRSS and QDRSS.

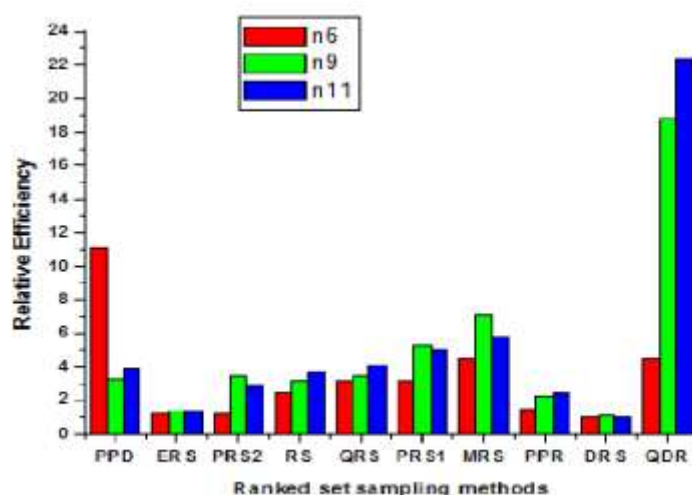


Figure 4 RE for exponential distributions with $(\lambda = 1)$

For Exponential Distribution:

- Regarding the sample size n , the efficiency of the mean estimators using the ranked methods increases as n increases for almost results.
- Figure 4 shows that estimating the population mean using QDRSS is substantially more efficient than the other methods and ERSS is less efficient than the other methods.
- For almost sample sizes, MRSS is more efficient than RSS, ERSS, QRSS, PRSS with ($p = 0.33$ and 0.8), PPRSS and DRSS.
- The results for RSS, ERSS, MRSS, QRSS, PRSS for ($p = 0.33$ and 0.8) and PPDRSS are very close more than the results from the other methods.

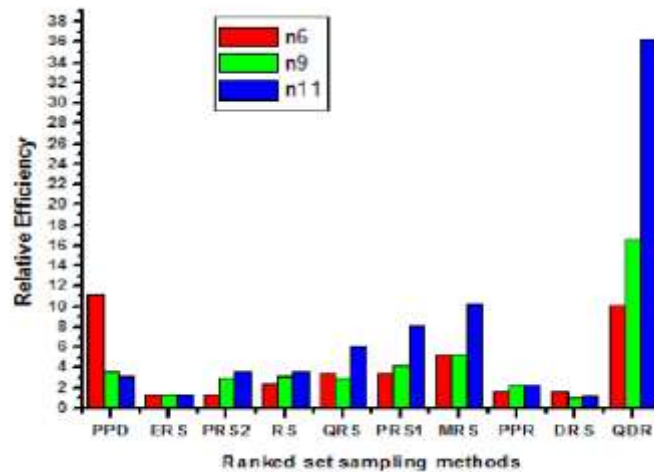


Figure 5 RE for gamma distribution with $(\alpha = 1, \beta = \frac{1}{2})$

For Gamma Distribution:

- This figure looks like figure 4 for all the methods because gamma distribution with shape parameter $\alpha = 1$ and scale parameter β is an exponential (β) distribution.
- Figure 5 shows that estimating the population mean using QDRSS is substantially more efficient than the other methods and ERSS is less efficient than the other methods.
- For almost sample sizes, MRSS is more efficient than RSS, ERSS, QRSS, PRSS with ($p = 0.33$ and 0.8), PPRSS and DRSS.
- The results for RSS, ERSS, MRSS, QRSS, PRSS for ($p = 0.33$ and 0.8) and PPDRSS are very close more than the results from the other methods.

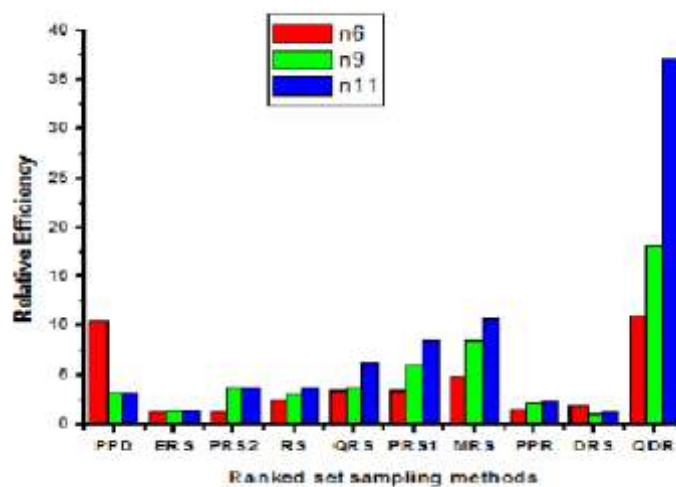


Figure 6 RE for weibull distribution with $(\lambda = 1, k = 2)$

For Weibull Distribution:

- This figure looks like figure 4 and figure 5 for all the methods because weibull $(1, k)$ random variable is an exponential random variable with mean k . Both the gamma and weibull distributions can be seen as generalizations of the exponential distribution.
- Figure 5 shows that estimating the population mean using QDRSS is substantially more efficient than the other methods and ERSS is less efficient than the other methods.
- For almost sample sizes, MRSS is more efficient than RSS, ERSS, QRSS, PRSS with $(p = 0.33$ and $0.8)$, PPRSS and DRSS.
- The results for RSS, ERSS, MRSS, QRSS, PRSS for $(p = 0.33$ and $0.8)$ and PPDRSS are very close more than the results from the other methods.

4. Summary and Conclusions

In this article, A comparative study for RE of the sample mean using RSS, PPRSS, ERSS, MRSS, QRSS, PRSS with $(p = 0.33$ and $0.8)$, DRSS, QDRSS and PPDRSS with respect to SRS is discussed. Based on the simulation results from the comparative study, the findings may be summarized as follows:

1. Regarding the sample size n , the efficiency of the mean estimators using the ranked set sampling methods increasing as the sample size n increasing, especially appear in symmetric distributions.
2. All the ranked set sampling methods are more efficient than SRS, the results given in list of figures confirmed this fact.
3. The QDRSS method is more efficient than the other methods which can called the best method for estimating the population mean.
4. The ERSS method is less efficient than the other methods which can called the worst method for estimating the population mean, especially appear in asymmetric distributions.

References

1. Al-Saleh, M. F., and Al- Kadiri, M. A., (2000). *Double ranked set sampling*. Statistics and Probability Letters 48, 205-212.

2. Al-Saleh , M. F., and Al- Omari, A. I., (2009). *Quartile double ranked set sampling for estimating the population mean*. Economic Quality control. 24, 243-253.
3. Balakrishnan, N., and Chen, W .W., (1997). *CRC Handbook of Tables for Order Statistics from Inverse Gussian Distributions with Applications*, CRC Press, Boca Raton.
4. Biradar, B. S., and Santosha, C. D., (2015). *Estimation of the Population Mean Based on Extremes Ranked Set Sampling*. American Journal of Mathematics and Statistics 5, 32-36.
5. Dell, T. R. and Clutter , J. L., (1972). *Ranked set sampling theory with order statistics background*. Biometrics 28, 545-555.
6. Haq, A., Brown, J., Moltchanova, E., and Al-Omari , A. I., (2013). *Partial ranked set sampling design*. Environmetrics 24, 201-207.
7. Haq, A., Brown, J., Moltchanova, E., and Al-Omari , A. I., (2016). *Paired double-ranked set sampling*. Communications in Statistics - Theory and Methods 10, 2873-2889.
8. Harter, H. L. and Balakrishnan, N., (1996). *CRC Handbook of Tables for the use of Order Statistics in Estimation*. CRC Press, Boca Raton.
9. Ibrahim, A., Al-Omari, A. I. , and Khalifa H. J., (2008). *Percentile Double Ranked Set Sampling*. Journal of Mathematics and Statistics 4, 60-64.
10. McIntyre, G. A., (1952). *A method of unbiased selective sampling, using ranked sets*. Australian J. Agricultural Research 3, 385-390.
11. Muttalak, H.A., (1996). *Pair ranked set sampling*. The Biometrical Journal 38, 879-885.
12. Muttalak, H. A., (1997). *Median ranked set sampling*. Journal of Applied Statistical Sciences 6, 245-255.
13. Muttalak, H. A., (1998). *Median ranked set sampling with concomitant variables and a comparison with ranked set sampling and regression estimator*. Environmetrics 9,255-267.

14. Muttlak, H. A., (2003). *Investigating the use of quartile ranked set samples for estimating the population mean*. Journal of Applied Mathematics and Computation 146,437-443.
15. Muttlak, H. A., (2003). *Modified ranked set sampling methods*. Pakistan Journal of Statistics 19, 315-323.
16. Patil, G. P., Sinha, A. K., Taillie, C., (1993). *Relative precision of ranked set sampling: Comparison with the regression estimator*. Environmetrics 4, 399-412.
17. Richard, S. L., Dennis, W.,and William, M., (2008). *Mathematical Statistics with Applications* .7th ed. Thomson Higher Education. USE.
18. Samawi, H. M., (2011). *Varied set size ranked set sampling with applications to mean and ratio estimation*. The International Journal of Simulation Modeling 31 , 6–13.
19. Samawi, H. M., Abu-Dayyeh, W. and Ahmed, S., (1996). *Extreme ranked set sampling*. The Biometrical Journal 30 , 577-586.
20. Samawi, H. M., and Muttlak, H. A., (1996). *Estimation of Ratio Using Rank Set Sampling* . Biometrical Journal 38 ,753-764.
21. Takahasi, K. and Wakimoto, K., (1968). *On unbiased estimates of the population mean based on the sample stratified by means of ordering*. Annals of the Institute of Statistical Mathematics 20 , 1-31.

جامعة القاهرة
معهد الدراسات والبحوث الإحصائية

المؤتمر السنوى الثانى والخمسين للإحصاء
وعلوم الحاسب وبحوث العمليات

الإحصاء التطبيقي

٢٥-٢٧ ديسمبر ٢٠١٧

فهرس الإحصاء التطبيقي

٢٢-١	أثر مشكلة التجميع على جودة توفيق نموذج الإنحدار الخطي " دراسة قياسية لمحددات القيمة السوقية لسعر السهم " مصطفى عبد المنعم الخواجة ، نبيلة أحمد سعد	١
٣٧-٢٣	أهمية قواعد وإجراءات السلامة والصحة المهنية على مراحل إدارة الإزمات والكوارث (دراسة تحليلية على إحدى شركات الإتصالات) عصام عطية عبد المنعم	٢
٥١-٣٨	مؤشرات إحصائية للحوادث المرورية بإستخدام تحليل السلاسل الزمنية بمنطقة عسير السعودية للفترة من (١٤١٧-١٤٣٤) محمد محمد أحمد المزاح	٣

أثر مشكلة التجميع على جودة توفيق نموذج الانحدار الخطي

" دراسة قياسية لمحددات القيمة السوقية لسعر السهم "

¹مصطفى عبد المنعم محمد الخواجة

²نبيلة احمد سعد على

المستخلص

تعتبر مشكلة التجميع أحد مشاكل القياس التي يتأثر بها نموذج الانحدار الخطي، والتي تنشأ نتيجة الإهتمام بالسلوك التجميعي والحصول على بيانات تجميعية وإستخدامها في تقدير النموذج وإتخاذ القرارات من خلاله. ويمثل الوصول لأثر مشكلة التجميع على جودة توفيق نموذج الانحدار الخطي أحد أهداف الدراسة الحالية، وفي حين تم استقراء العديد من الدراسات التي لخصت تلك الآثار. وقد خلصت الدراسة إلى بعض تلك الآثار بالتطبيق على محددات القيمة السوقية لسعر السهم في سوق الأوراق المالية المصرية.

ومن تلك الآثار انخفاض كفاءة مقدرات OLS، وزيادة قيم معامل التحديد إذا تم تجميع البيانات على مستوى الزمن. ولكن لم يتحدد أثر واضح للتجميع على قيم معامل التحديد إذا تم التجميع على مستوى الوحدات القطاعية. كما يُحسن التجميع على مستوى الزمن من قدرة نموذج الانحدار الخطي على التنبؤ، بينما يضعف التجميع على مستوى الوحدات القطاعية من قدرة النموذج على التنبؤ. وتنخفض دقة تنبؤات النموذج التجميعي الذي يعتمد على المجموع الكلي في التجميع بينما تزداد دقة التنبؤات التي تعتمد على أحد مقياس النزعة المركزية في التجميع.

الكلمات الدالية للبحث:

تجميع البيانات، جودة توفيق نموذج الانحدار الخطي، دقة تنبؤات نموذج الانحدار الخطي، قدرة نموذج الانحدار الخطي على التنبؤ، كفاءة مقدرات نموذج الانحدار الخطي.

1- مقدمة

يُطلق مصطلح التجميع Aggregation إذا تم أخذ قيمة واحدة في السنة كمتوسط، أو إجمالي عدة قيم لمتغير ما عند نقطة زمنية محددة. كما تهتم نظرية التجميع بتحويل العلاقات التفصيلية إلى علاقات تجميعية ودراسة العلاقة بينهما. ويهتم الكثير من الباحثين بوصف العديد من العلاقات الإقتصادية في

¹د.مصطفى عبد المنعم محمد الخواجة

أستاذ الإحصاء التطبيقي بقسم الإحصاء والرياضة والتأمين

كلية التجارة جامعة الاسكندرية

²نبيلة احمد سعد على

المعيدة بقسم الإحصاء والرياضة والتأمين

كلية التجارة جامعة الاسكندرية

شكل نموذج تجميعي، والذي يعتمد على بيانات يتم تجميعها على عدة مستويات تختلف وفقاً لنوع التجميع.

وذلك بالرغم من الآثار المختلفة لاستخدام هذا النموذج من حيث تأثيره على كفاءة التقديرات، والقوة التفسيرية للنموذج، والقدرة التنبؤية للنموذج ودقة التنبؤات. وترجع تلك الآثار لأن التعامل مع بيانات تفصيلية يسمح بوجود دقة أكبر في التوصيف والتقدير والتنبؤ. وتختلف تلك الآثار باختلاف نوع وأسلوب التجميع المتبع. وعليه يمكن تلخيص مشكلة الدراسة في التساؤلات الآتية:

- 1- ما هو أثر مشكلة التجميع على كفاءة مقدرات نموذج الانحدار الخطي؟.
- 2- بما تؤثر مشكلة التجميع على القوة التفسيرية لنموذج الانحدار الخطي؟.
- 3- كيف تتأثر القدرة التنبؤية لنموذج الانحدار الخطي بمشكلة التجميع، وكذلك دقة التنبؤات؟.

وتهدف الدراسة بصورة أساسية إلى الإجابة عن التساؤلات السابقة من خلال التعرف على مشكلة التجميع، وعرض أنواع التجميع المختلفة كالتجميع على مستوى الزمن والتجميع على مستوى الوحدات القطاعية، وذلك باستخدام المجموع الكلي أو أحد مقاييس النزعة المركزية كأساليب للتجميع. ودراسة آثار التجميع على كفاءة مقدرات نموذج الانحدار الخطي المقدر، والقوة التفسيرية والقدرة التنبؤية للنموذج، وأيضاً دقة تنبؤات النموذج. ويتم ذلك عملياً من خلال دراسة قياسية لبعض محددات القيمة السوقية لسعر السهم في سوق الأوراق المالية المصرية.

كما تتبع أهمية الدراسة في أنها :

- 1- تستعرض أهم الدراسات السابقة التي قدمت مشكلة التجميع كأحد مشاكل القياس، وأنواع وأساليب تجميع البيانات داخل العلاقات الخطية.
- 2- تقدم دراسة قياسية لبعض محددات القيمة السوقية لسعر السهم في سوق الأوراق المالية المصرية، وذلك لدراسة أثر مشكلة التجميع على جودة توفيق نموذج الانحدار الخطي.

2- أنواع التجميع

يمكن تقسيم تجميع البيانات في العلاقات إلى نوعين هما: تجميع البيانات الخطي، وتجميع البيانات غير الخطي. وفي كلا النوعين يمكن تجميع البيانات على مستوى الزمن Temporal Aggregation، أو تجميع البيانات على مستوى الوحدة القطاعية Longitudinal Aggregation. ويقصد بالوحدة القطاعية "المؤسسة أو القطاع أو البلد أو إقليم ما" ويتوقف ذلك على الظاهرة محل الدراسة.

مع مراعاة أن اختلاف طبيعة البيانات الخاصة بمتغيرات النموذج هي التي تفرض نوع التجميع الذي سيتطلب تطبيقه. أي أن التجميع على مستوى الزمن يمكن أن يُطبق على بيانات السلاسل الزمنية. في حين أن بيانات القطاعات المستعرضة وبيانات السلاسل القطاعية يمكن أن يُطبق عليها التجميع على مستوى الوحدة القطاعية، والتجميع على مستوى الزمن.

(1-2) التجميع غير الخطى للبيانات

تناول [8] Plassmann التجميع غير الخطى للبيانات حيث يتحقق ذلك النوع من التجميع في حالتين إحداهما: أن تكون معادلة النموذج التفصيلي غير خطية، والأخرى إذا تم استخدام أساليب التجميع بصورة غير خطية أي تكوين متغيرات النموذج التجميعية كدالة غير خطية. وقد أشار Garderen [1] et al. إلى مشكلة التجميع غير الخطى، حيث أكد أن التجميع غير الخطى يُعد مشكلة معقدة لا تصل إلى حلول مرضية إلا في ظل شروط قوية تتعلق بشكل وتكوين النموذج غير الخطى. والتي تعتمد على مدى التجانس في سلوك المؤسسات أو القطاعات محل الظاهرة على المستوى التفصيلي.

كما وضح [8] Plassmann أن تجميع البيانات في العلاقات غير الخطية يهتم بتحويل النموذج التفصيلي إلى نموذج تجميعي غير خطى، ثم أخذ اللوغاريتم لتلك المتغيرات التجميعية ذلك لتحويله إلى نموذج انحدار تجميعي خطى.

(2-2) التجميع الخطى للبيانات

وفيه بعض المفاهيم الأساسية لنظرية التجميع والتي طرحها كلٌّ من [10] Theil، و Plassmann [8]، و [6] Monteforte، و [5] Lutero. حيث أشار [10] Theil إلى نظرية تجميع البيانات في العلاقات الخطية و التي تهتم بتحويل العلاقات التفصيلية إلى علاقات تجميعية، ودراسة العلاقات بين النماذج على المستويين التفصيلي والتجميعي.

● **المستوى التفصيلي** ويتكون من عدد m من الوحدات القطاعية محل الظاهرة مثل: البلاد أو الأقاليم أو المناطق الجغرافية أو الأسر أو الشركات وغيرها، وكل مؤسسة يتاح لها عدد T من البيانات. بحيث يتم وصف كل مؤسسة بمتغير تابع واحد و عدة متغيرات تفسيرية، كما يمكن وصف النموذج التفصيلي كما في المعادلة الآتية:

$$Y_{(mT,1)} = X_{(mT,k)}\beta_{(k,1)} + u_{(mT,1)} \quad (1)$$

حيث إن:

$Y_{(mT,1)}$: يعبر عن متجه عمود للمتغير التابع التفصيلي.

$X_{(mT,k)}$: تعبر عن مصفوفة المتغيرات التفسيرية التفصيلية.

$\beta_{(k,1)}$: يعبر عن متجه المعاملات نتيجة انحدار المتغير التابع التفصيلي على المتغيرات التفسيرية التفصيلية.

$u_{(mT,1)}$: يعبر عن متجه عمود لحد الخطأ العشوائى في النموذج التفصيلي.

● **المستوى التجميعي** ويتم وصفه بمتغير تابع تجميعي واحد، و عدة متغيرات تفسيرية تجميعية ومن خلالها يتم تكوين مصفوفة المتغيرات التفسيرية التجميعية. كما يمكن وصف النموذج التجميعي بالمعادلة الآتية:

$$Y_{(T,1)} = X_{(T,k)}\beta_{(k,1)} + u_{(T,1)} \quad (2)$$

حيث إن:

$Y_{(T,1)}$: يعبر عن متجه عمود للمتغير التابع التجميعي.

$X_{(T,k)}$: تعبر عن مصفوفة المتغيرات التفسيرية التجميعية.
 $\beta_{(k,1)}$: تعبر عن متجه المعاملات نتيجة انحدار المتغير التابع التجميعي على المتغيرات التفسيرية التجميعية.
 $u_{(T,1)}$: يعبر عن متجه عمود لحد الخطأ العشوائي في النموذج التجميعي.

ويلاحظ من المعادلتين (1)،(2) أن الفارق بين المستويين التفصيلي والتجميعي هو تقليل عدد المشاهدات على المستوى التجميعي وذلك نتيجة أخذ المجموع الكلي أو أحد مقاييس النزعة المركزية. كما أن عدد معالم نموذج الانحدار لا تتأثر بتجميع البيانات، ولكن بالتأكد تختلف قيمة مقدرات تلك المعالم. ويمكن تجميع البيانات على عدة مستويات مختلفة مثل التجميع على مستوى الزمن والتجميع على مستوى الوحدة القطاعية محل الظاهرة، وذلك إذا كان تجميع البيانات خطي أو غير خطي ويمكن تناولهما كالاتي:

(2-3) التجميع على مستوى الزمن

ذكر [5] Lutero أن تطبيق التجميع على مستوى الزمن على المدى الطويل يُعد بمثابة تحويله لبيانات قصيرة المدى تُشاهد كل فترة قصيرة أي (بيانات تتكرر بصورة كبيرة) مثل البيانات الربع سنوية إلى بيانات طويلة المدى تُشاهد كل فترة طويلة أي (بيانات تتكرر بصورة أقل) مثل البيانات السنوية. ويُعرف بالتجميع على مستوى الزمن، واقتصر هذا النوع من التجميع على نماذج وحيدة المتغير، ثم توسعت الدراسات السابقة إلى نماذج متعددة المتغيرات.

وقد أشار [3] Hassler إلى وجود شكلين من بيانات السلاسل الزمنية التي يُجرى عليها التجميع على مستوى الزمن وهما: بيانات الأوراق المالية Stock Data، وبيانات التدفق Flow Data. وتُعتبر بيانات التدفق عن الاستهلاك الشهري حيث يتم تطبيق التجميع على مستوى الزمن بتحويل البيانات الشهرية إلى بيانات ربع سنوية أو بيانات سنوية. كما يُقصد ببيانات الأوراق المالية إنها الأسعار اليومية للأوراق المالية ومعدلات الصرف. وبتطبيق التجميع على مستوى الزمن يتم تحويل البيانات اليومية إلى بيانات أسبوعية أو بيانات شهرية، وذلك بأخذ أحد مقاييس النزعة المركزية لكل الأيام في الأسبوع للحصول على بيانات أسبوعية، أو أخذ مقاييس النزعة المركزية لكل الأيام في الشهر للحصول على بيانات شهرية.

وقد أكد [7] Petkovic & Veredas أن التجميع على مستوى الزمن لا يؤثر على توصيف النموذج في حالة النماذج الساكنة. إلا أنه يؤثر على تكوين وهيكل النموذج في حالة النماذج الديناميكية. وأكد [7] Petkovic & Veredas أن التجميع على مستوى الزمن بصفة عامة ينتج عنه مشكلة ارتباط ذاتي بين البواقي. كما استنتج أيضاً أن التجميع على مستوى الزمن يؤثر على خصائص المقدرات في النماذج الخطية. وذلك سواء كانت النماذج ساكنة أو ديناميكية. كما يُسبب التجميع على مستوى الزمن فقد في المعلومات والمشاهدات مما يقلل من كفاءة مقدرات النموذج التجميعي. وهذا يؤثر على مدى معقولية وقبول النموذج التجميعي.

وقد درس [9] Rossana & Seater أثر التجميع على مستوى الزمن على خصائص وثبات السلاسل الزمنية، وقد استنتج أن هذا النوع من التجميع يغير من خصائص السلاسل الزمنية للبيانات. ولذلك فإن الاختلافات الأساسية في البيانات التفصيلية تفقد كلياً مما يغير من خصائص السلسلة

الزمنية. أما عن الثبات فقد ذكر أن البيانات التجميعية تعد أكثر ثباتاً على المدى الطويل من البيانات التفصيلية.

ولكن حين تناول [3] Hassler الثبات في ظل مشكلة التجميع على مستوى الزمن تعارضت نتائج مع نتائج [9] Rossana & Seater . حيث أكد أن ثبات السلسلة الزمنية يختلف باختلاف مستوى التجميع، وقد أشار إلى عدم وجود أمر يقيني إن كان الثبات يتزايد أم يتناقص بسبب التجميع على مستوى الزمن. ولكنه أكد على أن ثبات السلسلة الزمنية يقل بزيادة الفترة التي يتم التجميع على أساسها.

وقد أشارت دراسة [2] Gordevicius et al. أن التجميع على مستوى الزمن يُعد معالجة هامة لتلخيص بيانات السلاسل الزمنية. وتتوافر بيانات سلاسل زمنية في معظم القطاعات كالقطاع المالي والقطاع الطبي وغيرها من القطاعات. كما يُستخدم التجميع على مستوى الزمن لتلخيص مجموعات كبيرة من البيانات بقيمة محددة عند نقطة زمنية معينة أو عند فترة زمنية معينة.

(2-4) التجميع على مستوى الوحدة القطاعية

لقد عرف [5] Lutero التجميع على مستوى الوحدة القطاعية أنه ذلك النوع الذي يهتم بالتجميع عبر النطاق الجغرافي، وعادة يكون على مستوى أقاليم أو مناطق إدارية أو مجموعة من المؤسسات أو القطاعات مثل الأسر أو الشركات وذلك باستخدام أحد أساليب التجميع. ويتم تناول و دراسة أساليب التجميع في المبحث التالي لما له من أهمية قصوى حيث لا يمكن الاقتصار على أسلوب بعينه للتعامل مع كافة أنواع البيانات، كما ينعكس أثر التجميع على خصائص نموذج الانحدار الخطي المقدر بتغيير أسلوب التجميع.

وبدأ [10] Theil التحدث عن مشكلة التجميع على مستوى المؤسسات في مجال الاقتصاد القياسي كأحد مشاكل القياس، وقد بدأ بافتراضين وهما: أن النموذجين التفصيلي والتجميعي يتم توصيفهما توصيفاً صحيحاً أمثلاً، وأن كل معادلة تفصيلية في النموذج التفصيلي لها نفس العدد من المتغيرات التفسيرية. وقد ذكر [7] Petkovic & Veredas أن هذا النوع من التجميع يتم لنماذج السلاسل القطاعية التي يتم مشاهدتها على أساس بعدين أحدهم الزمن والآخر الوحدات القطاعية. لذا يترتب على هذا النوع من التجميع ظهور مشكلة عدم ثبات التباين.

3- أساليب التجميع

بالعودة إلى مصطلح التجميع يمكن استنتاج أن أساليب التجميع متعددة كالتجميع باستخدام المجموع الكلي وذلك كما أشار [6] Monteforte، أو الاعتماد على أحد مقاييس النزعة المركزية كالوسط الحسابي كما عرض [10] Theil، والوسط التوافقي وذلك كما ذكر [1] Garderen et al، واستخدام الوسط المرجح بأوزان لعدد من السلاسل التفصيلية وهذا ما نوه عنه [8] Plassmann. وعليه يمكن أن تنقسم أساليب التجميع إلى المجموع الكلي أو استخدام أحد مقاييس النزعة المركزية.

وأكد [11] Yager أن أساليب التعامل مع البيانات تختلف وفقاً لنوع البيانات أي أنه إذا كانت البيانات كمية يصلح لها المجموع الكلي والوسط الحسابي والوسيط والمنوال كأساليب للتجميع. ولكن إذا كانت البيانات كمية ولها أوزان معينة خاصة بها فإن الأسلوب المناسب هو الوسط المرجح

بأوزان. أما إذا كانت البيانات غير كمية "وصفية" فإنه لا يمكن استخدام أى مقياس للنزعة المركزية سوى المنوال فى تلك الحالة. ويتم عرض بعض أساليب التجميع كما يلي:

(1-3) الوسط الحسابى للمشاهدات كأحد أساليب التجميع

يُعد [10] Theil أول من أشار إلى التجميع الخطى للبيانات، وفى ظل فروض [10] Theil بأن النموذجين التفصيلى والتجميعى يتم توصيفهما توصيفاً صحيحاً أمثلاً، وأن كل معادلة تفصيلية فى النموذج التفصيلى لها نفس العدد من المتغيرات التفسيرية. وبالرجوع إلى المعادلة التفصيلية (1)، وبتطبيق التجميع الخطى على مستوى الوحدات القطاعية وذلك بتكوين المتغير التجميعى باستخدام الوسط الحسابى كأحد مقاييس النزعة المركزية كالاتى:

$$\bar{Y} = \frac{1}{m} \sum_{i=1}^m Y_i \quad (3)$$

وبالتعويض من معادلة (1) فى معادلة (3)، فإن النموذج التجميعى يأخذ الشكل الآتى:

$$\bar{Y} = \frac{1}{m} \sum_{i=1}^m X_i \beta_i + \frac{1}{m} \sum_{i=1}^m u_i \quad (4)$$

وفى حالة إتاحة البيانات فى صورتها التجميعية، فإنه يتم التعامل مع متغيرات تجميعية (\bar{X}, \bar{Y}) أى التعامل مع نموذج تجميعى به علاقة خطية بسيطة فى متغيرات تجميعية. ومتجه معاملتها يناظر متجه المعاملات فى النموذج التفصيلى، والبواقي فى العلاقة التجميعية تساوى متوسط البواقي على المستوى التفصيلى المناظرة لها. ويتم التعامل مع المعادلة التجميعية التى تأخذ الشكل الآتى:

$$\bar{Y} = \bar{X} \beta + \frac{1}{m} \sum_{i=1}^m u_i$$

(5)

وبناءً على ما سبق فإن المعادلة التجميعية الصحيحة (4) لن تساوى المعادلة المقدره (5) إلا إذا كانت المعاملات على المستوى التفصيلى متساوية، بينما إذا كانت المعاملات على المستوى التفصيلى غير متساوية سيكون هناك خطأ فى التوصيف. ولذلك يذكر [10] Theil أن التجميع يعد حالة خاصة وتطبيقاً لتحليل التوصيف Specification analysis.

وقد استعان [10] Theil بإجراء الانحدار المساعد Auxiliary Regression، ذلك النوع من الانحدار الذى يستخدم لدراسة التناظر بين متغيرات النموذجين التفصيلى والتجميعى، وفيه يتم إجراء انحدار للمتغير التفسيري التفصيلى مضروباً فى مقلوب عدد المؤسسات $(\frac{1}{m} X_i)$ على المتغير التفسيري التجميعى \bar{X} . ويعبر عن معادلة الانحدار المساعد بصيغة المصفوفات بالشكل الآتى:

$$\frac{1}{m} X_i = \bar{X} B_i + \epsilon_i \quad (6)$$

حيث تعبر B_i عن مصفوفة معاملات نموذج الانحدار المساعد، وتأخذ المعادلة الآتية:

$$B_i = (\bar{X} \bar{X})^{-1} \bar{X} \frac{1}{m} X_i \quad (7)$$

وهى مصفوفة الأوزان الخاصة بالمعاملات المناظرة وغير المناظرة بين النموذجين التفصيلى والتجميعى. ويكون فيها مجموع الأوزان الخاصة بالمعاملات المناظرة يساوى الواحد الصحيح، ومجموع الأوزان الخاصة بالمعاملات غير المناظرة يساوى صفر. ويُقصد بالتناظر هنا على سبيل المثال أن توقع مقدر معلمة الميل على المستوى التجميعى يعتمد على معالم الميل على مستوى جميع

المؤسسات أو القطاعات (المستوى التفصيلي) في حين أن مقدرات معالم النموذج التجميعي الأخرى لا تتأثر بمعالم الميل على مستوى جميع المؤسسات أو القطاعات.

(2-3) المجموع الكلي للمشاهدات كأحد أساليب التجميع

لقد فرض [5] Lutero أن معادلة النموذج التفصيلي بصيغة المصفوفات للمؤسسة $i = 1, \dots, m$ ، وذلك عبر نقاط زمنية $t = 1, \dots, T$ تأخذ المعادلة (1). وبفرض أن التجميع الذي يتم تطبيقه تجميع خطي على مستوى المؤسسات أو القطاعات، وبتكوين المتغير التابع التجميعي باستخدام المجموع الكلي لمشاهدات المتغيرات التابعة على مستوى عدد m من المؤسسات بالشكل الآتي:

$$Y_{at} = \sum_{i=1}^m Y_{it} \quad (9)$$

حيث إن Y_{at} يعبر عن إجمالي قيم مشاهدات المتغير التابع وذلك على مستوى جميع المؤسسات محل الدراسة وعند نقطة زمنية t .

وبتكوين المتغير التفسيري التجميعي باستخدام المجموع الكلي لمشاهدات المتغيرات التفسيرية التفصيلية كما بالشكل الآتي:

$$X_{at} = \sum_{i=1}^m X_{it} \quad (10)$$

حيث X_{at} يعبر عن إجمالي قيم مشاهدات المتغير المستقل وذلك على مستوى جميع المؤسسات محل الدراسة وعند نقطة زمنية t .

فإن النموذج التجميعي المقدر يأخذ المعادلة (2)، بحيث تعرف متغيرات النموذج التجميعي باستخدام المجموع الكلي. ولدراسة العلاقة بين النموذجين التفصيلي والتجميعي والتناظر الإحصائي بينهما يتم الاستعانة بالانحدار المساعد Auxiliary Regression، وذلك بإجراء انحدار لمصفوفة المتغيرات التفسيرية في النموذج التفصيلي X_{it} لكل مؤسسة على حدة، وذلك الانحدار على المصفوفة القطرية لمتغيرات النموذج التجميعي التفسيرية وكما سبق بالمعادلة (6). ولكن في ظل النموذج التجميعي الذي تعرف متغيراته بالمجموع الكلي. وتأخذ معادلة الانحدار المساعد الشكل الآتي:

$$X_{it} = X_{at} \alpha_i + \epsilon_{it} \quad (11)$$

بحيث يعبر ϵ_{it} عن حد الخطأ العشوائي لمعادلة الانحدار المساعد، ويعبر α_i عن مصفوفة الأوزان الخاصة بالمعاملات المناظرة وغير المناظرة. ويكون مجموع تلك الأوزان الخاصة بالمعاملات المناظرة يساوي الواحد الصحيح في حين أن مجموع تلك الأوزان يساوي الصفر للمعاملات غير المناظرة.

وللوصول إلى العلاقة بين حد الخطأ العشوائي على المستوى التفصيلي، وحد الخطأ العشوائي على المستوى التجميعي يتم التعويض من المعادلة (11) في المعادلة (1) وبأخذ المجموع الكلي على مستوى المؤسسات أو القطاعات، ثم تقدير حد الخطأ العشوائي في النموذج التجميعي بالشكل التالي:

$$\hat{u}_{at} = \sum_{i=1}^m (\epsilon_{it} \beta_i + \hat{u}_{it}) \quad (12)$$

أي أن حد الخطأ العشوائي في النموذج التجميعي عبارة عن مجموع مقدارين حد الخطأ العشوائي في النموذج التفصيلي بعد التجميع، بالإضافة إلى خطأ التجميع وهو حاصل ضرب بواقي نموذج الانحدار المساعد ومقدر معلمة النموذج التفصيلي المناظرة له.

4- دراسة قياسية لبعض محددات القيمة السوقية لسعر السهم

لدراسة الآثار المترتبة على مشكلة التجميع، فقد تتضمن هذا الجزء اطار وبيانات الدراسة القياسية والأساليب الاحصائية المستخدمة ونتائج تلك الدراسة ومناقشة أهم النتائج والدلالات والمضامين الأكاديمية الخاصة بها.

(1-4) متغيرات وبيانات الدراسة

تشتمل الدراسة على ثلاثة متغيرات أحدهم يمثل المتغير التابع وهو القيمة السوقية لسعر السهم اليومي، والمتغيرين الاخرين هما التفسيرين ويمثلان أعلى سعر للسهم تم التعامل اليومي به وأقل سعر أيضاً. وتأخذ معادلة نموذج الانحدار المقدر الصورة التالية:

$$\hat{Y}_{it} = \hat{\beta}_0 + \hat{\beta}_1 X_{it} + \hat{\beta}_2 Z_{it}$$

بحيث أن:

\hat{Y}_{it} : يمثل القيمة المقدر لسعر اقبال السهم وهو آخر سعر تم التعامل اليومي به.

X_{it} : يمثل أعلى سعر يومي للسهم تم التعامل به، وهو أحد المتغيرات التفسيرية في نموذج الانحدار.

Z_{it} : يمثل أقل سعر يومي للسهم تم التعامل به، وهو أحد المتغيرات التفسيرية في نموذج الانحدار.

وكل من " $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ " يمثلوا مقدرات معالم نموذج الانحدار الخطي.

أما عن بيانات الدراسة تم الاستعانة ببيانات سلاسل زمنية الخاصة بعشر أوراق مالية الأعلى من حيث السيولة والنشاط كما في جدول (1) ووفقاً للأسبوع الأخير من شهر مارس لعام 2017، حيث يطبق على تلك البيانات التجميع على مستوى الزمن " شهرياً وربع سنوي ". وتجمع تلك البيانات في صورة سلاسل زمنية عن الفترة الممتدة من أول يناير 1998 إلى آخر يناير 2015 بشكل يومي بعد استبعاد الاجازات الرسمية وأيام العطلة والأيام التي لا يصدر عنها تقرير يومي.

جدول (1)

بيان بالشركات التي يتم بواسطتها تطبيق التجميع على مستوى الزمن

عدد المشاهدات على المستوى الربع سنوي	عدد المشاهدات على مستوى الشهر	عدد المشاهدات على المستوى التفصيلي	اسم الشركة محل الدراسة
29	86	1717	1- طلعت مصطفى
68	205	3404	2- مصر الجديدة
67	199	3771	3- السادس من أكتوبر
63	187	3554	4- هيرمس القابضة
68	204	4057	5- نصر للإسكان
35	103	1296	6- أجواء للصناعات
68	203	3873	7-النساجون الشرقيون
58	172	3463	8- جلوبال تيلكوم
37	109	2196	9- المصرية للاتصالات
68	204	4138	10- البنك التجاري الدولي

ولتطبيق التجميع على مستوى الوحدة القطاعية كالمؤسسات والشركات تم الإستعانة ببيانات فى صورة سلاسل مقطعية لعدد 14 قطاعاً إقتصادياً داخل سوق المقصورة ويحتوى جدول (2) على بيان بتلك البيانات، بحيث يحتوى كل قطاع على عدد من المؤسسات.

جدول (2)

بيان بالشركات التى يتم بواسطتها تطبيق التجميع على مستوى المؤسسات

عدد المشاهدات على مستوى التجميع	عدد المشاهدات على المستوى التفصيلي	عدد المؤسسات	اسم القطاع محل الدراسة
1940	22222	12	1- البنوك
1940	12949	7	2- الموارد الأساسية
1940	7617	4	3- الاتصالات
2051	35337	20	4- الأغذية والمشروبات
2051	34025	19	5- التشيد والبناء
2058	15492	8	6- الخدمات المالية
1940	10339	9	7- السياحة والترفيه
1940	38677	22	8- العقارات
1940	3612	2	9- الغاز والبتروول
2036	11296	6	10- الكيماويات
1940	20179	11	11- الخدمات والمنتجات الصناعية
2067	10195	7	12- الرعاية الصحية والأدوية
1940	13139	7	13- المنتجات المنزلية والشخصية
1939	3457	3	14- الموزعين وتجار التجزئة

(2-4) نتائج الدراسة

تم تجميع وتحليل البيانات باستخدام برمجة R وذلك مستعيناً بالأوامر الهامة كما ذكرها Kabacoff [4]. وتنقسم نتائج الدراسة إلى قسمين رئيسيين يتلخصان فى أثر مشكلة التجميع على مستوى الزمن وأثر مشكلة التجميع على مستوى الوحدات القطاعية، وذلك على كل من: القوة التفسيرية لنموذج الانحدار الخطى، وكفاءة مقدرات معالم النموذج، والقدرة التنبؤية لنموذج الانحدار الخطى، وأيضاً دقة تنبؤات نموذج الانحدار الخطى.

(1-2-4) أثر مشكلة التجميع على مستوى الزمن

أولاً: القوة التفسيرية للنموذج وتقاس تلك القوة التفسيرية لنموذج الانحدار الخطى المقدر باستخدام معامل التحديد، تلك القيمة التى تفسر نسبة التغيرات التى تحدث فى المتغير التابع، والتى يمكن إرجاعها إلى المتغيرات المستقلة محل الدراسة. ويمكن الحصول على معامل التحديد فى نموذج الانحدار الخطى من المعادلة الآتية:

$$R^2 = 1 - \frac{\sum_{t=1}^n u_t^2}{\sum_{t=1}^n y_t^2} \quad (13)$$

بحيث إن:

$$\sum_{t=1}^n u_t^2 : \text{يعبر عن مجموع مربعات البواقي فى نموذج الانحدار الخطى.}$$

$$\sum_{t=1}^n y_t^2 : \text{يعبر عن مجموع المربعات الكلية فى نموذج الانحدار الخطى.}$$

فى حين n تعبر عن حجم العينة وتساوى mT فى النموذج التفصيلى أى قبل تجميع البيانات، وتساوى T فى النموذج التجميعى. وبتطبيق التجميع على مستوى الزمن "شهرياً وربع سنوياً" باستخدام بيانات العشر شركات الأكثر نشاطاً وسيولة فى سوق البورصة المصرية، ثم حساب قيمة معامل التحديد فى النموذج التفصيلى، والقيمة المناظرة لها فى النموذج التجميعى وذلك كما بالمعادلة (13). وقد تم الحصول على النتائج المفروغة فى جدول (3) كما يلى.

جدول (3)

معامل تحديد نموذج الانحدار الخطى المتعدد " R^2 " فى ظل التجميع على مستوى الزمن

تجميع البيانات ربع سنوياً			تجميع البيانات شهرياً			نوع التجميع	
المجموع الكلى	الوسيط	الوسط الحسابى	المجموع الكلى	الوسيط	الوسط الحسابى	أسلوب التجميع	الشركات
						النموذج التفصيلى	
100%	99.99%	100%	100%	99.98%	99.99%	99.97%	1- طلعت مصطفى
100%	100%	100%	100%	99.99%	100%	99.99%	2- مصر الجديدة
100%	99.99%	100%	100%	100%	100%	99.98%	3- السادس من أكتوبر
100%	99.99%	100%	100%	99.99%	100%	99.98%	4- هيرمس القابضة
100%	100%	100%	100%	100%	100%	99.94%	5- نصر للإسكان
100%	99.98%	99.99%	100%	99.96%	100%	99.89%	6- أجواء للصناعات
99.94%	99.99%	99.92%	99.85%	99.98%	99.85%	97.98%	7- النساجون الشرقيون
100%	100%	100%	100%	99.99%	100%	99.98%	8- جلوبال تيلكوم
100%	99.97%	99.99%	99.99%	99.90%	99.97%	99.80%	9- المصرية للاتصالات
100%	100%	100%	100%	100%	100%	99.99%	10- البنك التجارى الدولى

وبناءً على الجدول السابق، فإن تجميع البيانات على مستوى الزمن سواءً تجميع على مستوى الشهر أو تجميعاً على المستوى الربع سنوياً، وبواسطة أي من أساليب التجميع كالمجموع الكلى أو أحد مقاييس النزعة المركزية كالوسط الحسابى أو الوسيط، فإن قيمة معامل التحديد فى النموذج التجميعى تزداد بالمقارنة بقيمة معامل التحديد المناظرة لها فى النموذج التفصيلى.

ثانياً: كفاءة مقدرات معالم النموذج ولدراسة أثر مشكلة التجميع على مستوى الزمن (شهرياً وربع سنوياً) على كفاءة مقدرات معالم النموذج يتم استخدام الخطأ المعياري كمقياس لكفاءة المقدرات. ويتم تلخيص ذلك الأثر فى حالة نموذج الانحدار الخطى المتعدد كما هو موضح فى جدول (4) التالى.

جدول (4): الخطأ المعياري لمقدرات معالم نموذج الانحدار الخطي وتأثره بالتجميع على مستوى الزمن

تابع جدول (4)

تجميع البيانات ربع سنوياً			تجميع البيانات شهرياً			نوع التجميع		الشركات
المجموع الكلي	الوسيط	الوسط الحسابي	المجموع الكلي	الوسيط	الوسط الحسابي	أسلوب التجميع	نموذج الانحدار الخطي المقدر	
						النموذج التفصيلي		
0.5910	0.0337	0.0107	0.0984	0.0344	0.0268	0.0088	قيمة المقدر (β_0)	طلعت مصطفى
0.3992	0.0135	0.0063	0.0865	0.0105	0.0056	0.0029	الخطأ المعياري للمقدر (β_0)	
0.5512	0.6354	0.5513	0.5387	0.6404	0.6774	0.5489	قيمة المقدر (β_1)	
0.0288	0.0482	0.0284	0.0174	0.0214	0.0130	0.0061	الخطأ المعياري للمقدر (β_1)	
0.4489	0.3629	0.4487	0.4616	0.3584	0.3247	0.4514	قيمة المقدر (β_2)	
0.0282	0.0473	0.0277	0.0171	0.0207	0.0126	0.0059	الخطأ المعياري للمقدر (β_2)	
0.3802	-0.0002	0.0007	0.1223	-0.0045	0.0081	0.0076	قيمة المقدر (β_0)	مصر الجديدة
0.3944	0.0156	0.0072	0.0959	0.0140	0.0055	0.0059	الخطأ المعياري للمقدر (β_0)	
0.4929	0.5229	0.4956	0.4879	0.6518	0.4929	0.4876	قيمة المقدر (β_1)	
0.0117	0.0259	0.0121	0.0059	0.0187	0.0067	0.0043	الخطأ المعياري للمقدر (β_1)	
0.5062	0.4772	0.5036	0.5110	0.3533	0.5062	0.5113	قيمة المقدر (β_2)	
0.0112	0.0248	0.0116	0.0056	0.0179	0.0063	0.0041	الخطأ المعياري للمقدر (β_2)	
-0.2024	-0.0049	-0.0056	-0.0683	0.0068	-0.0055	-0.0071	قيمة المقدر (β_0)	السادس من أكتوبر
0.3297	0.0134	0.0064	0.0671	0.0057	0.0038	0.0025	الخطأ المعياري للمقدر (β_0)	
0.4917	0.5634	0.4954	0.4891	0.5246	0.4904	0.5101	قيمة المقدر (β_1)	
0.0334	0.0395	0.0359	0.0166	0.0192	0.0170	0.0060	الخطأ المعياري للمقدر (β_1)	
0.5086	0.4397	0.5052	0.5111	0.4749	0.5101	0.4910	قيمة المقدر (β_2)	
0.5086	0.0379	0.0346	0.0160	0.0185	0.0164	0.0058	الخطأ المعياري للمقدر (β_2)	
0.1024	-0.0198	0.0020	0.0033	-0.0044	-0.0009	-0.0041	قيمة المقدر (β_0)	هيرمس القابضة
0.1645	0.0209	0.0038	0.0565	0.0116	0.0036	0.0034	الخطأ المعياري للمقدر (β_0)	
0.4919	0.5187	0.4929	0.4985	0.5159	0.4998	0.5095	قيمة المقدر (β_1)	
0.0057	0.0335	0.0066	0.0054	0.0177	0.0058	0.0039	الخطأ المعياري للمقدر (β_1)	
0.5076	0.4833	0.5067	0.5018	0.4842	0.5003	0.4911	قيمة المقدر (β_2)	
0.0055	0.03184	0.0062	0.0051	0.0168	0.0055	0.0037	الخطأ المعياري للمقدر (β_2)	
0.0736	-0.0048	0.0008	0.0361	0.0112	0.0015	0.0006	قيمة المقدر (β_0)	نصر للاسكان
0.3804	0.0101	0.0063	0.1273	0.0072	0.0063	0.0058	الخطأ المعياري للمقدر (β_0)	
0.5102	0.6126	0.5103	0.4901	0.3761	0.4979	0.5022	قيمة المقدر (β_1)	
0.0176	0.0201	0.0171	0.0144	0.0136	0.0141	0.0085	الخطأ المعياري للمقدر (β_1)	
0.4895	0.3916	0.4895	0.5087	0.6187	0.5012	0.4972	قيمة المقدر (β_2)	
0.0168	0.0191	0.0163	0.0138	0.0130	0.0134	0.0081	الخطأ المعياري للمقدر (β_2)	
0.0227	0.0169	-0.0099	0.0339	-0.0008	-0.0006	0.0075	قيمة المقدر (β_0)	أجواء للصناعات
0.2562	0.0272	0.0078	0.0871	0.0095	0.0071	0.0081	الخطأ المعياري للمقدر (β_0)	
0.6143	0.2662	0.6669	0.5730	0.6899	0.5981	0.5434	قيمة المقدر (β_1)	
0.0197	0.0665	0.0299	0.0144	0.0177	0.0186	0.0101	الخطأ المعياري للمقدر (β_1)	
0.3907	0.7203	0.3437	0.4290	0.3195	0.4067	0.4559	قيمة المقدر (β_2)	
0.0182	0.0605	0.0276	0.0133	0.0164	0.0172	0.0092	الخطأ المعياري للمقدر (β_2)	

الخطأ المعياري لمقدرات معالم نموذج الانحدار الخطي وتأثره بالتجميع على مستوى الزمن

جامعة القاهرة - معهد الدراسات والبحوث الإحصائية

تجميع البيانات ربع سنوياً			تجميع البيانات شهرياً			نوع التجميع		الشركات
المجموع الكلي	الوسيط	الوسط الحسابي	المجموع الكلي	الوسيط	الوسط الحسابي	النموذج التفصيلي	نموذج الانحدار الخطى المقدر	
0.587	0.0119	0.0155	0.3575	0.0114	0.0183	0.0152	قيمة المقدر ($\hat{\beta}_0$)	النساجون الشرقيون
1.155	0.0066	0.0223	0.3469	0.0064	0.0174	0.0146	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
0.3353	0.5242	0.3356	0.3079	0.4647	0.3375	0.3251	قيمة المقدر ($\hat{\beta}_1$)	
0.1152	0.0402	0.1216	0.0886	0.0243	0.0821	0.0373	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
0.6547	0.4749	0.6533	0.6798	0.5329	0.6515	0.6637	قيمة المقدر ($\hat{\beta}_2$)	
0.1106s	0.0389	0.1166	0.0849	0.0232	0.0787	0.0357	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	
0.0479	-0.0001	0.0007	0.0217	0.0021	0.0011	0.0014	قيمة المقدر ($\hat{\beta}_0$)	جلوبال تيلكوم القابضة
0.0449	0.0029	0.0009	0.0199	0.0025	0.0012	0.0009	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
0.5239	0.5095	0.5136	0.5429	0.5253	0.5106	0.5492	قيمة المقدر ($\hat{\beta}_1$)	
0.0121	0.0356	0.0129	0.0125	0.0254	0.0132	0.0056	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
0.4763	0.4912	0.4864	0.4578	0.4743	0.4892	0.4516	قيمة المقدر ($\hat{\beta}_2$)	
0.0118	0.0347	0.0125	0.0122	0.0247	0.0129	0.0055	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	
-0.821	-0.0029	-0.0257	-0.4680	-0.0306	-0.0115	-0.0045	قيمة المقدر ($\hat{\beta}_0$)	المصرية للاتصالات
1.101	0.0469	0.0258	0.3411	0.0475	0.0273	0.0152	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
0.5359	0.5985	0.5494	0.5297	0.5466	0.5415	0.5619	قيمة المقدر ($\hat{\beta}_1$)	
0.0219	0.0455	0.0225	0.0156	0.0252	0.0190	0.0063	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
0.4659	0.4042	0.4537	0.4726	0.4564	0.4604	0.4403	قيمة المقدر ($\hat{\beta}_2$)	
0.0213	0.0446	0.0217	0.0152	0.0242	0.0183	0.0061	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	
-0.2686	0.0051	-0.0018	-0.0994	-0.0065	-0.0001	-0.0061	قيمة المقدر ($\hat{\beta}_0$)	البنك التجاري الدولي
0.2036	0.0111	0.0039	0.0551	0.0058	0.0032	0.0024	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
0.5641	0.6336	0.5980	0.5489	0.5242	0.6251	0.4962	قيمة المقدر ($\hat{\beta}_1$)	
0.0201	0.0425	0.0218	0.0133	0.0164	0.0113	0.0052	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
0.4382	0.3704	0.4049	0.4529	0.4772	0.3785	0.5044	قيمة المقدر ($\hat{\beta}_2$)	
0.0196	0.0415	0.0212	0.0129	0.0160	0.0110	0.0051	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	
0.3944	0.0156	0.0072	0.0959	0.0140	0.0055	0.0059	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
0.4929	0.5229	0.4956	0.4879	0.6518	0.4929	0.4876	قيمة المقدر ($\hat{\beta}_1$)	
0.0117	0.0259	0.0121	0.0059	0.0187	0.0067	0.0043	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
0.5062	0.4772	0.5036	0.5110	0.3533	0.5062	0.5113	قيمة المقدر ($\hat{\beta}_2$)	
0.0112	0.0248	0.0116	0.0056	0.0179	0.0063	0.0041	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	

ويتضح من الجدول السابق زيادة الخطأ المعياري لمقدرات معالم النموذج التجميعي إذا تم التجميع بأسلوب المجموع الكلي أو أحد مقاييس النزعة المركزية كالوسط الحسابي والوسيط، وذلك بالنسبة لمقدرات معالم الحد الثابت ومقدرات معالم الميل. ولكن هناك إمكانية لانخفاض الخطأ المعياري لمقدرات معالم النموذج الخطي وذلك في حالة التجميع باستخدام الوسيط. وذلك كما حدث في شركة (النساجون الشرقيون للسجاد).

ثالثاً: القدرة التنبؤية لنموذج الانحدار الخطي

يتم تقييم القدرة التنبؤية لنموذج الانحدار الخطي من خلال اختبار مدى استقرار تقديرات معالم النموذج. والذي يعتمد على إضافة بيانات جديدة من خارج العينة محل الدراسة، وتقدير نموذج الانحدار الخطي قبل وبعد إضافة البيانات الجديدة. ثم مقارنة القيم الجديدة المقدرة للمتغير بالقيم

المقدرة التي تم الحصول عليها قبل إضافة البيانات الجديدة. ويتم الاستدلال عن الفرق بين متوسط القيم الفعلية للمتغير التابع بعد الحصول على بيانات جديدة وقبل الحصول على تلك البيانات $(\mu_{Y_2} - \mu_{Y_1})$. ويتم اختبار الفرق بين متوسطي مجتمعين مستقلين حيث يُصاغ فرض العدم كما يلي:

$$H_0: \mu_{Y_2} = \mu_{Y_1}$$

ويتضمن فرض العدم أن هناك استقراراً في تقديرات نموذج الانحدار الخطي. أي أن نموذج الانحدار الخطي له قدرة على التنبؤ حيث لا يوجد فرق بين متوسط القيم الفعلية للمتغير التابع قبل وبعد الحصول على البيانات الجديدة .

ولدراسة تأثير القدرة التنبؤية لنموذج الانحدار الخطي بمشكلة التجميع على مستوى الزمن، تم استخدام العشر شركات محل الاهتمام في الدراسة القياسية. و يوضح جدول (5) أثر التجميع على مستوى الزمن بتجميع البيانات على مستوى الشهر والتجميع الربع سنوياً لتلك البيانات وذلك باستخدام أساليب التجميع المختلفة. وباستخدام القيمة الإحصائية p-value اتخاذ القرار بقبول أو رفض فرض العدم محل الاختبار وحيث أن الاختبار ذو طرفين فإن القيمة الاحتمالية هي احتمال أن تكون إحصائية الاختبار المحسوبة أكبر من القيمة المطلقة للقيمة الجدولية المناظرة لها عند مستوى خطأ معين مسموح به وقيمهته 0.05.

جدول (5)

نتائج اختبار t لدراسة القدرة التنبؤية لنموذج الانحدار الخطي في ظل التجميع على مستوى الزمن

تجميع البيانات ربع سنوياً			تجميع البيانات شهرياً			نوع التجميع		
المجموع الكلي	الوسيط	الوسط الحسابي	المجموع الكلي	الوسيط	الوسط الحسابي	أسلوب التجميع	معايير الاختبار	الشركات
						النموذج التفصيلي		
26.9	0.326	0.320	7.52	0.380	0.371	0.2494	الفرق بين المتوسطين	طلعت مصطفى
0.68	0.47	0.46	0.90	0.93	0.91	2.79	قيمة "ت" المحسوبة	
0.498*	0.642*	0.648*	0.371*	0.353*	0.366*	0.005	القيمة الإحصائية	
75	1.59	1.57	28.4	1.33	1.33	1.540	الفرق بين المتوسطين	مصر الجديدة
0.37	0.49	0.49	0.72	0.71	0.71	3.21	قيمة "ت" المحسوبة	
0.711*	0.623*	0.625*	0.474*	0.479*	0.477*	0.001	القيمة الإحصائية	
2.2	0.11	0.11	2.4	0.107	0.106	0.104	الفرق بين المتوسطين	السادس من أكتوبر
0.03	0.11	0.11	0.18	0.18	0.18	0.74	قيمة "ت" المحسوبة	
0.973*	0.912*	0.915*	0.854*	0.860*	0.861*	0.458*	القيمة الإحصائية	
5.7	-0.04	-0.04	0.6	-0.003	-0.005	-0.016	الفرق بين المتوسطين	هيرمس القابضة
0.07	-0.03	-0.03	0.04	-0.00	-0.01	-0.09	قيمة "ت" المحسوبة	
0.947*	0.975*	0.975*	0.972*	0.996*	0.995*	0.929*	القيمة الإحصائية	
25	0.62	0.61	10.5	0.510	0.515	0.521	الفرق بين المتوسطين	نصر للإسكان
0.24	0.38	0.37	0.53	0.53	0.54	2.4	قيمة "ت" المحسوبة	
0.807*	0.704*	0.710*	0.597*	0.595*	0.592*	0.016	القيمة الإحصائية	
-2.6	-0.14	-0.14	-1.7	-0.139	-0.144	-0.335	الفرق بين المتوسطين	أجواء للصناعات
-0.04	-0.11	-0.11	-0.12	-0.19	-0.19	-1.57	قيمة "ت" المحسوبة	
0.972*	0.915*	0.912*	0.907*	0.853*	0.851*	0.117*	القيمة الإحصائية	
7.8	0.232	0.216	3.64	0.184	0.175	0.18	الفرق بين المتوسطين	النساجون الشرفيون
0.26	0.49	0.46	0.62	0.66	0.63	2.83	قيمة "ت" المحسوبة	
0.794*	0.627*	0.649*	0.538*	0.508*	0.528*	0.005	القيمة الإحصائية	
-0.1	-0.037	-0.039	-0.48	-0.024	-0.025	-0.024	الفرق بين المتوسطين	جلوبال تيلكوم القابضة
-0.00	-0.09	-0.09	-0.09	-0.10	-0.10	-0.42	قيمة "ت" المحسوبة	
0.997*	0.932*	0.930*	0.927*	0.924*	0.923*	0.675*	القيمة الإحصائية	

تابع جدول (5)

نتائج اختبار t لدراسة القدرة التنبؤية لنموذج الانحدار الخطى في ظل التجميع على مستوى الزمن

الشركات	نوع التجميع		تجميع البيانات شهرياً			تجميع البيانات ربع سنوياً		
	معايير الاختبار	أسلوب التجميع النموذج التفصيلي	المجموع الكلي	الوسيط	الوسط الحسابي	المجموع الكلي	الوسيط	الوسط الحسابي
المصرية للاتصالات			الفرق بين المتوسطين	-0.351	-0.357	-0.356	-7.09	-0.362
	قيمة "ت" المحسوبة	-4.59	-1.06	-1.05	-0.90	-0.63	-0.64	-0.29
	القيمة الإحتمالية	0.0001	0.291*	0.294*	0.371*	0.533*	0.523*	0.771*
البنك التجارى الدولى	الفرق بين المتوسطين	1.442	1.45	1.45	29.2	1.81	1.84	80
	قيمة "ت" المحسوبة	5.36	1.19	1.20	1.18	0.85	0.86	0.63
	القيمة الإحتمالية	0.0001	0.233*	0.232*	0.239*	0.399*	0.394*	0.529*

من الجدول السابق يمكن تلخيص بعض النتائج فيما يلى:

- إذا كان النموذج التفصيلي ليس لديه قدرة على التنبؤ، ويوجد عدم استقرار فى تقديرات معالم نموذج الانحدار كما فى بيانات الشركات الآتية: طلعت مصطفى القابضة، ومصر الجديدة للإسكان، ونصر للإسكان والتعمير، والنساجون الشرقيون للسجاد، والمصرية للاتصالات، والبنك التجارى الدولى. فإن القدرة التنبؤية للنموذج التجميعي بجميع أساليبه المجموع الكلي أو أحد مقاييس النزعة المركزية تتحسن.
- إذا كان النموذج التفصيلي لديه القدرة على التنبؤ، ويوجد استقرار فى تقديرات معالم نموذج الانحدار. فإن التجميع على مستوى الزمن يؤدي إلى زيادة القدرة التنبؤية للنموذج التجميعي بجميع أساليبه وذلك كما فى الشركات التالية: شركة جلوبال القابضة، وشركة أجواء للصناعات الغذائية، وشركة هيرمس القابضة، وشركة السادس من أكتوبر.

رابعاً: دقة تنبؤ نموذج الانحدار الخطى

لدراسة دقة تنبؤ نموذج الانحدار الخطى التجميعي تم قياس جذر متوسط مربعات الخطأ. وذلك فى نموذج الانحدار الخطى المتعدد، وعلى المستويين التفصيلي والتجميعي. ذلك النموذج التجميعي الذى يعتمد على الوسط الحسابي والوسيط والمجموع الكلي كأساليب للتجميع. ويتم حساب قيمة جذر متوسط مربعات الخطأ كما بالمعادلة التى يمكن صياغتها بالصورة التالية:

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_{it} - \hat{y}_{it})^2} \quad (15)$$

وبالتجميع على مستوى الزمن "شهرياً وربع سنوياً" لبيانات يومية، وذلك فى نموذج الانحدار المتعدد كما بجدول (6) التالى.

جدول (6)

جذر متوسط مربعات الخطأ RMSE لنموذج الانحدار الخطي المتعدد في ظل التجميع على مستوى الزمن

تجميع البيانات ربع سنوياً			تجميع البيانات شهرياً			نوع التجميع	
المجموع الكلي	الوسيط	الوسط الحسابي	المجموع الكلي	الوسيط	الوسط الحسابي	النموذج التفصيلي	أسلوب التجميع الشركات
0.6664	0.0249	0.0114	0.2857	0.0373	0.0198	0.0462	طلعت مصطفى القابضة
4.2619	0.0235	0.0767	2.2698	0.0393	0.1063	0.3928	النساجون الشريون للسجاد
0.1859	0.0120	0.0035	0.1425	0.0175	0.0083	0.0309	جلوبال تيلكوم القابضة
1.3925	0.0576	0.0247	0.5234	0.0429	0.0274	0.0838	السادس من أكتوبر
2.0438	0.0532	0.0332	1.1946	0.0671	0.0579	0.2413	نصر للتعمير والإسكان
0.6846	0.0766	0.0135	0.4163	0.0747	0.0226	0.0990	هيرمس القابضة
1.1786	0.1120	0.0327	0.6959	0.0718	0.0534	0.1998	أجواء للصناعات الغذائية
1.0867	0.0599	0.0209	0.5160	0.0551	0.0302	0.1028	البنك التجاري الدولي

ويتضح من الجدول السابق أن التجميع على مستوى الزمن وباستخدام المجموع الكلي كأسلوب للتجميع يؤدي إلى انخفاض دقة تنبؤ النموذج. أما باستخدام مقاييس النزعة المركزية في تجميع البيانات على مستوى الزمن فإن دقة تنبؤ النموذج التجميعي تزداد.

(2-2-4) أثر مشكلة التجميع على مستوى المؤسسات

أولاً: القوة التفسيرية للنموذج

للتعرف على أثر التجميع على مستوى المؤسسات على القوة التفسيرية في نموذج الانحدار الخطي المتعدد، تم الإستعانة ببيانات 14 قطاعاً إقتصادياً مختلفاً، وقياس قيم معامل التحديد قبل وبعد تجميع البيانات على مستوى المؤسسات "الشركات والبنوك". ويُخلص جدول (7) أهم تلك النتائج .

جدول (7)

معامل تحديد نموذج الانحدار الخطي المتعدد "R²" في ظل التجميع على مستوى المؤسسات

تجميع البيانات على مستوى المؤسسات			نوع التجميع	
المجموع الكلي	الوسيط	الوسط الحسابي	النموذج التفصيلي	أسلوب التجميع القطاعات
97.1%	91%	96.4%	95.8%	البنوك
85.4%	99.4%	85.9%	93.6%	الموارد الأساسية
83.9%	97.7%	83.3%	97.5%	الاتصالات
97.3%	59.5%	59.2%	91.3%	الأغذية والمشروبات
90.5%	59.4%	59.4%	94.1%	التشيد والبناء
99.9%	80.8%	95.8%	99.8%	الخدمات المالية
99.1%	97%	99.2%	98.4%	السياحة والترفيه
98.9%	96.7%	98.9%	98.2%	العقارات
97.2%	98.8%	98.8%	99.6%	الغاز والبتروك
64.8%	81.2%	75.5%	89.3%	الكيموايات
96.3%	98.8%	96.6%	97.0%	خدمات ومنتجات صناعية
92.1%	78.5%	78%	83.2%	الرعاية الصحية والأدوية
92.0%	98.0%	91.9%	97.2%	منتجات منزلية وشخصية
85.9%	84%	84%	77.4%	الموزعين وتجار تجزئة

ويتضح من الجدول السابق مدى تأثير القوة التفسيرية لنموذج الانحدار الخطي بمشكلة التجميع على مستوى المؤسسات وفيما يلي استعراض لأهم تلك النتائج:

- بتطبيق التجميع على مستوى المؤسسات باستخدام المجموع الكلى، فإن قيم معامل التحديد تتفاوت ما بين الزيادة والنقصان دون الوصول إلى أثر محدد. حيث يزداد معامل التحديد فى القطاعات الآتية: قطاع البنوك، وقطاع الأغذية والمشروبات، وقطاع الخدمات المالية باستثناء البنوك، وقطاع العقارات، وقطاع الرعاية الصحية والأدوية، وقطاع الموزعين وتجار التجزئة.
- وباستخدام أحد مقاييس النزعة المركزية كالوسط الحسابى أو الوسيط فى تطبيق ذلك النوع من التجميع، فإن قيم معامل التحديد تتخفف فى النموذج التجميعى وذلك فى جميع القطاعات فيما عدا:
 - قطاع البنوك وقطاع السياحة والترفيه وقطاع العقارات وقطاع الموزعين وتجار التجزئة وذلك فى حالة استخدام الوسط الحسابى.
 - أما باستخدام الوسيط فى تجميع البيانات على مستوى المؤسسات فإن قيم معامل التحديد تتخفف فى النموذج التجميعى وذلك فى جميع القطاعات فيما عدا: قطاع الموارد الأساسية وقطاع الاتصالات وقطاع الخدمات والمنتجات الصناعية وقطاع المنتجات المنزلية وقطاع الموزعين وتجار التجزئة.

ثانياً: كفاءة مقدرات معالم النموذج

لتلخيص والتعرف على مدى تأثير الخطأ المعيارى لمقدرات معالم النموذج بمشكلة التجميع على مستوى المؤسسات. فقد تم تجميع بيانات عدد 14 قطاعاً إقتصادياً فى نموذج الانحدار الخطى المتعدد. وجدول (8) التالى يلخص ذلك الأثر فى نموذج الانحدار الخطى.

جدول (8)

الخطأ المعيارى لمقدرات نموذج الانحدار الخطى المتعدد فى ظل التجميع على مستوى المؤسسات

تجميع البيانات شهرياً			نوع التجميع		
المجموع الكلى	الوسيط	الوسط الحسابى	أسلوب التجميع		القطاعات
			النموذج التفصيلى	نموذج الانحدار الخطى المقدر	
-1.0310	0.0134	-0.0162	-0.0627	قيمة المقدر ($\hat{\beta}_0$)	البنوك
0.7327	0.0878	0.0703	0.0279	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
-0.3489	-0.2603	-0.4029	0.0953	قيمة المقدر ($\hat{\beta}_1$)	
0.0762	0.0914	0.0793	0.0279	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
1.3097	1.2017	1.3569	0.8782	قيمة المقدر ($\hat{\beta}_2$)	
0.0736	0.0879	0.0767	0.0271	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	
-104.21	-0.4108	-15.490	-3.0187	قيمة المقدر ($\hat{\beta}_0$)	الموارد الاساسية
9.256	0.0448	1.349	0.6678	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
-0.7715	0.3648	-0.8543	-1.2066	قيمة المقدر ($\hat{\beta}_1$)	
0.1903	0.0209	0.1927	0.0803	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
1.7890	0.6329	1.8694	2.0144	قيمة المقدر ($\hat{\beta}_2$)	
0.1837	0.0194	0.1864	0.0783	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	
-11.66	-0.2564	-3.178	-0.0171	قيمة المقدر ($\hat{\beta}_0$)	الاتصالات
1.891	0.0463	0.4927	0.1468	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
0.1983	0.4492	0.1943	0.2259	قيمة المقدر ($\hat{\beta}_1$)	
0.1181	0.1157	0.1181	0.0618	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
0.8384	0.5560	0.8472	0.7498	قيمة المقدر ($\hat{\beta}_2$)	
0.1156	0.1120	0.1157	0.0601	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	

جدول (8)

الخطأ المعيارى لمقدرات نموذج الانحدار الخطى المتعدد فى ظل التجميع على مستوى المؤسسات

تجميع البيانات شهرياً			نوع التجميع		
المجموع الكلى	الوسيط	الوسط الحسابى	أسلوب التجميع	نموذج الانحدار الخطى المقدر	القطاعات
			النموذج التفصيلى		
-14.206	3.335	4.207	0.5019	قيمة المقدر ($\hat{\beta}_0$)	الأغذية والمشروبات
1.401	0.2413	0.2708	0.0355	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
0.1789	-2.002	-2.632	-0.3015	قيمة المقدر ($\hat{\beta}_1$)	
0.0549	0.1689	0.1695	0.0225	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
0.8273	2.629	3.224	1.2139	قيمة المقدر ($\hat{\beta}_2$)	
0.0505	0.1647	0.1638	0.0214	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	
-34.77	2.2328	-0.7286	-0.1361	قيمة المقدر ($\hat{\beta}_0$)	التشيد والبناء
4.487	0.3005	0.6452	0.9182	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
-1.0113	-1.845	-2.740	-0.9183	قيمة المقدر ($\hat{\beta}_1$)	
0.1437	0.1799	0.221	0.0449	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
1.980	2.566	3.603	1.8341	قيمة المقدر ($\hat{\beta}_2$)	
0.1437	0.1764	0.2141	0.0437	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	
-0.5575	0.1372	0.0106	-0.0258	قيمة المقدر ($\hat{\beta}_0$)	الخدمات المالية
0.0528	0.0441	0.0375	0.0043	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
0.5178	-0.8124	-0.0627	0.4499	قيمة المقدر ($\hat{\beta}_1$)	
0.0164	0.1280	0.0794	0.007	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
0.4898	1.693	1.0206	0.5472	قيمة المقدر ($\hat{\beta}_2$)	
0.0152	0.1227	0.0733	0.0066	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	
-2.129	-0.5932	-0.4283	-0.2021	قيمة المقدر ($\hat{\beta}_0$)	السياحة والترفيه
0.1283	0.0408	0.0233	0.0172	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
0.1787	0.2409	0.1956	0.1809	قيمة المقدر ($\hat{\beta}_1$)	
0.0499	0.0870	0.0463	0.0233	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
0.8267	0.7792	0.8147	0.8041	قيمة المقدر ($\hat{\beta}_2$)	
0.0476	0.0836	0.0442	0.0221	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	
-3.2007	-0.0751	-0.2889	-0.0858	قيمة المقدر ($\hat{\beta}_0$)	العقارات
0.7225	0.0497	0.0369	0.0133	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
0.1541	0.2957	0.2089	0.2797	قيمة المقدر ($\hat{\beta}_1$)	
0.0329	0.0361	0.0309	0.0113	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
0.8263	0.6854	0.7836	0.7019	قيمة المقدر ($\hat{\beta}_2$)	
0.0309	0.0342	0.0290	0.0106	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	
0.0621	-0.0781	-0.0781	0.0064	قيمة المقدر ($\hat{\beta}_0$)	الغاز والبتروك
0.2410	0.0904	0.0904	0.0474	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
0.4261	0.3523	0.3523	0.4062	قيمة المقدر ($\hat{\beta}_1$)	
0.0577	0.0443	0.0443	0.0417	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
0.5693	0.6452	0.6452	0.5897	قيمة المقدر ($\hat{\beta}_2$)	
0.0573	0.0435	0.0435	0.0409	الخطأ المعيارى للمقدر ($\hat{\beta}_2$)	
13.233	1.4810	3.1687	1.0453	قيمة المقدر ($\hat{\beta}_0$)	الكيمواويات
2.629	0.3604	0.5225	0.1550	الخطأ المعيارى للمقدر ($\hat{\beta}_0$)	
-1.2161	-1.1044	-1.3086	-0.9299	قيمة المقدر ($\hat{\beta}_1$)	
0.1870	0.3656	0.3403	0.0929	الخطأ المعيارى للمقدر ($\hat{\beta}_1$)	
2.055	1.9286	2.1193	1.8187	قيمة المقدر ($\hat{\beta}_2$)	

تابع جدول (8)

الخطأ المعيارى لمقدرات نموذج الانحدار الخطى المتعدد فى ظل التجميع على مستوى المؤسسات

جامعة القاهرة - معهد الدراسات والبحوث الإحصائية

0.1817	0.3625	0.3378	0.0913	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	الخدمات والمنتجات الصناعية
-0.529	-0.1107	-0.1068	0.2224	قيمة المقدر ($\hat{\beta}_0$)	
1.129	0.0271	0.1060	0.0486	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
-0.1411	0.5144	-0.0654	-0.1666	قيمة المقدر ($\hat{\beta}_1$)	
0.0721	0.0217	0.0689	0.0253	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
1.0957	0.4883	1.0292	1.0596	قيمة المقدر ($\hat{\beta}_2$)	
0.0672	0.0199	0.0641	0.0239	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	الرعاية الصحية والأدوية
-8.392	-1.0033	-2.1419	-0.7691	قيمة المقدر ($\hat{\beta}_0$)	
1.198	0.3503	0.4124	0.1611	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
-0.1310	-0.1586	-0.7054	-0.0124	قيمة المقدر ($\hat{\beta}_1$)	
0.1067	0.1028	0.1407	0.0497	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
1.1197	1.1256	1.6858	0.9766	قيمة المقدر ($\hat{\beta}_2$)	
0.1016	0.1028	0.1359	0.0483	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	المنتجات المنزلية
-3.017	-0.0619	-0.4646	-0.0594	قيمة المقدر ($\hat{\beta}_0$)	
1.033	0.0160	0.1507	0.0730	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
0.0493	0.4685	0.0485	0.0510	قيمة المقدر ($\hat{\beta}_1$)	
0.0935	0.0394	0.0927	0.0368	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
0.9388	0.5356	0.9410	0.9186	قيمة المقدر ($\hat{\beta}_2$)	
0.0911	0.0377	0.0903	0.0359	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	الموزعين وتجارة التجزئة
0.1695	0.0159	0.0159	0.1224	قيمة المقدر ($\hat{\beta}_0$)	
0.1035	0.0621	0.0621	0.0580	الخطأ المعياري للمقدر ($\hat{\beta}_0$)	
-0.7981	-0.6918	-0.6918	1.0948	قيمة المقدر ($\hat{\beta}_1$)	
0.1609	0.1441	0.1441	0.1178	الخطأ المعياري للمقدر ($\hat{\beta}_1$)	
1.7077	1.6199	1.6199	1.9908	قيمة المقدر ($\hat{\beta}_2$)	
0.1573	0.1403	0.1403	0.1139	الخطأ المعياري للمقدر ($\hat{\beta}_2$)	

ومن الجدول السابق تُعرض أهم النتائج كالتالي:

- يزداد الخطأ المعياري لمقدرات معالم النموذج التجميعي إذا تم التجميع بأسلوب المجموع الكلي أو الوسط الحسابي، وذلك بالنسبة لمقدرات معالم الحد الثابت و مقدرات معالم الميل.
- تتفاوت قيم الخطأ المعياري لمقدرات معالم نموذج الانحدار الخطي التجميعي الذي يعتمد على الوسيط في تكوين متغيراته، وذلك حيث يزداد الخطأ المعياري في 12 قطاعاً إقتصادياً ويقل في قطاع الموارد الأساسية، وقطاع الخدمات والمنتجات الصناعية.

ثالثاً: القدرة التنبؤية لنموذج الانحدار الخطي

يُلخص جدول (9) كيفية تأثير القدرة التنبؤية لنموذج الانحدار الخطي المقدر بمشكلة تجميع البيانات على مستوى المؤسسات، وذلك باستخدام بيانات يومية عن الفترة من أول يناير 2007 م إلى آخر يناير 2015م لعدد 14 قطاعاً إقتصادياً. وباختبار مدى استقرار التقديرات والذي يعتمد على إضافة بيانات جديدة من أول فبراير 2015م إلى آخر أغسطس 2015م. ثم اختبار الفرق بين متوسط القيم الفعلية بعد زيادة البيانات الجديدة μ_{Y_2} ، ومتوسط القيم الفعلية قبل زيادة البيانات الجديدة μ_{Y_1} باستخدام اختبار t للفرق بين متوسطي مجتمعين مستقلين وذلك بنسبة خطأ مسموح بها تساوي 0.05.

جدول (9)

اختبار t لدراسة القدرة التنبؤية لنموذج الانحدار الخطي في ظل لتجميع على مستوى الزمن

جامعة القاهرة - معهد الدراسات والبحوث الإحصائية

تجميع على مستوى المؤسسات			نوع التجميع		
المجموع الكلي	الوسيط	الوسط الحسابي	النموذج التفصيلي	معايير الاختبار	أسلوب التجميع القطاعات
1.19	0.002	0.191	0.174	الفرق بين المتوسطين	البنوك
0.63	0.01	1.25	1.43	قيمة "ت" المحسوبة	
0.528*	0.99*	0.211*	0.152* ³	القيمة الإحصائية	
-41.3	-0.797	-6.10	-5.64	الفرق بين المتوسطين	الموارد الأساسية
-2.83	-1.91	-2.75	-1.69	قيمة "ت" المحسوبة	
0.005	0.056*	0.006	0.09*	القيمة الإحصائية	
-3.08	-0.335	-0.811	-0.81	الفرق بين المتوسطين	الاتصالات
-1.84	-1.5	-1.97	-0.75	قيمة "ت" المحسوبة	
0.065*	0.133*	0.049*	0.452*	القيمة الإحصائية	
8.10	-0.894	-0.655	-0.386	الفرق بين المتوسطين	الأغذية و المشروبات
2.27	-4.44	-3.31	-3.7	قيمة "ت" المحسوبة	
0.023	0.0001	0.001	0.0001	القيمة الإحصائية	
10.26	-1.246	-0.706	-0.820	الفرق بين المتوسطين	التشيد والبناء
1.70	-5.76	-2.45	-1.95	قيمة "ت" المحسوبة	
0.090*	0.0001	0.014	0.051*	القيمة الإحصائية	
1.361	-0.188	-0.107	-0.143	الفرق بين المتوسطين	الخدمات المالية باستثناء البنوك
1.50	-3.62	-0.97	-1.44	قيمة "ت" المحسوبة	
0.133*	0.0001	0.333*	0.151*	القيمة الإحصائية	
0.807	-4.193	-2.976	-0.371	الفرق بين المتوسطين	السياحة والترفيه
0.84	-31.08	-19.36	-2.67	قيمة "ت" المحسوبة	
0.400*	0.0001	0.0001	0.007	القيمة الإحصائية	
-2.01	-0.071	-0.110	-0.113	الفرق بين المتوسطين	العقارات
-0.76	-0.74	-0.89	-1.22	قيمة "ت" المحسوبة	
0.448*	0.462*	0.371*	0.221*	القيمة الإحصائية	
-0.782	-0.647	-0.647	-0.577	الفرق بين المتوسطين	الغاز والبتترول
-1.59	-1.74	-1.74	-0.86	قيمة "ت" المحسوبة	
0.112*	0.082*	0.082*	0.391*	القيمة الإحصائية	
-0.95	-6.121	-5.352	-1.177	الفرق بين المتوسطين	الكيمويات
-0.58	-9.46	-8.54	-2.31	قيمة "ت" المحسوبة	
0.562*	0.0001	0.0001	0.021	القيمة الإحصائية	
4.55	-0.223	0.349	0.375	الفرق بين المتوسطين	الخدمات والمنتجات الصناعية
1.42	-1.96	1.11	1.09	قيمة "ت" المحسوبة	
0.154*	0.050*	0.265*	0.275*	القيمة الإحصائية	
2.61	-1.220	-0.805	-0.607	الفرق بين المتوسطين	الرعاية الصحية والأدوية
1.26	-3.78	-2.63	-2.45	قيمة "ت" المحسوبة	
0.207*	0.0001	0.009	0.014	القيمة الإحصائية	
4.52	-0.100	0.627	0.637	الفرق بين المتوسطين	المنتجات المنزلية والشخصية
2.00	-2.01	1.94	1.13	قيمة "ت" المحسوبة	
0.045	0.044	0.053*	0.258*	القيمة الإحصائية	
0.054	0.0562	0.0562	0.0751	الفرق بين المتوسطين	الموزعين وتجارة التجزئة
0.40	0.87	0.87	1.21	قيمة "ت" المحسوبة	
0.692*	0.382*	0.382*	0.226*	القيمة الإحصائية	

ومن الجدول السابق يمكن الحصول على النتائج التالية:

³ يأخذ القرار بقبول أو رفض فرض العدم وفقاً للقيمة الاحتمالية p-value وذلك بنسبة خطأ مسموح به مقداره 0.05. كما يعبر الرمز "*" على قبول فرض العدم وتحقق استقرار تقديرات معالم نموذج الانحدار.

- إذا كان النموذج التفصيلي له قدرة تنبؤية فإن التجميع على مستوى المؤسسات يحافظ على تلك القدرة التنبؤية، واستقرار تقديرات معالم النموذج الخطي في بعض القطاعات. وذلك كما في قطاع البنوك، وقطاع الاتصالات، وقطاع الخدمات المالية باستثناء البنوك، وقطاع العقارات، وقطاع الغاز والبترو، وقطاع الخدمات والمنتجات الصناعية، وقطاع الموزعين وتجار التجزئة.
- بينما تقل القدرة التنبؤية بسبب مشكلة تجميع البيانات على مستوى المؤسسات وذلك كما في قطاع الموارد الأساسية، وقطاع التشيد والبناء، وقطاع المنتجات المنزلية والشخصية.
- إذا كان النموذج التفصيلي ليس لديه القدرة على التنبؤ فإن تجميع البيانات على مستوى المؤسسات لا يبعكس استقرار تقديرات معالم النموذج وانعكاسها على قدرته التنبؤية. وذلك كما في القطاعات الآتية: قطاع الأغذية والمشروبات، وقطاع السياحة والترفيه، وقطاع الكيماويات، وقطاع الرعاية الصحية والأدوية. وإذا حدث تحسن في قدرة النموذج التجميعي على التنبؤ فإنه ذلك لا يحدث إلا في ظل التجميع باستخدام المجموع الكلي.

رابعاً: دقة تنبؤ نموذج الانحدار الخطي

وبالتجميع على مستوى المؤسسات، فقد تم تلخيص أهم نتائج جذر متوسط مربعات الخطأ في نموذج الانحدار الخطي المتعدد كما في جدول (10) التالي.

جدول (10)

جذر متوسط مربعات الخطأ RMSE لنموذج الانحدار الخطي المتعدد في ظل التجميع على مستوى المؤسسات

تجميع البيانات على مستوى المؤسسات			نوع التجميع	
المجموع الكلي	الوسيط	الوسط الحسابي	النموذج التفصيلي	أسلوب التجميع
				القطاعات
10.201	1.2126	0.91916	2.6259	البنوك
174.76	1.0682	26.060	69.527	الموارد الأساسية
21.035	1.0884	5.2935	10.713	الاتصالات
20.085	2.0486	4.7449	4.1563	الأغذية والمشروبات
63.467	4.8669	5.9622	13.586	التشيد والبناء
1.1439	0.7647	0.7285	0.4255	الخدمات المالية
2.9931	0.9929	0.5277	1.2784	السياحة والترفيه
8.922	0.56212	0.4205	1.7429	العقارات
2.6228	1.3172	1.3172	1.9225	الغاز والبترو
30.282	12.543	13.364	12.828	الكيماويات
19.298	0.39349	1.8223	5.8647	الخدمات الصناعية
19.019	5.1700	5.0237	7.3069	الرعاية الصحية والأدوية
19.876	0.22639	2.8636	7.6278	منتجات منزلية وشخصية
1.6005	0.80607	0.80607	1.2099	الموزعين وتجار التجزئة

ومن الجدول السابق يمكن استنتاج الآتي:

- يؤدي التجميع على مستوى المؤسسات إلى زيادة قيم RMSE في النموذج التجميعي، والذي يعتمد على المجموع الكلي في تكوين المتغيرات التجميعية. مما يؤكد على انخفاض دقة التنبؤ للنموذج بسبب التجميع على مستوى المؤسسات.

- إذا تم استخدام أحد مقاييس النزعة المركزية للتجميع على مستوى المؤسسات كالوسط الحسابي أو الوسيط، فإن قيمة جذر متوسط مربعات الخطأ تقل في النموذج التجميعي بالمقارنة بقيمته في النموذج التفصيلي مما يدل على ارتفاع دقة تنبؤ النموذج بالتجميع.

وذلك كما في جميع القطاعات 14 محل الدراسة، ولكن فيما عدا قطاع الأغذية والمشروبات، وقطاع الخدمات المالية باستثناء البنوك، وقطاع الكيماويات حيث تقل دقة تنبؤ النموذج التجميعي الذي يعتمد على مقاييس النزعة المركزية مثل: الوسط الحسابي أو الوسيط.

5- الخلاصة

تؤدي مشكلة التجميع إلى العديد من الآثار على جودة توفيق نموذج الانحدار، وقد تم قياس ذلك من خلال القوة التفسيرية لنموذج الانحدار وقد تم قياس معامل التحديد كمقياس لها. كما تم قياس الخطأ المعياري لمقدرات معالم النموذج وذلك لدراسة أثر مشكلة التجميع على كفاءة تلك مقدرات. وللتعرف على مدى تأثير القدرة التنبؤية لنموذج الانحدار الخطي، تم استخدام اختبار مدى استقرار التقديرات لنموذج الانحدار الخطي. وأخيراً لقياس أثر مشكلة التجميع على دقة تنبؤات نموذج الانحدار الخطي، تم الاستعانة بجذر متوسط مربعات الخطأ كمقياس لها. وذلك بتجميع البيانات على مستوى الزمن ومستوى المؤسسات محل الدراسة وباستخدام المجموع الكلي ومقاييس النزعة المركزية مثل: الوسط الحسابي والوسيط كأساليب لتجميع البيانات. وفي ظل البيانات محل الدراسة تتلخص أهم نتائج الدراسة القياسية فيما يلي:

- يتوقف أثر مشكلة التجميع على القوة التفسيرية على نوع التجميع حيث تؤدي تلك المشكلة إلى زيادة قيم معامل التحديد إذا تم تجميع البيانات على مستوى الزمن. ولكن لم يتحدد أثر واضح للتجميع على قيم معامل التحديد إذا تم التجميع على مستوى الوحدات القطاعية.
- تؤدي مشكلة التجميع إلى انخفاض كفاءة مقدرات OLS بغض النظر عن أسلوب أو نوع التجميع.
- يتوقف أثر مشكلة التجميع على القدرة التنبؤية للنموذج على نوع التجميع حيث يحسن التجميع على مستوى الزمن من قدرة نموذج الانحدار الخطي على التنبؤ، بينما يضعف التجميع على مستوى المؤسسات من قدرة النموذج على التنبؤ.
- يتوقف أثر مشكلة التجميع على دقة التنبؤات على أسلوب التجميع حيث تنخفض دقة تنبؤات النموذج التجميعي الذي يعتمد على المجموع الكلي في التجميع وزيادة دقة التنبؤات التي تعتمد على أحد مقاييس النزعة المركزية في التجميع مثل الوسط الحسابي والوسيط.

6- المراجع

- [1] Garderen, K.J.V., Lee, K. and Pesaran, M.H. (2000). " Cross Section Aggregation of Non-Linear Models." *Journal of Econometrics*, 95, pp.285-331.
- [2] Gordevicius, J., Gamper, J. and Bohlen, M. (2012). "Parsimonious Temporal Aggregation." *The VLDB Journal*, 21, pp.309-332.
- [3] Hassler, U. (2014). " Persistence under Temporal Aggregation and differencing. " *Economics Letters*, 124, pp.318-322.

- [4] Kabacoff, R.I. (2011). "R in Action Data Analysis and Graphics with R." Shelter Island.
- [5] Lutero, G. (2010). " The Aggregation Problem in its Historical Perspective: A Summary overview ." Methods Development of quarterly national accounts. ISTAT, National Accounts Directorate.
- [6] Monteforte, L. (2006). "Aggregation Bias in Macro Models : Does it matter for the Euro area? " Economic Modelling , 24,pp.236-261.
- [7] Petkovic, A. and Veredas, D. (2009). " Aggregation of Linear Models for Panel Data." International Association For Research and Teaching, working paper no .12.
- [8] Plassmann, E. (2002). "Econometric Modelling of European Money Demand." Physica -Verlag Berlin Heidelberg New York, pp.81-88.
- [9] Rossana , R.J. and Seater , J.J. (1995). " Temporal Aggregation and Economic Time Series." Journal of Business & Economic Statistics , 13, pp. 441-451.
- [10] Theil, H. (1954). " Linear Aggregation of Economic relationships." Amsterdam: North Holland.
- [11] Yager,R.,R.(1998). " Fusion of ordinal Information using Weighted Median Aggregation." International Journal of Approximate Reasoning 18,pp.35-52.

اهمية قواعد وإجراءات السلامة والصحة المهنية علي مراحل إدارة الأزمات والكوارث

(دراسة تحليلية على إحدى شركات الإتصالات)

عصام عطية عبدالمنعم

الملخص

تعد إدارة السلامة والصحة المهنية وفي ظل العديد من الازمات والكوارث فى الاونة الاخيرة من المداخل العلمية الحديثة التي تمارسها المنظمات الانتاجية والخدمية على حد سواء لتوفير بيئة عمل آمنة من المخاطر المختلفة ورفع مستوى كفاءة وسائل الوقاية، ويهدف البحث الى تحليل العلاقة بين ادارة السلامة والصحة المهنية والأزمات والكوارث ودور إدارة السلامة والصحة المهنية في الحد من الأزمات والكوارث وكلما استطاعت الشركة المصرية للاتصالات الساعية الى الاستمرار والنمو في ظل العديد من الازمات والكوارث فى الاونة الاخيرة وفي ظل المنافسة المتزايدة من تطبيق قواعد وإجراءات السلامة والصحة المهنية كلما أدى ذلك الى مواجهة الازمات والكوارث التي قد تجابهها بالاسلوب الذي يمكنها من معالجتها وأيجاد السبل الكفيلة لتفاديها مستقبلاً عند تكرارها خاصة وان الظروف الاقتصادية التي تمر بها الشركة المصرية للاتصالات تتطلب المزيد من التحسين والانتباه الى مصادر الازمات والكوارث وادارتها بالاسلوب العلمي المناسب لتجاوزها دون اضرار ناهيك عن الخسائر المادية والمعنوية التي قد تتعرض لها الشركة جراء هذه الازمات . وقد توصلت الدراسة الى عدة استنتاجات اعتمدت عليها التوصيات المناسبة .

الكلمات الدالة

- عينة طبقية
- معاملات الارتباطات البسيطة
- معاملات الانحدار القياسية

طالب دكتوراة جودة فى معهد الدراسات والبحوث الاحصائية

مقدمة البحث

لقد أدى التطور التقني الذي شهده العالم إلى ظهور العديد من المخاطر والازمات التي ينبغي على الإنسان إدراكها وتجنب الوقوع في مسبباتها ، فأماكن العمل المتعددة والمختلفة من ورش ومصانع ومختبرات ومعامل تعتبر بيئات عمل تكثر فيها العديد من المخاطر المهنية التي من الممكن ان تتسبب في حدوث العديد من الازمات والكوارث في بيئة العمل. (موقع دليل السلامة والصحة المهنية – <http://www.education.gov.bh>)

لذلك فإن توفير بيئة عمل آمنة من مخاطر الصناعات المختلفة ورفع مستوى كفاءة وسائل الوقاية سيؤدي بلا شك إلى الحد من الإصابات والأمراض المهنية ، وحماية العاملين من الحوادث والتي قد تتسبب في حدوث العديد من الازمات والكوارث .

(المركز العالمي للصحة والسلامة المهنية - <http://www.salama-libya.org>)

وفي ظل التأثير السلبي الذي تخلفه الأمراض والإصابات المرتبطة بالعمل على صحة العمال وإنتاجيتهم وبالتالي على عائلاتهم والوضع الاقتصادي والرفاهية الاجتماعية في البلاد، ازدادت التوعية حول الصحة والسلامة المهنية والازمات والكوارث في العالم بأكمله. (خوان سومافيا ، 2005).

وما زالت الأمراض والحوادث المهنية أهم أسباب الإصابات والوفيات للعاملين في بيئة العمل ويمكن تفادي أغلب الحوادث والازمات والكوارث عبر "وقاية سليمة تدعمها ممارسات ملائمة للتفتيش والتقرير وترشدها اتفاقيات منظمة العمل الدولية". (منظمة العمل الدولية ، جنيف، 2007)

ومن هنا زاد الاهتمام بالسلامة والصحة المهنية وسنت القوانين والتشريعات وانهقدت المؤتمرات المحلية والدولية المتلاحقة الهادفة لحماية الإنسان في العمل مثل الاتفاقية الدولية (اتفاقية السلامة والصحة والمهنية) رقم (155) لعام 1981 والتي تهدف إلى الوقاية من الحوادث والإصابات الصحية الناجمة عن العمل والتقليل من المخاطر المرتبطة ببيئة العمل بالإضافة إلى قانون العمل المصري رقم (12) الصادر لسنة (2003). ومن أجل تمكين الشركات من تحقيق أهدافها والقيام بدورها لابد من توفير إجراءات السلامة في الشركات وإيجاد بيئة عمل خالية من المخاطر والازمات وإلى تعريف العاملين بقواعد السلامة بهدف الوقاية من مخاطر العمل وتحقيق أكبر قدر من الصحة المهنية وأقل قدر من الخسائر المادية والبشرية وهذا يساعد على زيادة الإنتاج ويخفض التكلفة (أبو عبدون، 2003).

لذلك فإن هذه الدراسة سوف تسلط الضوء على دور إدارة السلامة والصحة المهنية في الحد من الازمات والكوارث

مشكلة البحث

تعتبر المصرية للاتصالات شركة رائدة في مجال الاتصالات السلكية واللاسلكية ونقل المعلومات ذات السرعات الفائقة وهي اعرق شركات الاتصالات في افريقيا والشرق الاوسط .
إلا إن القصور في تطبيق اجراءات ومتطلبات السلامة والصحة المهنية في هذه الشركة وعدم مطابقتها بالمعايير الدولية الموضوعه من قبل منظمة العمل الدولية (ILO) وإدارة السلامة والصحة المهنية (OSHA) ومنظمة الصحة العالمية (WHO) يؤدي ذلك إلى وقوع بعض الحوادث والازمات وقد تشمل الشركة نفسها أو قد تمتد إلى خارجها (المديفر، 2005) لذا يجب أن يتوفر الجهد لرفع مستوى الوعي لدى كل من العاملين والادارة بأهمية السلامة والصحة المهنية في مواقع العمل بالشركة، كما لا بد من التركيز على مجال التدريب والاستفادة من برامج السلامة الشاملة (كوهن وآخرون، 1995) ،

مما تقدم فان مشكلة البحث تبرز من خلال ما يلي :-

- هل هناك ارتباط بين عناصر قواعد واجراءات السلامة والصحة المهنية مع عناصر مراحل إدارة الازمات والكوارث في الشركة المصرية للاتصالات ؟
- ترتيب عناصر قواعد واجراءات السلامة والصحة المهنية وفقا للعلاقة مع عناصر مراحل إدارة الازمات والكوارث في الشركة المصرية للاتصالات .

أهمية البحث

إنبثقت أهمية الدراسة من تحديد الآثار الإيجابية من تطبيق ادارة السلامة والصحة المهنية في الحد من الأزمات و الكوارث من خلال تطبيق منظومة ادارة السلامة والصحة المهنية بكفاءة و فاعلية مما تساعد في مقاومة الأزمات و الكوارث و تعجل من سرعة إحتوائها و إعادة الأمن و الإستقرار و إزالة آثارها السلبية و كذلك دراسة الأزمات و الكوارث السابقة بما يكفل فاعلية مواجهتها و إحباط آثارها مستقبلاً ،
وإن عملية تقدير القيمة الاقتصادية للعنصر البشري في أي وحدة إنتاجية يجب أن تنال اهتماماً كبيراً وأن يكون هناك أسلوب دقيق يجب إتباعه واتساع دائرة تطبيقه في جميع الشركات للخروج بصورة دقيقة عن قيمة الثروة البشرية ومدى العائد من استثماره في التدريب والتعليم والخبرة حيث توضح أهمية الدراسة فيما يلي :

- تساهم هذه الدراسة للباحث في إثراء معلوماته حول السلامة والصحة المهنية ، وكيفية تجنب المخاطر المهنية والازمات والوقاية منها .

- تسهم الدراسة في معرفة العاملين بإشترطات السلامة والصحة المهنية عند أدائهم لأعمالهم وبالتالي التقليل من حجم الخسائر سواء المادية أو البشرية في الشركة المصرية للاتصالات
- تساعد هذه الدراسة العاملين بالشركات في الأخذ بكافة احتياطات السلامة المهنية والقواعد والإجراءات الواجب إتباعها في بيئة العمل ، وذلك من أجل خلق جو آمن وبيئة خالية من المخاطر والازمات في جميع المجالات.
- توضح الدراسة مدى احتياج إدارة الشركة المصرية للاتصالات لنتائج هذه الدراسة والتي من شأنها القيام بتطوير قواعد وإجراءات السلامة والصحة المهنية واساليب إدارة الازمات والكوارث .
- تساعد نتائج هذه الدراسة المسؤولين عن الشركة المصرية للاتصالات في تفعيل قواعد وإجراءات السلامة والصحة المهنية عند قيامهم بعملهم في الشركة.

أهداف البحث

أن الهدف الرئيسى من هذا البحث هو ابراز اهمية دور تطبيق قواعد واجراءات السلامة والصحة المهنية بالشركة المصرية للاتصالات في الحد من الازمات والكوارث وذلك من خلال قياس المتغيرات المستقلة المتمثلة فى قواعد واجراءات السلامة والصحة المهنية ، والمتغيرات التابعة المتمثلة فى مراحل إدارة الازمة أو الكارثة ، وذلك ببعض العبارات التى تناسب كل متغير.

منهج البحث وخطواته

سوف يتم إستخدام المنهج الوصفى التحليلى عن طريق مدخل المسح بإستخدام الإستبانة كأداة لجمع البيانات بالإضافة الى اساليب اخرى وفيما يلى توضيحها .
جمع البيانات

سوف يعتمد اسلوب العمل فى أنجاز الجانب النظري من الدراسة الأسلوب الوصفى باستخدام المراجع العلمية والدوريات وأدبيات إدارة الإنتاج والعمليات فضلا عن شبكة الانترنت ومحتوياتها الحديثة. اما الجانب الميداني من الدراسة سوف يعتمد بشكل رئيسي على استمارات الاستبيان التي جاءت متوافقة ومستمدة من الجانب النظري من الدراسة بما يمكن تحديد قواعد واجراءات السلامة والصحة المهنية ومراحل إدارة الازمات والكوارث بالشركة المصرية للاتصالات اذ وزعت (275) مائتان وخمسة وسبعون استمارة على مسؤولي الوحدات ورؤساء الاقسام زيادة على ذلك المقابلات الشخصية والزيارات الميدانية للادارات ومواقع الشركة المختلفة بالمحافظات المختلفة . كما اعتمد اسلوب ليكرت الخماسى على تحليل الاستمارة وفضلاً عن الارتباط بين متغيرات الدراسة الى جانب

التكرارات والنسب المئوية والمتوسطات والانحرافات المعيارية ومعامل الاختلاف لوصف وتشخيص متغيرات الدراسة

مجتمع الدراسة :

تقوم الشركة المصرية للاتصالات بدور حيوى وهام فى خدمة المجتمع من خلال توفير خدمة الانترنت فائق السرعة بما ساهم فى سهولة تداول البيانات بما يرفع من الدخل القوى للبلاد .
وقد قام الباحث بسحب عينة طبقية من مختلف المستويات الوظيفية بالادارة العامة للسلامة والصحة المهنية وجميع الاقسام التابعة لها فى قطاعات الشركة المختلفة بالمحافظات وكذلك الادارة العامة لإدارة وتحليل الازمات وجميع الاقسام التابعة لها فى قطاعات الشركة المختلفة بالمحافظات ومبنى سنترال القبة وحيث أن مجتمع الدراسة المتمثل فى الادارة العامة للسلامة والصحة المهنية وجميع الاقسام التابعة لها فى قطاعات الشركة المصرية للاتصالات المختلفة بالمحافظات والبالغ عدد مفردات الدراسة من واقع سجلات شئون العاملين وهم (90) مفردة وكذلك الادارة العامة لإدارة وتحليل الازمات وجميع الاقسام التابعة لها فى قطاعات الشركة المصرية للاتصالات المختلفة بالمحافظات والبالغ عدد مفردات الدراسة من واقع سجلات شئون العاملين وهم (50) مفردة .
وكذلك إدارة مبنى سنترال القبة والبالغ عدد مفردات الدراسة من واقع سجلات شئون العاملين وهم (360) مفردة .
وبالتالى يكون اجمالى مفردات مجتمع الدراسة البالغ (500) مفردة مقسمة كما فى جدول رقم (1) .

جدول رقم (1)

توزيع مفردات المجتمع على الإدارات

البيان	العدد	النسبة من المجتمع (%)
إدارة مبنى سنترال القبة	360	72
الإدارة العامة للسلامة والصحة المهنية وجميع الاقسام التابعة لها فى الفروع	90	18
الادارة العامة لإدارة وتحليل الازمات وجميع الاقسام التابعة لها فى الفروع	50	10
المجموع	500	100

ويتضح من الجدول رقم (1) توزيع مفردات مجتمع الدراسة على الإدارات .

عينة الدراسة :

وتم سحب عينة طبقية تمثل (55%) من مجتمع الدراسة اى تمثل (275) مفردة منهم (60) مفردة من الادارة العامة للسلامة وجميع الاقسام التابعة لها ومنهم ايضا (35) مفردة من الادارة العامة لإدارة وتحليل الازمات وايضا (180) مفردة من العاملين بإدارة مبنى سنترال القبة كما فى جدول رقم (2) . وسوف يتم نشر (275) إستمارة إستقصاء صالحة للتحليل الاحصائى وذلك نظراً لضيق الوقت ولعدم إمكانية نشر (500) إستمارة وإعادة جمعهم وتحليلهم وإستخراج النتائج والتوصيات

جدول رقم (2)

توزيع مفردات العينة على الإدارات

البيان	العدد	النسبة من العينة (%)	النسبة من مفردات المجتمع (%)
إدارة مبنى سنترال القبة	180	65	50
الادارة العامة للسلامة والصحة المهنية وجميع الاقسام التابعة لها فى الفروع	60	22	67
الادارة العامة لإدارة وتحليل الازمات وجميع الاقسام التابعة لها فى الفروع	35	13	70
المجموع	275	100	-----

ويتضح من الجدول رقم (2) توزيع مفردات عينة الدراسة على الإدارات ونسبتها من مفردات المجتمع

متغيرات الدراسة

حيث تلاحظ ان متغيرات الدراسة متمثلة فى الاستبيان والذى يوضح مجموعتين من المتغيرات وهما متغيرات المجموعة الأولى (X) الخاصة بتطبيق قواعد واجراءات السلامة والصحة المهنية وتتكون من اربعة مؤشرات وهى كالتالى :

- التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية (X_1)
- الالتزام بتوفير وسائل السلامة والصحة المهنية في بيئة العمل (X_2)
- تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية (X_3)
- الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة والمهنية (X_4)

ومتغيرات المجموعة الثانية (Y) الخاصة بمراحل إدارة الازمات والكوارث وتتكون من خمسة مؤشرات وهى كالتالى:

- مرحلة اكتشاف إشارات الإنذار المبكر (y_1)
- مرحلة الاستعداد والوقاية (y_2)
- مرحلة احتواء الأضرار (y_3)
- مرحلة استعادة النشاط (y_4)
- مرحلة التعلم (y_5)

وفضلا عن ايضاحها في استمارة الاستبيان وكيفية قياس كل متغير فيها.

الاسلوب البحثي

أستخدم الأسلوب التحليلي في الدراسة حيث استخدم التحليل العاملي Factor Analysis لاستخلاص المتغيرات المستقلة والتابعة من عناصر صحيفة الاستبيان ثم تم حساب معاملات الارتباط البسيطة في شكل مصفوفة الارتباطات لتحديد شدة الارتباط بين المتغيرات المستقلة (قواعد واجراءات السلامة والصحة المهنية) والمتغيرات التابعة (مراحل إدارة الازمة والكارثة). وحساب قيم Beta Coefficient التي تمثل معاملات الأنحدار في صيغتها القياسية.

نتائج الدراسة

هذا الجزء يستعرض اهم النتائج التي توصل اليها البحث وتتماشا مع الأهداف البحثية

جدول رقم (3)

الأوزان الناتجة من التحليل العاملي

العوامل (Factors)	الاوزان القياسية	المتغيرات (Variables)
التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية ← (X_1)	.654 .659 .229 .595 .673 .468	<ul style="list-style-type: none"> • يوجد سياسة سلامة واضحة للعمل من قبل إدارة الشركة • يوجد تقييم وتحليل للمخاطر والكوارث الطبيعية والصناعية المتوقعة واعداد خطة طوارئ لحماية الشركة والعمال عند وقوع الكارثة من قبل الإدارة • يوجد بالشركة اجهزة لقياس مخاطر بيئة العمل من قبل الإدارة • هناك مشاركة للعاملين في عملية تقييم المخاطر وتطوير إجراءات السلامة والصحة المهنية • تشعر بوجود تعاون بينك وبين الإدارة لتحقيق تطبيق قواعد وإجراءات السلامة • يتم توفير الإمكانيات الفنية والمالية لتطبيق إجراءات السلامة والصحة المهنية • يوجد خطط وبرامج للسلامة المهنية والعمل على متابعتها

	<p>361- 115- 331 553 369 729 117</p>	<ul style="list-style-type: none"> • هناك تنسيق وتعاون من قبل الإدارة العليا مع المنظمات الأهلية بشأن تطبيق إجراءات السلامة والصحة المهنية • هناك وصف وظيفي لكل وظيفة يشمل تحديد إجراءات السلامة والصحة المهنية التي تنطوي عليها الوظيفة وإطلاع الموظف عليها. • يوجد كفاءات وخبرات في الإدارة العليا تستطيع أن تعمل على تطوير إدارة أنظمة السلامة والصحة المهنية • يوجد لجان للسلامة المهنية لتمارس دورها في تطوير إجراءات السلامة والصحة المهنية • تهتم الإدارة بالبحوث والدراسات العلمية التي تتعلق بتطوير إجراءات السلامة المهنية • هناك إدارة عامة بالشركة متخصصة في السلامة والصحة المهنية
<p>الالتزام بتوفير وسائل السلامة والصحة المهنية في بيئة العمل ← (X_2)</p>	<p>375 499 818 734 259 064- 110- 309- 029- 294 229 169- 280- 555- 126 088</p>	<ul style="list-style-type: none"> • بناء السنترال (الجدران والأبواب والنوافذ) من مواد غير قابلة للاشتعال ومقاومة للمواد الكيماوية والانفجارات • تجهيزات السنترال من (إضاءة - تهوية - حرارة - ضوضاء) تعتبر مناسبة • يهتم المتخصصون بالفحص الدوري (الماء ، الكهرباء ، الغاز) • من السهل الحصول على معدات السلامة الشخصية (معطف -قفاز -خوذة - نظارة -كامرة -أحذية - بطانية إطفاء -قناع الوجه) • هناك أنظمة إنذار متوفرة داخل السنترال و مناسبة لنوعية السنترال وبأعداد كافية وأماكن مناسبة. • يوجد ترخيص وسجل صيانة دورية للمصاعد • يوجد ملصقات ولوحات تحذيرية لتوعية العاملين بالسلامة في أماكن واضحة ومرئية • الطفايات اليدوية والمتحركة بأنواعها (لبودرة أو الرغوة أو السائلة أو الغازية) وجرادال الرمل أعدادها كافية وأماكنها مناسبة. • كافة أجهزة وادوات الإطفاء المستخدمة مطابقة للمواصفة القياسية المصرية • هناك صيانة دورية لطفايات الحريق وكذلك للاجهزة والمعدات من قبل مختصين • يوجد نظام يقوم بالفحص الطبي للعاملين الجدد • هناك سجل طبي للعاملين يتعلق بحوادث العمل داخل السنترال • يوجد مداخل ومخارج للطوارئ بالسنترال بأعداد كافية وأماكن مناسبة والممرات به ممهدة ومزودة بوسائل الإضاءة والاشارات • يوجد خزانات للإسعافات الأولية جاهزة للاستخدام في محيط العمل • وجود تعاون وتنسيق بين العاملين في السنترال في تطبيق وسائل السلامة • يوجد مشرف للسلامة والصحة المهنية والبيئة بكل سنترال
<p>تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية ← (X_3)</p>	<p>584 530 506 292 581 725 421 272</p>	<ul style="list-style-type: none"> • يتلقى العاملون تدريباً على خطة الإخلاء والطوارئ ومكافحة الحريق • يتم اختبار فاعلية خطط الطوارئ وإدارة الأزمات والكوارث الصناعية والطبيعية واجراء تدريبات عملية عليها للتأكد من كفاءتها بصفة دورية • يتم التدريب على طرق التعامل مع أنظمة وأجهزة السلامة بالسنترال • هناك تدريب للعاملين على الإسعافات الأولية وكيفية استخدام مهمات الوقاية الشخصية أيضاً • يتم التدريب على إجراءات السلامة عن طريق (التدريب العملي-المحاضرات - النشرات المطبوعة- ورش العمل) • يتم توعية وتدريب العاملين بالمخاطر التي يواجهونها والزامهم باستخدام وسائل الوقاية المقررة • هناك مشاركة للعاملين في الندوات والمؤتمرات الدولية المتعلقة بالسلامة والصحة المهنية) (منظمة العمل الدولية -منظمة الصحة العالمية) • تلقيت تدريب على إجراءات السلامة من قبل جهات خارجية(الدفاع المدني - وزارة الصحة -وزارة العمل -مؤسسات أهلية)
<p>الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية ← (X_4)</p>	<p>541 162 454- 195- 037-</p>	<ul style="list-style-type: none"> • هناك لوائح وقوانين خاصة بالسلامة والصحة المهنية داخل الشركة • توجد قوانين الزامية لاستخدام معدات الوقاية الشخصية • يتوفر منشورات وكتيبات إرشادية لكيفية استخدام الوسائل الوقائية • يتوفر داخل الشركة لوائح وقوانين تهتم بإجراءات السلامة والصحة المهنية • هناك التزام من قبل الشركة باللوائح والقوانين المصرية (قانون العمل المصري -قانون الصحة العامة) • هناك توجيهات وإرشادات للوائح المرتبطة بتطبيق الأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية

	.857 .829 .300 .279 .153	<ul style="list-style-type: none"> • هناك التزام من قبل العاملين باللوائح والقوانين المتعلقة بقواعد السلامة المهنية • تفرض عقوبات رادعة على غير الملتزمين باستخدام وسائل الوقاية داخل السنترال • اللوائح والأنظمة الموجودة متوافقة مع المعايير الدولية (OSHA) و واضحة ومحددة • هناك قواعد وقوانين تطبق مبدأ التعويض والتأمين للعاملين
مرحلة اكتشاف إشارات الإنذار المبكر ← (y_1)	.798 .920 .816	<ul style="list-style-type: none"> • تولى الإدارة العليا في الشركة اهتماما ودعما لرصد مؤشرات حدوث الأزمات. • يتم مسح بيئة عمل الشركة بصورة شاملة للتعرف على مؤشرات احتمال حدوث الأزمة. • تهتم إدارة الشركة بعمليات تصنيف وتبويب مؤشرات حدوث الأزمات.
مرحلة الاستعداد والوقاية ← (y_2)	.552 .793 .839	<ul style="list-style-type: none"> • يتم تشكيل فرق مختلفة ومتعددة لحل العديد من الأزمات المحتملة في الشركة • يسهل الحصول على الإمكانات المادية والبشرية والتقنية والمعلوماتية المطلوبة من الإدارات والأقسام الأخرى عند الحاجة لها من أجل التعامل مع الأزمات. • تتوفر برامج وخطط كافية وجاهزة لإدارة الأزمات في الشركة ويتم العمل على مراجعتها وتطويرها باستمرار.
مرحلة احتواء الأضرار ← (y_3)	.650 .925 .858	<ul style="list-style-type: none"> • تقوم إدارة الشركة بالاستجابة الفورية لاحتواء الأزمة عن طريق توزيع المهام وتحديد الصلاحيات بفترة قصيرة ومناسبة عند حدوث الأزمة. • يتم التأثير في الأحداث واستخدام إجراءات الطوارئ التي تقلل وتحد من الأضرار التي تسببها الأزمة بكفاءة. • تقوم إدارة الشركة بإعداد غرفة عمليات مناسبة ومجهزة بالتقنيات الحديثة لاحتواء أسباب وأضرار الأزمة.
مرحلة استعادة النشاط ← (y_4)	.563 .842 .729	<ul style="list-style-type: none"> • تحافظ إدارة الشركة في ظروف الأزمات على اتخاذ الإجراءات اللازمة لمواصلة ممارسة النشاطات الاعتيادية في الشركة دون أي تأخير. • تقوم إدارة الشركة بتحديد الاحتياجات اللازمة للمواقع المختلفة التي تأثرت بالأزمة لمعالجة تأثيرات الأزمة واستعادة النشاط الاعتيادي. • تعمل إدارة الشركة على اتخاذ كافة الإجراءات اللازمة للتخفيف من آثار الأزمة والحد من استمرار حدوثها.
مرحلة التعلم ← (y_5)	.811 .848 .686	<ul style="list-style-type: none"> • تستخلص إدارة الشركة بصورة فعالة الدروس والعبر من الأزمات التي واجهتها سابقا في محاولة للاستفادة منها مستقبلا. • تقوم إدارة الشركة بتقييم خطط وبرامج إدارة الأزمات السابقة بقصد تطويرها وتحسينها من أجل التعامل مع الأزمات المستقبلية. • تعمل إدارة الشركة على الاستفادة من أساليب معالجة الأزمات في الإدارات الأخرى في البلد أو في الدول الأخرى ذات التجارب المشابهة.

ويتضح من جدول رقم (3) الأوزان الناتجة من التحليل العاملي ،

و كانت اهم النتائج تمثلت في:

- أعلى الاوزان القياسية في مؤشر التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية (X_1) هو 729. والتي تمثل بعبارة تهتم الإدارة بالبحوث والدراسات العلمية التي تتعلق بتطوير إجراءات السلامة المهنية .
- أعلى الاوزان القياسية في مؤشر الالتزام بتوفير وسائل السلامة والصحة المهنية في بيئة العمل (X_2) هو 818. والتي تمثل بعبارة يهتم المتخصصون بالفحص الدوري (الماء ، الكهرباء ، الغاز)

- أعلى الاوزان القياسية فى مؤشر تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية (X_3) هو 725. والتي تمثل بعبارة يتم توعية وتدريب العاملين بالمخاطر التي يواجهونها والزامهم باستخدام وسائل الوقاية المقررة
- أعلى الاوزان القياسية فى مؤشر الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية (X_4) هو 857. والتي تمثل بعبارة هناك توجيهات وإرشادات للوثائق المرتبطة بتطبيق الأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية
- أعلى الاوزان القياسية فى مؤشر مرحلة اكتشاف إشارات الإنذار المبكر (y_1) هو 920. والتي تمثل بعبارة يتم مسح بيئة عمل الشركة بصورة شاملة للتعرف على مؤشرات احتمال حدوث الأزمة.
- أعلى الاوزان القياسية فى مؤشر مرحلة الاستعداد والوقاية (y_2) هو 839. والتي تمثل بعبارة تتوفر برامج وخطط كافية وجاهزة لإدارة الأزمات في الشركة ويتم العمل على مراجعتها وتطويرها باستمرار .
- أعلى الاوزان القياسية فى مؤشر مرحلة احتواء الأضرار (y_3) هو 925. هو والتي تمثل بعبارة يتم التأثير في الأحداث واستخدام إجراءات الطوارئ التي تقلل وتحد من الأضرار التي تسببها الأزمة بكفاءة.
- أعلى الاوزان القياسية فى مؤشر مرحلة استعادة النشاط (y_4) هو 842. والتي تمثل بعبارة تقوم إدارة الشركة بتحديد الاحتياجات اللازمة للمواقع المختلفة التي تأثرت بالأزمة لمعالجة تأثيرات الأزمة واستعادة النشاط الاعتيادي .
- أعلى الاوزان القياسية فى مؤشر مرحلة التعلم (y_5) هو 848. والتي تمثل بعبارة تقوم إدارة الشركة بتقييم خطط وبرامج إدارة الأزمات السابقة بقصد تطويرها وتحسينها من أجل التعامل مع الأزمات المستقبلية.

جدول رقم (4) مصفوفة الارتباطات البسيطة بين متغيرات الدراسة										
		X_1	X_2	X_3	x_4	y_1	y_2	y_3	y_4	y_5
X_1	Pearson Correlation	1	-.083	.408	.384	.402	-.371	.276	.162	-.234
	Sig. (2-tailed)		.171	.000	.000	.000	.000	.000	.007	.000
X_2	Pearson Correlation	-.083	1	-.040	.037	.030	.052	-.205	.162	.181
	Sig. (2-tailed)	.171		.510	.537	.626	.393	.001	.007	.003
X_3	Pearson Correlation	.408	-.040	1	.098	-.089	-.208	.258	.161	-.286
	Sig. (2-tailed)	.000	.510		.103	.139	.001	.000	.007	.000
x_4	Pearson Correlation	.384	.037	.098	1	.328	-.237	.053	-.110	-.021
	Sig. (2-tailed)	.000	.537	.103		.000	.000	.381	.069	.723
y_1	Pearson Correlation	.402	.030	-.089	.328	1	-.099	-.142	.157	.061
	Sig. (2-tailed)	.000	.626	.139	.000		.103	.018	.009	.316
y_2	Pearson Correlation	-.371	.052	-.208	-.237	-.099	1	.137	.186	.260
	Sig. (2-tailed)	.000	.393	.001	.000	.103		.023	.002	.000
y_3	Pearson Correlation	.276	-.205	.258	.053	-.142	.137	1	-.098	-.450
	Sig. (2-tailed)	.000	.001	.000	.381	.018	.023		.106	.000
y_4	Pearson Correlation	.162	.162	.161	-.110	.157	.186	-.098	1	.172
	Sig. (2-tailed)	.007	.007	.007	.069	.009	.002	.106		.004
y_5	Pearson Correlation	-.234	.181	-.286	-.021	.061	.260	-.450	.172	1
	Sig. (2-tailed)	.000	.003	.000	.723	.316	.000	.000	.004	

ويتضح من جدول رقم (4) انه يوجد ارتباط بين أغلب بين متغيرات قواعد واجراءات السلامة والصحة المهنية التابعة لإدارة السلامة والصحة المهنية والمتمثلة في اربعة مؤشرات وهم (X_1 X_2 X_3 X_4) ومتغيرات مراحل إدارة الازمات والكوارث التابعة لإدارة الازمات والكوارث والمتمثلة في خمسة مؤشرات وهم (y_1 y_2 y_3 y_4 y_5) ،

و كانت اهم النتائج تمثلت في:

- مرحلة اكتشاف اشارات الإنذار المبكر ترتبط ارتباط معنوي مع (التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية و الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية) وذلك نظرا لأن Sig. اقل من 05 .

- مرحلة الاستعداد والوقاية ترتبط ارتباطاً معنوي مع (التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية و تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية و الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية) وذلك نظراً لأن Sig. اقل من 05 .
- مرحلة احتواء الأضرار ترتبط ارتباطاً معنوي مع (التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية و الالتزام بتوفير وسائل السلامة والصحة المهنية في بيئة العمل و تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية) وذلك نظراً لأن Sig. اقل من 05 .
- مرحلة استعادة النشاط ترتبط ارتباطاً معنوي مع (التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية و الالتزام بتوفير وسائل السلامة والصحة المهنية في بيئة العمل و تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية) وذلك نظراً لأن Sig. اقل من 05 .
- مرحلة التعلم ترتبط ارتباطاً معنوي مع (التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية و الالتزام بتوفير وسائل السلامة والصحة المهنية في بيئة العمل و تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية) وذلك نظراً لأن Sig. اقل من 05 .

جدول رقم (5)

معاملات الانحدار القياسية - وترتيب أهميتها

المتغيرات	التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية	الالتزام بتوفير وسائل السلامة والصحة المهنية في بيئة العمل	تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية	الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية
مرحلة اكتشاف اشارات الإنذار المبكر	(الاول) .46	(الرابع) .05	(الثاني) -.29	(الثالث) .18
مرحلة الاستعداد والوقاية	(الاول) -.29	(الرابع) .03	(الثالث) -.08	(الثاني) -.18
مرحلة احتواء الأضرار	(الاول) .21	(الثاني) -.18	(الثالث) .17	(الرابع) -.04
مرحلة استعادة النشاط	(الاول) .22	(الثالث) .19	(الرابع) .10	(الثاني) -.21
مرحلة التعلم	(الثالث) -.15	(الثاني) .16	(الاول) -.22	(الرابع) .05

ويتضح من جدول رقم (5) معاملات الانحدار القياسية وترتيب أهميتها وذلك باستخدام قيم **Beta Coefficient** ،

و كانت اهم النتائج تمثلت فى :

● **مرحلة اكتشاف اشارات الإنذار المبكر** أن العوامل التى تؤثر فيها وبترتيبها حسب الدرجة الاعلى فى الاهمية وهى التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية ثم يليها تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية ثم يليهم الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية ثم يليهم فى المرحلة الاخيرة الالتزام بتوفير وسائل السلامة والصحة المهنية فى بيئة العمل وذلك بحساب قيم Beta.

● **مرحلة الاستعداد والوقاية** أن العوامل التى تؤثر فيها وبترتيبها حسب الدرجة الاعلى فى الاهمية وهى التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية ثم يليها الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية ثم يليهم تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية ثم يليهم فى المرحلة الاخيرة الالتزام بتوفير وسائل السلامة والصحة المهنية فى بيئة العمل وذلك بحساب قيم Beta.

● **مرحلة احتواء الأضرار** أن العوامل التى تؤثر فيها وبترتيبها حسب الدرجة الاعلى فى الاهمية وهى التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية ثم يليها الالتزام بتوفير وسائل السلامة والصحة المهنية فى بيئة العمل ثم يليهم تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية ثم يليهم فى المرحلة الاخيرة الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية وذلك بحساب قيم Beta.

● **مرحلة استعادة النشاط** أن العوامل التى تؤثر فيها وبترتيبها حسب الدرجة الاعلى فى الاهمية وهى التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية ثم يليها الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية ثم يليهم الالتزام بتوفير وسائل السلامة والصحة المهنية فى بيئة العمل ثم يليهم فى المرحلة الاخيرة تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية وذلك بحساب قيم Beta.

● **مرحلة التعلم** أن العوامل التى تؤثر فيها وبترتيبها حسب الدرجة الاعلى فى الاهمية وهى تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية ثم يليها الالتزام بتوفير وسائل السلامة والصحة المهنية فى بيئة العمل ثم يليهم التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية ثم يليهم فى المرحلة الاخيرة الالتزام بتوفير وتطبيق اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة المهنية وذلك بحساب قيم Beta.

خلاصة الدراسة

ويتضح ان التزام الإدارة العليا لتطبيق إجراءات السلامة والصحة المهنية لها الاولوية فى مراحل إدارة الازمات والكوارث ويليهما كلا من الالتزام بتوفير وسائل السلامة والصحة المهنية فى بيئة العمل و الالتزام بتوفير وتطبيق

اللوائح والأنظمة والقوانين المتعلقة بالسلامة والصحة والمهنية ولهما نفس الأهمية وثم يليهم في الأهمية المرحلة الأخيرة تدريب وتوعية العاملين بإجراءات السلامة والصحة المهنية

المراجع العربية

م	العربية
1	العباسي ، عبد الحميد ، (2016) " تحليل البيانات باستخدام الحزم الإحصائية " معهد الدراسات والبحوث الإحصائية ، جامعة القاهرة .
2	الروسان وآخرون (2009) : الأمن الصناعي والسلامة المهنية ، الطبعة الثانية ، مكتبة المجتمع العربي للنشر والتوزيع ، عمان .
3	الشيخ خليل، علي (2008) : تقييم وسائل الوقاية والسلامة المستخدمة في مستشفيات قطاع غزة الحكومية وأثرها على أداء العاملين ، (رسالة ماجستير) ، الجامعة الإسلامية، غزة.
4	قطيشات، تالا وآخرون (2007) : مبادئ في الصحة والسلامة العامة ، الطبعة الثانية ، دار المسيرة للنشر والتوزيع ، عمان ، الأردن .
5	المغني، أميمه (2006) : واقع وإجراءات الأمن والسلامة المهنية المستخدمة في منشآت قطاع الصناعات التحويلية في قطاع غزة ، (رسالة ماجستير) ، الجامعة الإسلامية ، غزة.
6	المديفر، فهد (2005) : مدى فعالية تطبيق أنظمة الأمن والسلامة المهنية والتقنية ، دراسة مسحية على معامل الأقسام العلمية بكليات البنات ، الرياض (رسالة ماجستير) ، جامعة نايف العربية للعلوم الأمنية ، الرياض.
7	السرحي، حسن وآخرون (2002) : الصحة والسلامة في المكاتب المؤتمتة ، دراسة حالة على مكتبة جامعة الملك عبد العزيز ، مجلة الاتجاهات الحديثة في المكتبات والمعلومات بالقاهرة ، المجلد 7 ، العدد 13.
8	الشعلان ، فهد احمد (2002) . إدارة الأزمات : الاسس – المراحل – الآليات . (ط 2) ، الرياض " أكاديمية نايف العربية للعلوم الأمنية .
9	الهانسي ، مختار (1991) : " مقدمه في طرق التحليل الاحصاء " قسم الاحصاء والرياضة والتأمين ، كلية التجارة ، جامعة الاسكندرية
10	الحملوي، محمد رشاد (1995) إدارة الأزمات : تجارب محلية وعالمية" ، مكتبة عين شمس، القاهرة، ط ٢
11	الحملوي، محمد رشاد (1995) التخطيط لمواجهة الأزمات : عشر كوارث هزت مصر " ، القاهرة، مكتبة عين شمس.
12	عليوة، السيد (2003) : إدارة الوقت والأزمات والإدارة بالأزمات ، القاهرة، دار الأمين للنشر والتوزيع.
13	برى ، عدنان و هندی ، محمود (2009) : " مبادئ الاحصاء والاحتمالات " ، مكتبة الشقري للنشر والتوزيع

14	طبية ، احمد عبدالسميع (2008) : " مبادئ الاحصاء " ، دار البداية ، عمان ، ط1
15	الهانسي ، مختار (1991) : " مقدمه فى طرق التحليل الاحصاء " قسم الاحصاء والرياضة والتأمين ، كلية التجارة ، جامعة الاسكندرية
16	طه ، محمد عبدالحميد (1985) " مقدمه فى الاحصاء " ، الهيئة العامة للتعليم التطبيقي والتدريب
17	محمد ، على ابو القاسم (1987) " اساليب الاحصاء التطبيقي " ، المعهد العربى للتخطيط بالكويت، دار الشباب للنشر والترجمة
18	منظمة العمل الدولية (2007) ،أماكن عمل آمنة وصحية: جعل العمل اللائق واقعاً حقيقياً. مكتب العمل الدولي، جنيف .
19	عبد الرؤوف، مصطفى (2002) : الأمن والسلامة في المباني التعليمية واستخدام الحاسب الآلي في تقييم مستوى السلامة لمدارس التعليم الأساسي ، (رسالة دكتوراه) ، جامعة القاهرة ، مصر .
20	وزارة القوى العاملة (2003) : قانون العمل المصري رقم (12) لعام (2003) ،

المراجع الأجنبية

21	Bartlett, M.S. (1939), "A Note On Test Of Significance In Multivariate Analysis", Proceedings Of Cambridge Philosophical Society, 35.
22	Cristine Person and Ian Mitroff (1993): " From Crisis Prone To Crisis Prepared: A Framework For Crisis Management", Academy of Management Executive . Vol 7, No 1.
23	Graham Allison: Essence of Decision, U.S.A., Little Brown and Company, 1971, P.5.
24	Hotelling. H. (1936), "Relations Between Two Sets Of Variates", Biometrika, Vol28, Pp.321-336
25	Sterwart, D.K, And Love, W.A.(1968), "A General Canonical Correlation Index", Psychological Bulletin Vol,70, Pp.160-163.
26	Dejoy , D., Schaffer , B. & Wilson, N. , (2003) : " Creating Safer Workplaces : assessing the determinants and role of Safety Climate " , Journal of Safety Research , USA , 2003.
27	Lin, J. & Mill, A. (2001). Measuring the occupational health and safety performance of construction companies in Australia, Facilities, vol. 19, No. 3/4, pp. 131-138

المواقع الالكترونية

28	http://www.ilo.org/global
29	http://www.alolabor.org
30	http://www.education.gov.bh -
31	http://www.safety-eng.com -
32	http://www.salama-libya.org -

مؤشرات إحصائية للحوادث المرورية باستخدام تحليل السلاسل الزمنية بمنطقة عسير السعودية للفترة من (1417-1434) محمد محمد أحمد المزاح¹

الملخص

تعتبر الحوادث المرورية من أكبر الأخطار التي تواجه المجتمع وتستنزف موارده وطاقاته البشرية، حيث تناول هذا البحث تحليلاً إحصائياً زمنياً لأعداد الحوادث المرورية بمنطقة عسير للفترة 1417-1434هـ، مع التركيز على تعيين التوزيعات الاحتمالية المناسبة للبيانات، ومن ثم تقدير الطور (phase) العشوائي المولد للسلسلة الزمنية بهدف التعرف على الوضع الراهن لأعداد الحوادث، والتطورات المستقبلية لها، وإظهار السلوك المستقبلي لهذه الطوريات بناءً على النموذج المتنبأ به.

وقد أظهر شكل الانتشار للسلسلة الزمنية وجود حركة اهتزازية أقرب ما تكون للاهتزازات التوافقية، والتي تمثل المركبة الدورية، وعدم وجود المركبة الموسمية، وأن النموذج المناسب لعرض بيانات السلسلة الزمنية هو النموذج التجميعي. وأن مركبة الاتجاه العام كانت بصورة نماذج غير خطية، أي أن النموذج المتنبأ به هو النموذج "اللا خطي". وتعد نتائج هذا البحث تقيماً لواقع فترة زمنية مهمة في تاريخ الحوادث، ونسبة الإصابة بالمنطقة. وأن أعداد الحوادث المرورية تسير بشكل متنامي، ومن المتوقع أن يزيد عدد الحوادث إلى حوالي (66173) حادثة في حلول عام 1443، ومتناقصة لنسبة الإصابة بالإضافة إلى وجود عوامل لها دور فاعل ومؤثر في الحوادث المرورية، وأن موضوع البحث بجانبه النظري والتطبيقي يفتح مجالاً واسعاً أمام الباحثين لاستخدام نماذج إحصائية تنبؤية أخرى.

الكلمات المفتاحية:

التذبذبات التوافقية، الحوادث المرورية، السلسلة الزمنية اللاخطية، نسبة الإصابة، النموذج التجميعي.

مقدمة

إن المتأمل لسير الحوادث المرورية بالسعودية، يلاحظ أنها صارت تمثل هاجساً وقلقاً وشعباً مخيفاً لكافة أفراد المجتمع، بل أضحت واحدة من أهم المشكلات التي تستنزف الموارد المادية والطاقات البشرية، وتستهدف أهم مقومات الحياة وهو العنصر البشري. وتنبع أهمية البحث من خطورة الدور الذي تلعبه الحوادث المرورية في الحياة العامة. وتشير الدراسات المرورية ومنها ما أوضحه رصد وحدة التقارير الاقتصادية بصحيفة الشرق الأوسط من تراجع الحوادث في المملكة بنسبة 9% خلال العام 2014م وانخفاض الوفيات إلى 20.5 حالة وفاة يومياً، وتصدرت منطقة الرياض من حيث الأكثر حوادث، ومنطقة مكة المكرمة من حيث الأكثر حالات الوفاة وتلتها منطقة الشرقية ثم منطقة عسير، كما تصدرت منطقة مكة المكرمة من حيث عدد الإصابات (الشرق الأوسط، 2015).

وأوضحت الدراسة التي أجراها عصام وآخرون (2013) أن 80% من الحوادث المرورية تقع داخل المدن، في حين 18% فقط تقع خارج المدن أو على الخطوط السريعة، وقدرت الدراسة الخسائر الاقتصادية كاملة لوفيات الحوادث المرورية بحوالي 79.92 مليار ريال، في حين يقدر إجمالي تكاليف علاج الإصابات البسيطة بحوالي 170.73 مليون ريال، والإصابات البليغة بـ 135 مليون ريال، وتكاليف الإضرار بالمتلكات "السيارات" 6.94 مليار ريال، أما إجمالي التكاليف الشاملة للحوادث المرورية تقدر بحوالي 87.17 مليار ريال (jazirahonline.com/news).

يعد تحليل السلاسل الزمنية من الأساليب العلمية والإحصائية لتفسير سلوك الظاهرة، حيث أشار الباحثان النعيمي والشاروط (2000) بدراستهما استخدام التدخل في السلاسل الزمنية لتقدير عدد المصابين بالأورام الخبيثة بتأثير عامل التدخل لتحديد أفضل نموذج للتنبؤ بعدد المصابين، ولاحظا أن أعداد المصابين بالأورام الخبيثة غير مستقرة في الوسط الحسابي والتباين، وأن هناك اتجاهًا عامًا وواضحًا في السلسلة بعد عام 1973. أيضاً استخدم العلكة (2009) السلاسل الزمنية للتنبؤ بأعداد طلاب الماجستير ونسبة الخريجين منهم في الجامعات السورية، والذي توصل إلى أن أعداد طلاب الماجستير تسير بشكل متنامي (متزايد) في حين نسبة الخريجين منهم تتضاءل بشكل كبير. وقام الجبوري (2009) بدراسة السلاسل الزمنية ثنائية المتغيرات للتنبؤ بنسبة التضخم وعلاقته بسعر صرف الدولار الأمريكي مقابل الدينار العراقي، وتوصل إلى أن متجه التضخم النقدي يتبع نموذج الانحدار الذاتي ثنائي المتغيرات غير المستقر من الرتبة الثانية. كما قام محمد (2010) باقتراح نموذج إحصائي باستخدام تحليل السلاسل الزمنية لحوادث الحركة المرورية في ولاية الخرطوم، حيث قام بتقسيم بيانات السلاسل الزمنية إلى ثلاث سلاسل زمنية الأولى حوادث مرورية بسيطة، والثانية جسمية، والثالثة حوادث أموات، وقد توصل إلى أن السلسلة الزمنية تسير بشكل متزايد وأن بعضها مستقر والبعض الآخر غير مستقر. أما طعمة (2012) فقد حدد في دراسته أن أفضل نموذج للتنبؤ بأعداد المصابين بالأورام الخبيثة والذي أظهر تزايد أعداد المصابين بالأورام بتأثير عامل

¹ قسم الرياضيات- كلية العلوم والآداب - جامعة الملك خالد - المملكة العربية السعودية.

قسم الرياضيات والحاسوب - كلية العلوم - جامعة إب - الجمهورية اليمنية.

التدخل، وقد لوحظ أن أعداد المصايين بالأورام الخبيثة غير مستقرة في الوسط الحسابي، وأن هناك اتجاهًا عامًا واضحًا في السلسلة بعد عام 2009 و أن النموذج الملائم للبيانات هو نموذج الانحدار من الرتبة الثانية. وسيقدم هذا البحث عرضاً وصفية مبسطة وبشكل أساسي على بحث سلوك السلسلة الزمنية الخاصة بالحوادث المرورية. وتعتبر نتائج هذا البحث أداة للتعرف على واقع فترة زمنية في تاريخ الحوادث المرورية، ويهدف هذا البحث إلى التنبؤ المستقبلي بأعداد الحوادث، والإصابات، والتطورات المستقبلية.

1. بعض الأسس النظرية لتحليل السلاسل الزمنية

السلسلة الزمنية هي مجموعة المشاهدات مرتبة وفق حدوثها، فهي بذلك عبارة عن سجل تاريخي يتم اعتماده لبناء التوقعات المستقبلية (فاندل، 1992)، وتتميز أي سلسلة زمنية بأن بياناتها مرتبة بالنسبة للزمن، وأن المشاهدات المتتالية عادة ما تكون غير مستقلة، مما يساعد على التوصل إلى تنبؤات موثوقة بها (Kirchgssner and Wolters, 2007) والغرض الأساسي من تحليل السلسلة هو الوصول لنموذج مناسب لتقدير أو قياس التغيرات، وبالتالي دراسة علاقاتها بالظروف المختلفة، من خلال التخلص من آثار العوامل المؤثرة في التغيرات وخاصة الاتجاه العام والتغيرات الدورية (Quenouille, 1957 and Murray, 2003).

نموذج السلسلة الزمنية:

يفترض نموذج السلسلة الزمنية أن قيم السلسلة دالة في مجموعة من العناصر المكونة لها مع الزمن: $Y = f(T_t, S_t, C_t, I_t)$ وفي هذا السياق

يبرز نموذجان لتجسيد العلاقة بين العناصر المتداخلة والتي يتفاعلها لتشكيل السلسلة الزمنية (Larsen and Marx, 1993)، ومنها:

1. النموذج التجميعي: يقوم على افتراض أن كل قيمة لكل من مكونات السلسلة لا تؤثر ولا تتأثر بقيمة غيرها من المكونات، ويكتب بالصيغة:

$$Y = T_t + S_t + C_t + I_t$$

2. النموذج الضربي: يقوم على افتراض أن قيم الظاهرة تساوي حاصل ضرب مكوناتها وأنها تعتمد على بعضها البعض أي:

$$Y = T_t \cdot S_t \cdot C_t \cdot I_t \quad (\text{Maestas and Robert, 2000})$$

قياس أثر الاتجاه العام T في حالة الاتجاهات غير الخطية Non-Linear Trends

يمكن تعيين عنصر الاتجاه العام بعدة طرق مختلفة منها طريقة المربعات الصغرى (البلداوي، 1997)، وسنركز على النماذج اللاخطية لتقدير معالم مركبة

الاتجاه العام $Y = ax^2 + bx + c \rightarrow (1)$ حيث a, b, c معاملات النموذج اللاخطي (Box and Jcnkins, 1976). ويتم تعيين قيم a, b, c من خلال العلاقات الآتية:

$$\left. \begin{aligned} \sum y &= nc + b \sum x + a \sum x^2 \\ \sum xy &= c \sum x + b \sum x^2 + a \sum x^3 \\ \sum x^2 y &= c \sum x^2 + b \sum x^3 + a \sum x^4 \end{aligned} \right\} \rightarrow (2)$$

أي أن:

$$\hat{a} = \frac{n \sum x^2 y - (\sum x^2)(\sum y)}{n \sum x^4 - (\sum x^2)^2} \rightarrow (3)$$

$$\hat{b} = \frac{\sum xy}{\sum x^2} \rightarrow (4)$$

$$\hat{c} = \frac{\sum y \sum x^4 - \sum x^2 y \sum x^2}{n \sum x^4 - (\sum x^2)^2} \rightarrow (5)$$

ويشترط لتحليل أية سلسلة زمنية أن تكون ساكنة أو تنسم بالسكون أي تحقق الشروط: ثبات المتوسط الحسابي والتباين أي يحقق

: $Var(Z_t) = E[(Z_t - \mu)^2] = \sigma^2$ (Brock and Davis, 1991)، وقد أشار (شعراوي، 2005) وأيضاً (Shumway and

Stoffer, 2006; Mosteller and Tukey, 1977; Clive and Granger, 2004) أن الفرضية الأساسية لتطبيق طريقة المربعات الصغرى في

النموذج الخطي هي عدم وجود ظاهرة الارتباط الذاتي. ونخلص مما سبق إلى أن الارتباط الذاتي يعتمد فقط على $(t-s)$ أي:

اختبار عشوائية أخطاء التنبؤ. $\rho_{t-s} = \frac{E[(Z_t - \mu)(Z_s - \mu)]}{\sigma^2} = \frac{\sigma_{ts}}{\sigma^2}$ وتستخدم دالة الارتباط الذاتي للبواري (RACF) لفحص ملاءمة النموذج عن طريق

نتائج تحليل السلاسل الزمنية

منهجية البحث

اعتمد البحث على بيانات السلسلة الزمنية لأعداد الحوادث المرورية والإصابات والتي تم الحصول عليها من الإدارة العامة للمرور بمنطقة عسير للفترة (1417-1434)، معتمداً المنهج الوصفي التحليلي من خلال الاطلاع على عدد من الأدبيات في هذا المجال، والتطبيق عليها وإجراء المعالجات الإحصائية

والعروض البيانية اللازمة من خلال بعض البرامج الحاسوبية والإحصائية (SPSS, Data Fit, Easy fit)

أولاً: أعداد الحوادث المرورية

بالرجوع إلى الإحصاءات السنوية الصادرة عن الإدارة العامة للمرور بمنطقة عسير خلال الفترة من (1417-1434) كانت بيانات الحوادث المرورية بمنطقة عسير كما هو موضح في جدول (1).

جدول (1): أعداد الحوادث المرورية خلال الفترة (1417-1434)

العام	1417	1418	1419	1420	1421	1422	1423	1424	1425
عدد الحوادث	7351	8095	7932	8274	9273	9521	10801	20042	18447
العام	1426	1427	1428	1429	1430	1431	1432	1433	1434
عدد الحوادث	22348	19550	20553	22954	25289	29469	27794	30530	31279

وبناءً على تحليل البيانات نلاحظ أن التوزيع الاحتمالي التوفيقى Fitting of Distribution المناسب للبيانات المقدمة وفقاً لاختبار كلموجروف-سميرنوف هو توزيع Johnson SB Distribution

ذو المعالم $\gamma, \zeta, \lambda > 0, \delta > 0$ ، وأن دالة الكثافة الاحتمالية لهذا التوزيع تعطى بالصورة الآتية:

$$f(x) = \frac{1}{\lambda \sqrt{2\pi} z(1-z)} \exp \left[-\frac{1}{2} \left(\gamma + \delta \ln \left(\frac{z}{1-z} \right) \right)^2 \right] \rightarrow (6)$$

حيث $Z = \frac{X - \zeta}{\lambda}$ وأن مدى التوزيع الاحتمالي هو: $\zeta < X < \zeta + \lambda$ وأشارت البيانات تحت الدراسة أن:

$$\gamma = 0.02976, \quad \delta = 0.4322, \quad \zeta = 5955.4, \quad \lambda = 25230$$

ومن الجدول (2) نلاحظ أن توزيع بيانات الحوادث المرورية شبه ملتو ناحية اليمين التواءً بسيطاً موجباً، وهذا يعني أن عدد الحوادث المسجلة في السنوات الأولى للفترة قيد البحث سار بشكل رتيب حول قيمة البدء، ومن ثم سار بشكل متذبذب في الفترات اللاحقة، ثم سار على نحوٍ مطرد في السنوات الأخيرة للفترة الزمنية قيد البحث، أما من حيث التفرطح فإننا نلاحظ تذبذباً قليلاً لتوزيع البيانات قيد الدراسة وهذا يعني أن تشتت البيانات حول وسطها الحسابي قليل نسبياً.

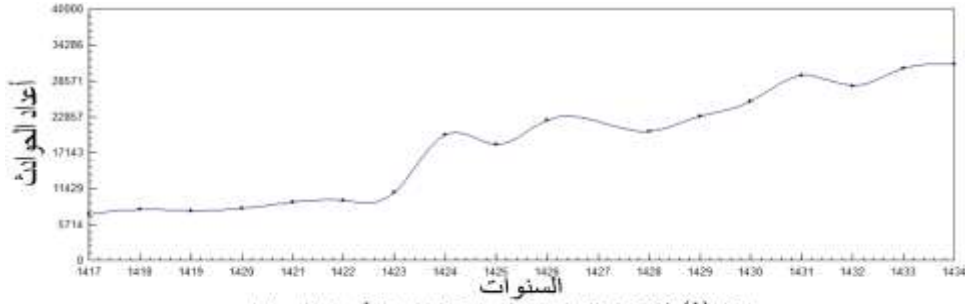
جدول (2): المقاييس الوصفية للبيانات الخاصة بالحوادث المرورية

اصغر قيمة	أكبر قيمة	المجموع	المتوسط الحسابي	الانحراف المعياري	معامل الالتواء	معامل التفرطح
7351	31279	329502	18305.67	8664.76	0.045	-1.528

التحليل الزمني للبيانات

إن دراسة السلسلة الزمنية تهدف إلى الحصول على وصف دقيق للملامح الخاصة بالطور (phase)

العشوائي المولد للسلسلة الزمنية بغرض بناء نموذج رياضي لتفسير وشرح سلوك السلسلة، بالاعتماد على معلومات الماضي، كما يفيد النموذج الرياضي الذي سيتم بناءه في تحديد القوى المؤثرة في السلسلة الزمنية بغرض الحصول على تنبؤات أكثر دقة، ومن المعلوم أن كل سلسلة زمنية سيرافقها طور عشوائي $\{Z_t\}_{t \in T}$ حيث يمكن تمثيل قيم هذه السلسلة الزمنية لقيم الطور العشوائي والتي يرمز لها بالرمز X_t من أجل $t \in T$ (العكلة، 2009)، وفي هذا البحث ستقوم بدراسة السلاسل الزمنية والتي أخذت مشاهداتها على فترات زمنية متساوية الطول (عام كامل) بحسب طبيعة البيانات قيد الدراسة، والعرض البياني الممثل بالشكل (1) يوضح انتشار السلسلة الزمنية الخاصة بأعداد الحوادث، حيث نلاحظ من الشكل وجود اتجاه عام Trend للسلسلة والذي يمثل تغيرات قيم الطور العشوائي المولد للسلسلة الزمنية على المدى البعيد.



شكل (1) الانتشار للسلسلة الزمنية الخاصة بأعداد الحوادث

ويعد أسلوب الانحدار من أكثر هذه الطرق شيوعاً في تحديد السلسلة الزمنية، ولاختبار الارتباط العشوائي للبيانات قيد الدراسة سوف نستخدم طريقة الفروق الأولى لـ موريسيس Wallis-Moore عند مستوى معنوية $\alpha = 0.05$ ، لإزالة أثر الاتجاه العام (Clive and Granger, 2004; George and Canavos, 1984).

فرضية العدم: H_0 : البيانات قيد البحث تمثل بيانات مستقلة عشوائياً أي أن: $H_0: \rho = 0$.

الفرضية البديلة: H_1 : البيانات قيد البحث مرتبطة عشوائياً أي أن: $H_1: \rho > 0$.

وللتحقق من ذلك تم إنشاء الجدول (3) الذي يعرض الفروق الأولى (الفروق المتتالية)، والتي تتم بطرح قيم مشاهدات السلسلة من بعضها البعض في ترتيب زمني محدد (العكس، 2009). والذي يبين أن عدد الإشارات الموجبة (+) يساوي $k = 13$ (علمياً بأن k هو متغير عشوائي يتوزع طبيعياً

حيث $\zeta \sim N\left(\frac{n-1}{2}, \frac{n+1}{12}\right)$ (Larsen and Marx, 1993)، وحيث $(n < 30)$ فإن الإحصاء الخاص بهذا الاختبار نحصل عليه من:

$$TG = \frac{\left| \zeta - \frac{n-1}{2} \right| - \frac{1}{2}}{\sqrt{\frac{n+1}{12}}} = \frac{\left| 13 - \frac{18-1}{2} \right| - \frac{1}{2}}{\sqrt{\frac{18+1}{12}}} = 3.1789 \rightarrow (7)$$

وبما أننا لاحظنا وجود اتجاه عام للسلسلة الزمنية فإننا سوف نستخدم اختبار ذو اتجاه واحد، وتكون المنطقة الحرجة عند مستوى معنوية 5% هي

1.64

وبالمقارنة نجد أن $TG < Z_{0.05}$ وبناء على ذلك نرفض الفرضية الصفرية H_0 وقبل بالفرضية البديلة H_1 عند مستوى معنوية $\alpha = 0.05$ أي أن البيانات قيد الدراسة مرتبطة عشوائياً.

جدول (3): الفروق الأولى (الفروق المتتالية)

السنة	1417	1418	1419	1420	1421	1422	1423	1424	1425
عدد الحوادث المرورية	7351	8095	7932	8274	9273	9521	10801	20042	18447
$Sig(x_{t+1}, x_t)^*$		+	-	+	+	+	+	+	-
السنة	1426	1427	1428	1429	1430	1431	1432	1433	1434
عدد الحوادث المرورية	22348	19550	20553	22954	25289	29469	27794	30530	31279
$Sig(x_{t+1}, x_t)^*$	+	-	+	+	+	+	-	+	+

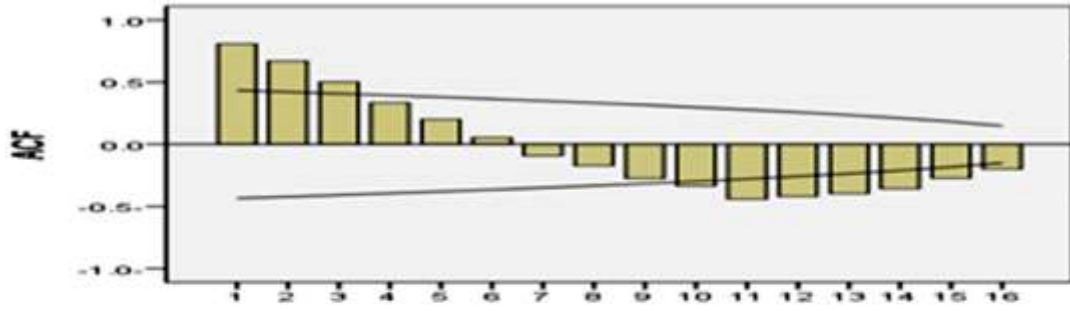
وبعد تطبيق دالة الارتباط الذاتي ACF بواسطة برنامج SPSS على بيانات الجدول (1) حصلنا على معاملات الارتباط الذاتي الموضحة في الجدول (4) وكذلك من شكل (2) نلاحظ أن دالة الارتباط الذاتي تتناقص بشكل ملحوظ وهذا دليل على هدوء السلسلة، ولاختبار معنوية معاملات الارتباط الذاتي عند الفجوة الأولى $k=1$ بفترة ثقة 95% نلاحظ أن القيمة الجدولية 1.96 والمحسوبة 3.43

حيث $Z_c = \frac{r}{\sqrt{1/n}}$ وحيث أن القيمة المحسوبة أكبر من الجدولية نرفض فرضية انعدام معاملات الارتباط عند $k=1$. وعند اختبار الفرضية عند $k=6$ نجد أن القيمة المحسوبة 0.229 وهذا يعني انعدام معاملات الارتباط الذاتي عند الفجوة 6، وبذلك يمكننا القول بهدوء

السلسلة الزمنية قيد البحث بعد الفجوة الخامسة. ومن شكل (1) نلاحظ ظهور حركة اهتزازية اقرب ما تكون للاهتزازات التوافقية والتي تمثلها المركبة الدورية للسلسلة الزمنية، في حين نرى أن المركبة الموسمية لا وجود لها في شكل الانتشار، لأنه من المعلوم بأن المركبة الموسمية تعبر عن التغيرات والتذبذبات الموسمية أو الفصلية الناتجة عن التغيرات في الفصول بسبب تأثير عوامل خارجية في مواسم محددة تعيد نفسها بشكل شبه منتظم على فترات زمنية متساوية الطول. وأن النموذج التجميعي هو النموذج المناسب عند انعدام إحدى المركبات (Granger,1977).

جدول (4) معاملات الارتباط الذاتي

Lag	1	2	3	4	5	6	7	8
Auto Corr	0.808 (0.000)	0.669 (0.000)	0.502 (0.000)	0.334 (0.000)	0.200 (0.000)	0.054 (0.000)	-0.092 (0.003)	-0.171 (0.000)
Lag	9	10	11	12	13	14	15	16
Auto Corr	-0.274 (0.000)	-0.335 (0.000)	-0.442 (0.000)	-0.419 (0.000)	-0.397 (0.000)	-0.357 (0.000)	-0.273 (0.000)	-0.200 (0.000)



شكل (٢) :دالة الارتباط الذاتي للبقاقي للنموذج اللاخطي

تقدير عنصر الاتجاه العام للسلسلة الزمنية

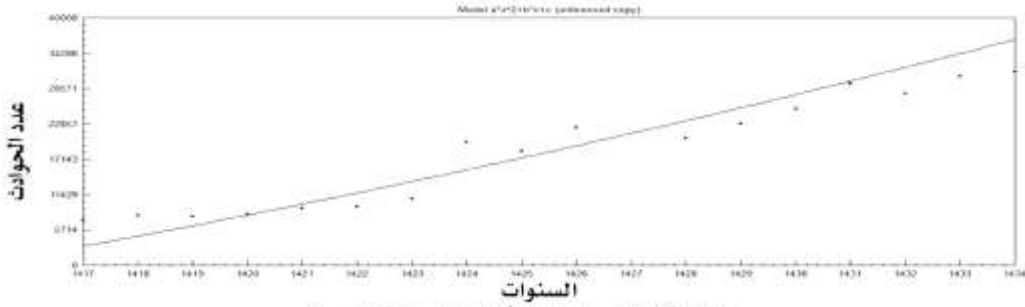
لتقدير عنصر الاتجاه العام $T(t)$ سوف نفترض ضمناً أن المكونين الدوري والموسمي محتواة في عنصر الخطأ (سواء كانت موجودة أم لا) وتكتب بالصورة الآتية: $I_t = T(t) + I_t$ وفي هذه الحالة يجب أن نُحدِث الخطأ العشوائي I_t أثناء عملية التقدير ويضاف كمتغير عشوائي بعد الانتهاء من عملية التقدير. ولذلك لن يتم إدخاله في عملية تقدير عنصر الاتجاه العام والعنصر الدوري، ويعتمد تقدير عنصر الاتجاه العام على التعرف على خط مستقيم يصف لنا حركة السلسلة الزمنية في فترة زمنية طويلة نسبياً، وبما أن البيانات قيد الدراسة هي بيانات سنوية فإن انحرافات قيم السلسلة عن خط الاتجاه العام لا تمثل فقط التغيرات العشوائية أو العرضية، بل تمثل التغيرات الدورية، وهذا يعني تعيين خط انحدار مناسب لبيانات السلسلة قيد الدراسة، وهذا يتطلب منا توفيق منحنى fitting curve لبيانات السلسلة الزمنية (Hamilton and James,1994). ولتوفيق هذا المنحنى نستخدم برنامج Curve(Data fit)، الذي يقوم بتوفيق المنحنى اعتماداً على طريقة المربعات الصغرى من جهة، ومن جهة أخرى يقوم البرنامج بتعيين المنحنى ذو الخطأ المعياري الأصغر أولاً، ومن ثم تعيين أكبر قيمة لمعامل الارتباط، حيث تم تقدير معالم نموذج الاتجاه التربيعي، فقد أوضحت نتائج التحليل أن القيم التقديرية لمعاملات النموذج في (3,4,5) a, b, c هي:

$$\hat{a} = 13.416, \hat{b} = 3667.719, \hat{c} = 25038286.026$$

أي يأخذ النموذج اللاخطي الصورة:

$$\hat{T}(t) = 13.416t^2 - 3667.719t + 25038286.026$$

وبخطأ معياري $S = 94.04$ ومعامل ارتباط $r = 0.967$ ومعامل التحديد $R^2 = 0.94$ ، وأن كفاءة النموذج حسب اختبار درين-وتسن الإحصائي هو: $d = 1.592$ وأن $d_u = 1.39$ ، $d_L = 1.16$ ، وحيث أن $d_U < d$ فلا يوجد لدينا ما يدفعنا لرفض الفرضية الصفرية H_0 عند مستوى معنوية $\alpha = 0.05$ ، والشكل (3) يوضح مركبة الاتجاه العام لهذه السلسلة الزمنية.



شكل (3) الاتجاه العام لأعداد الحوادث المرورية

تقدير العنصر الدوري للسلسلة الزمنية

يمكن تقدير العنصر الدوري بطريقة البواقي (Residual method) وذلك بفضله من عنصر الاتجاه العام وبالتالي يبقى اثر التغيرات الدورية لهذه

$$\zeta_t = \hat{T}(t) + C(t) + I(t) \rightarrow (8)$$

$$\tilde{\zeta}_t = \zeta_t - \hat{T}(t) = C(t) + I(t) \rightarrow (9)$$

والذي يمكن كتابته بالصورة: وتسمى هذه العلاقة باسم النموذج المختزل للسلسلة الزمنية، وبموجب الفرضيات الخاصة بالعنصر العشوائي، وحيث ($n < 30$) واهتزازات هذه السلسلة أقرب ما تكون إلى الاهتزازات التوافقية وبأدوار متساوية الطول ودورة كل منها تساوي $P = 5$ على وجه التقريب (وهو عدد القياسات في كل دورة) وبناء على ذلك فإن الطريقة المناسبة لتقدير العنصر الدوري هي طريقة جيب التمام لهالبرغ (Halberg Cosine method) والتي تقوم على عدد من الفرضيات (Hamilton and James, 1994)، وتكتب بالصيغة الآتية

$$C(t) = \mu + A \cos(\omega t + \psi) \rightarrow (10)$$

حيث μ : هي القيمة المتوسطة لقيم الفروق Residual Values وأن قيمتها نظرياً $\mu = 0$ لأن المعاملات \hat{a} , \hat{b} , \hat{c} اختيرت

بشكل يحقق المعادلة (2) حتى وإن كانت قيمتها العددية $\mu \neq 0$ و A : سعة الاهتزاز و ω : السرعة الزاوية للاهتزاز والتي تساوي

$$\omega = \frac{2\pi}{P} \text{ أي أن: } \omega = \frac{2\pi}{5} = 1.257 \text{ و } \psi: \text{ فرق المسار (الطور) بين السلسلة الزمنية ودالة جيب التمام، وحيث أن السرعة الزاوية}$$

للاهتزاز ω معلوم دائماً من بيانات العينة المأخوذة من مسار السلسلة الزمنية. ويتم تقدير المعامل μ , A , ψ بالاعتماد على قيم الفروق \tilde{X}_t بين القيم الفعلية X_t والقيم المقدرة \hat{X}_t وفقاً للمنحنى التوافقي الخاص بالبيانات تحت الدراسة، وسنقدم هنا تقديرات لمعامل العنصر الدوري مباشرة دون اللجوء إلى العمليات الحسابية المطولة، وبفرض n حجم العينة المأخوذة من مسار السلسلة الزمنية وأن القيمة المقدرة لـ μ , α , ψ هي:

$$\mu = -1.36 \times 10^{-7}, A = 93.524, \psi = -1.431 \text{ علماً بأن } \mu = -1.36 \times 10^{-7} \text{ تعتبر تافهة بالمقارنة مع بقية القيم}$$

وبالتعويض نجد أن الدالة المقدرة للمركبة الدورية $\hat{C}(t)$ هي:

$$\hat{C}(t) = -1.36 \times 10^{-7} + 93.524 \cos(\pi/2t - 1.431)$$

وبناءً على ما سبق نلاحظ أن الصيغة المثالة للنموذج المقدّر للطور (phase) العشوائي المولد للسلسلة الزمنية لأعداد الحوادث المرورية يأخذ

الصيغة الآتية:

$$\hat{\zeta}_t = 13.416t^2 - 36676.862t + 25038387.436 - 1.36 \times 10^{-7} + 93.524 \cos(\pi/2t - 1.431) + I(t)$$

ومن خلال استعراضنا لأحد المسارات المقدّرة للطور العشوائي المولد للسلسلة المعطاة نلاحظ أن القيمة المقدّرة للعام 1425 و 1434 وفقاً للنموذج العشوائي المقدّر للطور العشوائي المولد للسلسلة الزمنية هي على الترتيب (17281 و 32050) حيث كان الخطأ المعياري $S = 101.83$ والقيم الحقيقية لهذين العامين هي (19342 و 31279) ما زالتا تقعان ضمن التقدير الجيد، وفقاً للطور العشوائي المقدّر الذي قمنا بتعيينه، مما يزيد ثقتنا في استخدامنا لهذا النموذج في التنبؤ، والشكل (4) يوضح ذلك:



يمكن استخدام النموذج المقدر لطور العشوائى المولد للسلسلة للتنبؤ بمشاهدات السلسلة مستقلاً، ويعتبر التنبؤ في هذه الحالة من المستوى طويل الأمد، والذي يغطي فترة زمنية أكثر من خمس سنوات، ولهذا فإنه وعلى المدى المسموح به للتنبؤ وفي هذه الحالة (يجوز لنا استخدام هذا المنحنى للتنبؤ حتى نهاية فترة طولها يساوي نصف طول الفترة التي استخدمت للدراسة) ومن الشكل (5) يلاحظ أن عدد الحوادث المرورية في المنطقة تظهر تنامياً (تزايداً) واضحاً حتى على مدى السنوات المقبلة، والذي من المتوقع أن يزيد العدد إلى حوالي (49090) حادثة في عام 1443هـ.



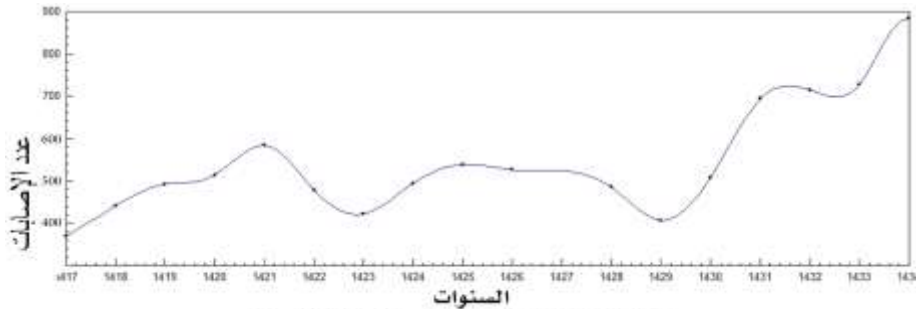
ثانياً: التحليل الزمني لأعداد الحوادث ذات الإصابات

بالرجوع إلى الإحصاءات السنوية الصادرة عن الإدارة العامة للمرور بمنطقة عسير خلال الفترة من (1417-1434) فإن الجدول (5) يبين الإحصاءات السنوية لأعداد الحوادث المرورية ذات الإصابات لمنطقة عسير

جدول (5): أعداد الحوادث المرورية ذات الإصابات

العام	1417	1418	1419	1420	1421	1422	1423	1424	1425
عدد الإصابات	369	441	492	513	584	479	422	494	538
العام	1426	1427	1428	1429	1430	1431	1432	1433	1434
عدد الإصابات	526	470	485	407	508	694	714	727	884

العرض البياني الممثل بالشكل (6) يوضح الانتشار للسلسلة الزمنية الخاصة بأعداد الحوادث المرورية التي حدثت فيها إصابات، يشير إلى وجود اتجاه عام Trend متزايد للسلسلة الزمنية يتوافق مع السلوك العام لأعداد الحوادث المرورية. حيث نلاحظ ظهور حركة اهتزازية اقرب ما تكون للاهتزازات التوافقية والتي تمثلها المركبة الدورية للسلسلة الزمنية، في حين أن المركبة الموسمية لا وجود لها في شكل الانتشار، وأن النموذج التجميعي هو النموذج المناسب لعرض هذه السلسلة الزمنية (Granger,1977).



شكل (6) الانتشار لإعداد الحوادث ذات الإصابات

وبلاحظ من خلال الشكل(6) أن أعداد الإصابات بدءاً من العام 1430 بدأت تتضاعف عن الأعوام السابقة وهذا يعود إلى أن إدارة المرور قد طورت نظام رصد الحوادث من خلال أمن الطرقات وكمرات المراقبة، والدوريات الأمنية. واختبار الارتباط العشوائي للبيانات قيد الدراسة نستخدام طريقة الفروق الأولى عند مستوى معنوية $\alpha = 0.05$ لإزالة أثر الاتجاه العام. ومن جدول(6) نلاحظ أن عدد الإشارات الموجبة(+) يساوي $k = 12$ فإن قيمة الفروق المحسوبة لهذا الاختبار يمكن الحصول عليها من العلاقة رقم(7). وبما أننا لاحظنا وجود اتجاه عام للسلسلة الزمنية فإننا سوف نستخدم اختبار ذو اتجاه واحد، وتكون المنطقة الحرجة عند مستوى معنوية 5% هي 1.64

وبالمقارنة نجد أن $Z_{\alpha} < TG$ وبناء على ذلك نرفض الفرضية الصفرية H_0 وقبل بالفرضية البديلة H_1 عند مستوى معنوية $\alpha = 0.05$ أي أن البيانات قيد الدراسة مرتبطة عشوائياً، ويمكننا بحث هدوء(سكون) السلسلة الزمنية قيد الدراسة. جدول (6): الفروق الأولى(الفروق المتتالية)

العام	1417	1418	1419	1420	1421	1422	1423	1424	1425
عدد الإصابات	369	441	492	513	584	479	422	494	538
$sig(x_{t+1}, x_t)$		+	+	+	+	-	-	+	+
العام	1426	1427	1428	1429	1430	1431	1432	1433	1434
عدد الإصابات	526	470	485	407	508	694	714	727	884
$sig(x_{t+1}, x_t)$	-	-	+	-	+	+	+	+	+

ومن خلال تفحص معاملات الارتباط الذاتي $\hat{\rho}_k$ الموضحة في جدول(7) لبيانات العينة الممثلة لمسار السلسلة الزمنية قيد الدراسة، وتطبيق دالة الارتباط الذاتي ACF بواسطة برنامج SPSS على بيانات الجدول (5) نلاحظ أن دالة الارتباط الذاتي تتناقص بشكل ملحوظ؛ وهذا دليل على هدوء السلسلة الزمنية، وإذا أردنا اختبار انعدام معاملات الارتباط الذاتي عند الفجوة الأولى $k = 1$ بفترة ثقة 95% تكون القيمة المحسوبة 3.084 والقيمة الجدولية 1.96، وحيث أن القيمة المحسوبة أكبر من القيمة الجدولية نرفض فرضية انعدام معاملات الارتباط عند $k = 1$ ، وبالمثل عند $k = 5$ نجد أن القيمة المحسوبة 0.127 وهذا يعني انعدام معاملات الارتباط الذاتي عند 5 درجات، وبذلك يمكننا القول بهدوء السلسلة الزمنية قيد البحث بعد الفجوة الرابعة، أي أن U_t مستقلة عن U_{t-1} .

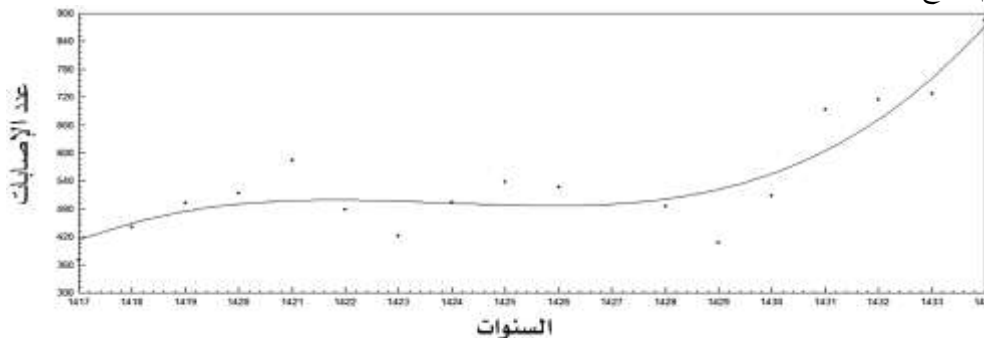
جدول(7): معاملات الارتباط الذاتي

Lag	1	2	3	4	5	6	7	8
Auto Corr	0.727 (0.000)	0.465 (0.000)	0.245 (0.001)	0.072 (0.001)	0.030 (0.003)	-0.015 (0.007)	-0.067 (0.013)	-0.094 (0.020)
Lag	9	10	11	12	13	14	15	16
Auto Corr	-0.110 (0.029)	-0.130 (0.036)	-0.164 (0.036)	-0.190 (0.028)	-0.214 (0.01)	-0.300 (0.002)	-0.299 (0.000)	-0.268 (0.000)

ولتقدير الطور(phase) العشوائي المولد للسلسلة الزمنية نقوم بتقدير العناصر التي يلاحظ وجودها.

تقدير عنصر الاتجاه العام للسلسلة الزمنية

بنفس الأسلوب السابق سنقوم بتقدير عنصر الاتجاه العام $T(t)$ ومن خلال الشكل(7) نلاحظ أن أحسن خط انحدار لبيانات السلسلة الزمنية باستخدام برنامج Curve(Data fit) هو:



شكل (7) مركبة الاتجاه العام لأعداد الحوادث ذات الاصابة

$$T(t) = at^4 + bt^3 + ct^2 + dt + e \rightarrow (11) \quad \text{والذي معادلته}$$

حيث أوضحت نتائج التحليل أن القيم التقديرية لـ a, b, c, d, e هي:

$$\hat{a} = 3.8441 \times 10^{-3}, \quad \hat{b} = -21.52086, \quad \hat{c} = 45167.999,$$

$$\hat{d} = -42119619.28, \quad \hat{e} = 1472412340.03$$

وبناء على ذلك فإن نموذج الاخطي يأخذ الصورة:

$$\hat{T}(t) = 3.8441 \times 10^{-3} t^4 - 21.52086 t^3 + 45167.999 t^2 - 42119619.28 t + 1472412340.03$$

وبخطأ معياري لمعادلة الانحدار مقداره: $S = 83.04$ ومعامل ارتباط $r = 0.918$ ومعامل التحديد $R^2 = 0.842$ ، وأن كفاءة النموذج حسب اختبار ديرن-وتسن الإحصائي هو: $d = 1.56$ وأن $d_u = 1.39$ ، $d_L = 1.16$ ، وحيث أن $d_U < d$ ومن ثم فلا يوجد لدينا ما يدفعنا لرفض الفرضية الصفرية H_0 عند مستوى معنوية $\alpha = 0.05$. ومن العلاقة السابقة نلاحظ أن الاتجاه العام لهذه السلسلة الزمنية لها شكل دالة اهتزازية، ومتزايدة بدرجة ملحوظة خلال السنوات العشر القادمة.

تقدير العنصر الدوري للسلسلة الزمنية

وبشكل مماثل لما سبق يمكن تقدير العنصر الدوري، وتحت الفرضيات الخاصة بالعنصر العشوائي أن لهذا العنصر العرض الآتي وفقاً لطريقة التجيب ل هالبرغ، وبافتراض أن العنصر الدوري العرض

$$p(t) = \mu + \alpha \cos(\omega t + \psi) \quad \text{فنجد أن قيمة مقدر } \mu \text{ هي: } \hat{\mu} = -0.0001847 \text{ علماً بأن}$$

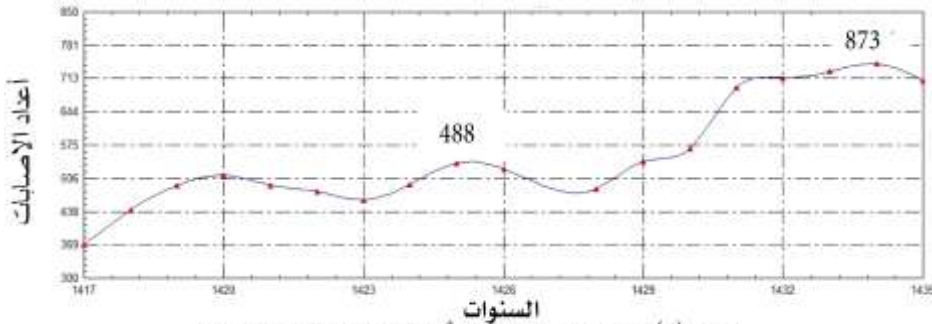
$$w = \frac{2\pi}{p} = 1.257 \quad \text{وأن قيمة } \hat{\alpha} = 18.735, \quad \hat{\psi} = 4.632 \quad \text{وبالتعويض في صيغة المركبة الدورية نجدها تعرض حسب الصيغة الآتية:}$$

$$\hat{p}(t) = -0.0001847 + 18.735 \cos(1.257t + 4.632)$$

وبالتعويض في المعادلة (8) نجد أن الصيغة المثلة للنموذج المقدر للطور العشوائي المولد للسلسلة الزمنية الخاص بأعداد الحوادث ذات الإصابات يأخذ الصورة الآتية:

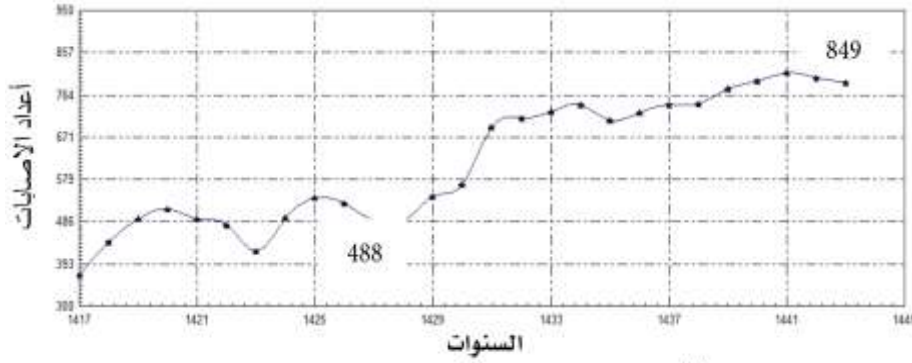
$$\hat{Y}_t = 3.8441 \times 10^{-3} T^4 - 21.52086 T^3 + 45167.999 T^2 - 42119619.28 T + 1472412340.3 - 0.0001647 + 18.735 \cos(1.257t + 4.632)$$

وبناء على ما سبق يصبح لأحد المسارات المقدر للطور العشوائي المولد للسلسلة قيد البحث بالشكل (8)



شكل (8) أحد المسارات المقدر لأعداد الحوادث ذات الإصابات

ومن خلال استعراضنا لأحد المسارات المقدر للطور العشوائي المولد للسلسلة الزمنية الموضح في الشكل (9) نلاحظ أن القيم المقدر ما زالت تقع ضمن التقدير الجيد، وفقاً للطور العشوائي المقدر الذي قمنا بتعيينه، مما يزيد ثقتنا في استخدامنا لهذا النموذج في التنبؤ.



شكل (9) النموذج المقدر للتنبؤ بأعداد الحوادث ذات الإصابات

ونلاحظ مما سبق أن أعداد المصابين في الحوادث المرورية بالمنطقة يظهر تنامياً (تزايداً) واضحاً حتى على مدى السنوات المقبلة، والذي من المتوقع أن يزيد العدد إلى حوالي (849) حادثة في حلول عام 1443هـ (بفرض استقرار الظروف).

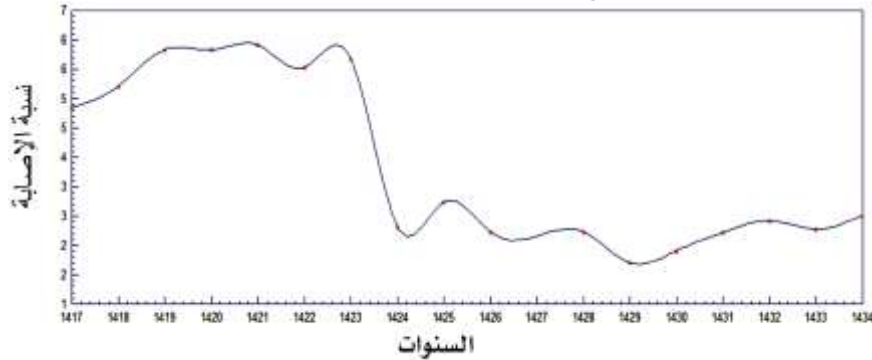
ثالثاً: التحليل الزمني للنسبة المئوية لأعداد الحوادث ذات الإصابات

من العوامل المهمة في مثل هذه المسائل معرفة النسب المئوية لأعداد الحوادث ذات الإصابات لأنها تعطينا تصوراً واضحاً حول طبيعة الحوادث في هذه المرحلة ومدى توفر وسائل الحماية والأمن والسلامة، ونوعية المركبة، وخدمات الطرق. والجدول (8) يشير إلى النسب المئوية لأعداد الحوادث ذات الإصابات.

جدول (8): النسب المئوية لأعداد الحوادث التي حدثت ذات الإصابات.

العام	1417	1418	1419	1420	1421	1422	1423	1424	1425
نسبة الإصابات	5.019	5.448	6.203	6.200	6.298	5.833	6.015	2.581	3.089
العام	1426	1427	1428	1429	1430	1431	1432	1433	1434
نسبة الإصابات	2.469	2.528	2.478	1.843	2.092	2.462	2.707	2.528	2.826

العرض البياني في الشكل (10) الخاص بنسبة الإصابة يشير إلى أن الاتجاه العام لهذه السلسلة الزمنية متذبذباً بمدى ثابت وتكرر ثابت، كما يبين ظهور حركة اهتزازية أقرب ما تكون للاهتزازات التوافقية. أي أنه يشير إلى أن النسبة المئوية لأعداد الحوادث ذات الإصابات بدأت من العام 1434هـ تتناقص عن الأعوام السابقة بشكل ملحوظ، وهذا قد يرجع إلى توافر بعض العوامل المؤثرة مثل وسائل الحماية وخدمات الطرق.



شكل (10) الانتشار لنسبة الحوادث ذات الإصابات

و يشير جدول (9) إلى الارتباط العشوائي للبيانات باستخدام طريقة الفروق الأولى لإزالة أثر الاتجاه العام.

جدول (9): الفروق الأولى (الفروق المتتالية)

العام	1417	1418	1419	1420	1421	1422	1423	1424	1425
نسب الإصابات	5.019	5.448	6.203	6.200	6.298	5.833	6.015	2.581	3.089
$sig(x_{t+1}, x_t)^*$		+	+	-	+	-	+	-	+
العام	1426	1427	1428	1429	1430	1431	1432	1433	1434
نسب الإصابات	2.469	2.528	2.478	1.843	2.092	2.462	2.707	2.528	2.826
$sig(x_{t+1}, x_t)^*$		-	+	-	-	+	+	+	-

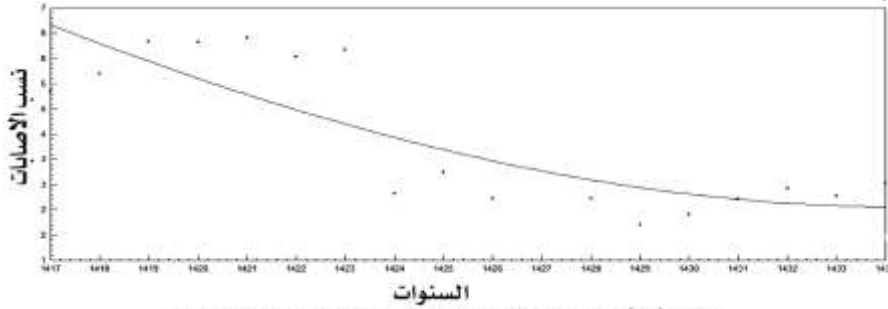
ومن الجدول (9) نلاحظ أن عدد الإشارات الموجبة (+) يساوي $10 = \zeta$ ، علماً بأن القيمة المحسوبة للفروق الأولى وفقاً للمعادلة (7) هي $TG = 0.7947$ وتكون المنطقة الحرجة عند مستوى معنوية 5% هي 1.64. وبالمقارنة نجد أن $TG < Z_\alpha$ ومن ثم ليس لدينا ما يدفعنا لرفض الفرضية الصفرية H_0 عند مستوى $\alpha = 0.05$. وكما سبق ومن خلال فحص معاملات الارتباط الذاتي المتناقصة بشكل ملحوظ كما في جدول (10)، وبنا نستطيع القول بهدوء السلسلة الزمنية.

جدول (10): يوضح معاملات الارتباط الذاتي

Lag	1	2	3	4	5	6	7	8
Auto Corr	0.835 (0.000)	0.705 (0.000)	0.485 (0.001)	0.280 (0.000)	0.061 (0.000)	-0.138 (0.000)	-0.289 (0.000)	-0.338 (0.020)
Lag	9	10	11	12	13	14	15	16
Auto Corr	-0.372 (0.000)	-0.392 (0.000)	-0.390 (0.000)	-0.322 (0.000)	-0.246 (0.00)	-0.176 (0.000)	-0.115 (0.000)	-0.063 (0.000)

تقدير عنصر الاتجاه العام للسلسلة الزمنية

وعلى نحو مماثل لما سبق سنقوم بتقدير عنصر الاتجاه العام $T(t)$ فنلاحظ من خلال الشكل (11) أن أفضل خط انحدار لبيانات السلسلة الزمنية باستخدام برنامج Curve (Data fit) هو:



شكل (11) مركبة الاتجاه العام لنسبة الحوادث ذات الاصابة

والذي معادلته $(12) \rightarrow T(t) = aX^3 + bX^2 + cx + d$ حيث أوضحت نتائج التحليل أن القيم التقديرية هي: $\hat{d} = -436548.596$ ، $\hat{c} = 942.0416$ ، $\hat{b} = -0.67406$ ، $\hat{a} = 1.6066$ وبناء على ذلك فإن النموذج اللاخطي يأخذ الصورة: $\hat{T}(t) = 1.6066T^3 - 0.6741T^2 + 942.042T - 438548.596$ وبخطأ معياري $S = 23.04$ ومعامل ارتباط $r = 0.837$ ومعامل التحديد $R^2 = 0.6998$ ، وأن كفاءة النموذج حسب اختبار ديرين-وتسون الإحصائي هو $d = 1.001$ وأن $d_u = 1.39$ ، $d_L = 1.16$ ، وحيث أن $d < d_L$ ومن ثم فلا يوجد لدينا ما يدفعنا لقبول الفرضية الصفرية H_0 عند مستوى معنوية $\alpha = 0.05$. أي أنه لو استمر عدد الإصابات بالحوادث المرورية على النحو الذي تم في الماضي فإن نسبة الإصابات ستتضاعف خلال العقد المقبل، نتيجة توافر بعض العوامل المؤثرة مثل وسائل الحماية، وخدمات الطرق.

تقدير العنصر الدوري للسلسلة الزمنية

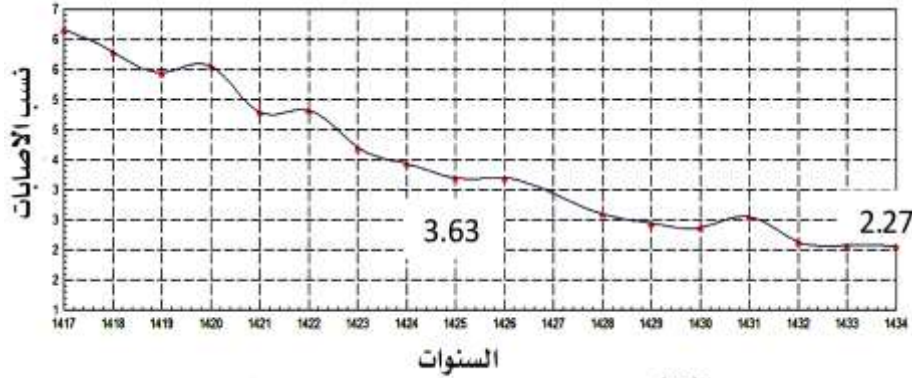
بشكل مماثل لما سبق يمكن تقدير العنصر الدوري، وتحت نفس الفرضيات نلاحظ أن الدالة المقدرة للمركبة الدورية لها الشكل الآتي:

$$\hat{p}(t) = 2.2 \times 10^{-6} + 1.351 \cos(1.257t - 3.172)$$

وبالتالي يصبح المقدر الطور العشوائي المولد للسلسلة الزمنية الصيغة الرياضية الآتية:

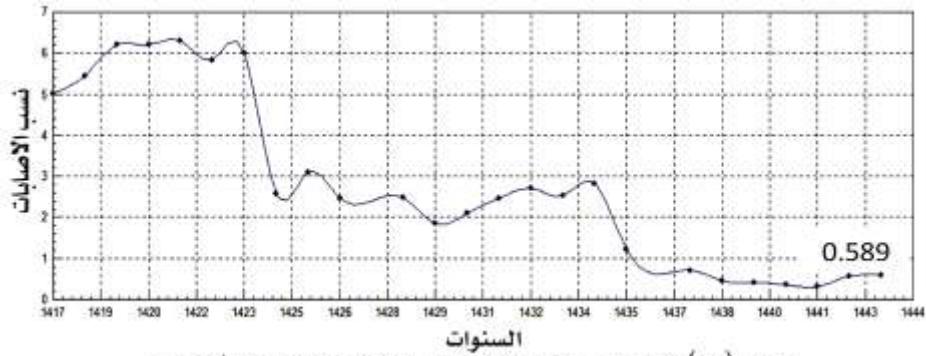
$$\hat{\zeta}_t = 1.61X^3 - 0.67X^2 + 942.04X - 438548.6 + 2.2 \times 10^{-6} + 1.35 \cos(1.257t - 3.172)$$

ومن ثم يصبح لأحد المسارات المقدرة للطور العشوائي المولد للسلسلة الزمنية قيد البحث بالشكل (12):



شكل (12) أحد المسارات المقدرة لنسبة الحوادث ذات الاصابة

وكما سبق سنستعرض أحد المسارات المقدرة للطور العشوائي المولد للسلسلة الزمنية في الشكل (13) نلاحظ أن القيم المقدرة ما زالت تقع ضمن التقدير الجيد، وفقاً للطور العشوائي المقدّر الذي قمنا بتعيينه، مما يزيد ثقتنا في استخدامنا لهذا النموذج في التنبؤ على المدى المسموح به للتنبؤ.



شكل (13) النموذج المقدّر للتنبؤ بنسبة الحوادث ذات الاصابة

ومما سبق نلاحظ أن نسبة المصابين في الحوادث المرورية يظهر تناقصاً واضحاً على مدى السنوات المقبلة، والذي من المتوقع أن تتضاءل نسبة الإصابة إلى حوالي 0.589 حادثة في حلول عام 1443هـ.

4.4 الاستنتاجات

نخلص إلى أن الحوادث المرورية بمنطقة عسير كانت خلال الفترة الزمنية 1417-1434 تسير بشكل متناهي في معظم الحالات، إلا أن هذا السلوك المتناهي أشار إلى مستقبل غير متفائل في أعداد الحوادث، نتيجة الارتفاع المقدّر إذا استمر الوضع على ما كان عليه في الماضي، إلا أنه متفائل بخصوص أعداد ونسبة الإصابات في الحوادث، نتيجة لتوفر وسائل الأمن والسلامة، والذي تجلّى من خلال الصيغ الممثلة لمقدرات الاتجاه العام لكل من السلسلة الزمنية لأعداد الحوادث، والإصابات وكذا نسبة الإصابة، كما أظهرت النتائج وجود عوامل لها دور فاعل في زيادة أعداد الحوادث منها عدم الوعي بأخطار الحوادث، كما أدى توفر وسائل الحماية، وتطور المركبات إلى تناقص أعداد ونسبة الإصابة بشكل ملحوظ.

5. المراجع

1. الإحصاء السنوي للفترة (1417-1434). الإدارة العامة للمرور، منطقة عسير.
2. عبد الحميد البلداوي. (1997): الإحصاء للعلوم الإدارية والتطبيقية. ط1، دار الشروق، عمان.الأردن.
3. وليد دهان الجبوري. (2009): التنبؤ بمستوى التضخم في أسعار المستهلك الشهرية في العراق باستخدام السلاسل الزمنية ثنائية المتغيرات. رسالة ماجستير في الإحصاء، جامعة المستنصرية.
4. الشرق الأوسط. (2015): جريدة العرب الدولية، العدد (7981)، المملكة العربية السعودية، طبعة السعودية.
5. سمير مصطفى شعراوي (2005): مقدمة في التحليل الحديث للسلاسل الزمنية. جامعة الملك عبد العزيز، جدة: السعودية.
6. سعدية عبد الكريم طعمة. (2012): استخدام السلاسل الزمنية لتنبؤ بأعداد المصابين بالأورام الخبيثة في محافظة الأنبار. مجلة جامعة الأنبار للعلوم الاقتصادية والإدارية، المجلد 4 العدد 8.
7. حميد العكلة. (2009): تحليل إحصائي لطلاب دراسات الماجستير في الجامعات السورية بين (1991-2006). مجلة جامعة دمشق، المجلد 25، العدد 1، علوم أساسية، سوريا.

8. والتر فاندل. ترجمة عبدالمرضي حامد عزام، أحمد حسين هارون.(1992): السلاسل الزمنية من الوجهة التطبيقية ونماذج بوكس-جنكيز - ، دار المريخ، الرياض، السعودية.
9. عصام كوثر وآخرون.(2013): الآثار المترتبة على الحوادث المرورية وسبل تقليصها تاريخ استرجاع 24/2/2014.نشر بموقع: <http://www.al-jazirahonline.com/news/2014/20140224/13208>
10. انتصار أوتله محمد.(2010): استخدام تحليل السلاسل الزمنية لبناء نماذج الحركة لولاية الخرطوم. رسالة ماجستير في الإحصاء، كلية العلوم- جامعة السودان للعلوم والتكنولوجيا.
11. محمد عبدالعال النعمي ومحمد حبيب الشاروط.(2000): استخدام التدخل في السلاسل الزمنية لتقدير عدد المصابين بالأورام الخبيثة. مجلة القادسية للعلوم الإدارية والاقتصادية، المجلد 3 العدد 1. العراق.
12. Box G.E.P, and Jenkins, G.M .1976. Time Series Analysis Forecasting and Control (Rev. Ed.). San Francisco: Holden-Day, Inc. First Ed., 19.
13. Brock Well , P.J. and Davis, R.A .1991. Time Series Theory and Methods, 2nd ed , Springer Verlag New York Inc , New York .
14. Clive, W. and Granger .2004. Time Series Analysis Cointegration , and Applications. Paper 2004-02 University of California, San Diego.
15. George C. and Canavos, C .1984. Applied probability and Statistical Methods. Little , Brown and Company-Boston, Toronto 608.
16. Granger C.W.J .1977. Comment on Relationships-and the Lack Thereok- Between Economic Time Series with Special Reference to Money and Interest Rates. by D.A. Pierce. Journal of the American Statistical Association, 72, 357 (March),22-23.
18. Hamilton, J. and James .1994. Time Series Analysis, Princeton U. press, Princeton, NJ ISBN 0-691-04289-6[SJS].
19. Kirchgssner G. and Wolters J .2007. Introduction to Modern Time Series Analysis. SPRINGER-Verlag, Berlin Heidelberg.
20. Larsen, R. J. and Marx, M.L .1993. An introduction to Mathematical Statistics and its Applications. Prentice-Hall International, Inc. USA596.
21. Maestas, C. and Robert, R .2000. Modeling Volatility in Political Time Series. Electoral Studies 19:95–110.
22. Mosteller, F., and Tukey, J.W .1977. Data Analysis and Regressions :A Second Course in Statistics. Reading, MA: Addison-Wesley.
23. Murray, R.and Spiegel .2003.Theory and problems of statistics.Mc GRAW Hill Edition, New York.
24. Quenouille, N.H .1957. Analysis of Multiple Time Series. New York, Hafner Publishing Co.
- Shumway, R.H. and Stoffer, D.S .2006. Time Series Analysis and Its Applications. SPRINGER, New York.

**Statistical Indicators of Traffic Accidents Using Time Series Analysis in Asir (KSA) during (1417-1434) Hijri
Mohammed M.A. Almazah⁽¹⁾ and Esmail H. A. Alsabri⁽²⁾**

Mathematics Dept., College of sciences and Arts (Muhiyl), King Khalid Unvi., KSA; Ibb Unvi., Yemen.

ABSTRACT

Traffic accidents constitute one of the biggest safety threats facing community and have a big impact on its material and human resources. This research provides a time-based statistical analysis of the number of traffic accidents in Asir Region during 1417-1434 hijri (equivalent to 1996-2013), with a focus on determining appropriate probability distributions of the data. It then gives an estimate of the random phase generating the time series in order to reveal the current condition of accidents number, predict future expectations, and anticipate the future behavior of these phases based on a prediction model. The scatter diagram of the time series has revealed a vibrating movement somehow close to harmonic vibrations which represent a cyclical compound and also lack of seasonal compound. It has also been revealed that the appropriate model to display the time series data is the aggregation model. The general trend compound came in the form of non-linear models: that is, the prediction model is non-linear. The results of this research can be considered as an assessment of an important period of time in the history of the rate of traffic accidents and casualties in the region. It has been observed that the number of traffic accidents and casualties has been growing, and is expected to increase to about 66173 incidents by 1443 hijri (2021), despite the expected decrease in the rate of traffic casualties. In addition, the study has pointed out factors which have an active and influential role in traffic accidents rates. It is hoped that this study- in its both theoretical and practical sides- would encourage researchers to apply various predictive models to discuss this subject.

Key Words: Aggregation model, Casualties rate, Harmonic oscillations

Nonlinear Time Series, Traffic accidents.

جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوى الخمسين للإحصاء
وعلوم الحاسب وبحوث العمليات

إحصاء سكانى

٢٥-٢٧ ديسمبر ٢٠١٧

فهرس الإحصاء السكاني

٣٤-١	خريطة الإعاقة فى مصر ٢٠١٦ سها أحمد حسن ، مصطفى ربيع	١
٦٤-٣٥	العوامل المؤثرة على نية ريادة الأعمال بين الشباب المصرى إيمان أحمد إسماعيل	٢
٨٣-٦٥	دوافع المهاجرين الحاليين والعائدين فى مصر محمد مصطفى حسن ، محمد عبد المنعم عبد الرؤف ، أيمن زهرى	٣

خريطة الإعاقة فى مصر 2016

مصطفى ربيع²

سها أحمد حسن متولى¹

المُلخَص

مازال توفر إحصاءات دقيقة عن الإعاقة يشكل تحدياً كبيراً يواجه العديد من الدول ومنها مصر، وهو ما يعوق عملية صنع القرار المبنية على المعلومات والأدلة، ويهدف هذا البحث إلى رسم خريطة الإعاقة فى مصر والتعرف على مستويات شدة الإعاقة المختلفة (بسيطة/ شديدة/ تامة)، وأكثر أنواع الإعاقات انتشاراً فى مصر وفقاً لمجموعة أسئلة مجموعة عمل واشنطن المعنى بإحصاءات الإعاقة المختصرة، بالإضافة إلى التعرف على خصائص ذوي الإعاقة بحسب الحالة التعليمية والحالة العملية والحالة الزوجية، والمستوى المعيشي للأسرة. ويعتمد البحث على بيانات المسح القومى "مسح أحوال الأسرة المصرية لعام 2016- الدورة 13"، والذي نفذه مركز المعلومات ودعم اتخاذ القرار، وتبلغ عينة الدراسة 49431 فرد على مستوى الجمهورية. وقد أظهرت نتائج البحث أن نسبة الإعاقة حسب تعريف بسيطة فأكثر فى مصر تبلغ 10.4% فى عام 2016 من إجمالى السكان، ولا يوجد اختلاف فى انتشار الإعاقة حسب النوع أو حسب محل الإقامة حضر/ريف، بينما تتباين حسب المناطق الجغرافية، وتشكل الإعاقات البسيطة النسبة الأكبر من الإعاقات، كما تعد الإعاقات الخاصة بالرؤية أكثر الإعاقات التي يعاني منها الأفراد فى مصر، يليها الحركة ثم السمع، وترتفع نسب الإعاقة بشكل ملحوظ مع ارتفاع العمر. كما أظهرت النتائج أن الإعاقة لها تأثير سلبي على تمتع الأفراد ذوي الإعاقة بحقوقهم، سواء فى التعليم أو العمل أو الزواج، كما يعيش الأفراد ذوي الإعاقة فى أسر فقيرة مقارنة بباقي بأسر الأفراد غير المعاقين. وجدير بالذكر إلى أنه مازال حاجة إلى إجراء مسوح قومية تتضمن بيانات أكثر تفصيلاً عن المعاقين وظروف إعاقاتهم، وهو ما يسمح برسم خريطة أكثر تفصيلاً عن أوضاع ذوي الإعاقة فى مصر.

الكلمات الدالة: الإعاقة – أنواع الإعاقة – مستويات الإعاقة - مصر

مقدمة

يُقدَّر عدد الأفراد ذوي الإعاقة فى العالم بحوالى مليار فرد، أي ما يُعادل 15% من سكان العالم (World Health Organization 2011)، وهذه الأعداد فى تزايد مستمر خاصة مع تزايد أعداد المسنين فى العالم، وانتشار الأمراض غير المعدية، وزيادة وتيرة حوادث الطرق، والكوارث الطبيعية (United Nations 2012). وتكمن خطورة الإعاقة فى تأثيرها على جودة حياة الأفراد ذوي الإعاقة، لذا نالت قضايا الإعاقة الكثير من الاهتمام على المستوى الدولى، وقد توجت

¹ مدرس بقسم الإحصاء الحيوى والسكانى – معهد الدراسات والبحوث الإحصائية – جامعة القاهرة

² باحث إحصائى – مركز المعلومات ودعم اتخاذ القرار – رئاسة الوزراء

تلك الجهود بإعلان الاتفاقية الدولية لحقوق الأفراد ذوي الإعاقة في عام 2006 (الأمم المتحدة 2006)، والتي أكدت على كافة الحقوق الإنسانية التي يجب أن يتمتع بها الأفراد ذوي الإعاقة. كما أكدت أهداف التنمية المستدامة 2015 – 2030 على حقوق المعاقين، وحددت أهداف تضمن لهم الحياة الكريمة وتكافؤ الفرص والعدالة في الحصول على الخدمات (الأمم المتحدة 2015).

وفي مصر تعاني إحصاءات الإعاقة من قصور في التقدير، وكذلك التباين الكبير في التقديرات عند المقارنة بالدول الأخرى والذي يرجع إلى عدة أسباب من بينها الاختلاف في أساليب القياس، إلى جانب سيطرة التقاليد التي تميل إلى عدم الرغبة بالإدلاء عن وجود إعاقات بين الأفراد. ويعد الوقوف على حجم وأنواع الإعاقات أحد القضايا الهامة من منطلق مدخلين أساسيين، الأول يرتبط بكون الإعاقة تُعد أحد قضايا حقوق الإنسان، حيث يعاني الأفراد ذوي الإعاقة من عدم المساواة (World Health Organization 2011)، والثاني يرتبط بكون الإعاقة تُعد أحد قضايا التنمية (Mitra et al 2011).

ويُعدُّ توفر بيانات دقيقة وحديثة عن حجم ذوي الإعاقة وشدة الإعاقة أمر هام لأنه يضمن بقاء ذوي الإعاقة على أجندة صانعي القرار، وتساعد هذه البيانات في صياغة برامج الاستهداف المستقبلية وذلك من خلال إلقاء الضوء على أوجه الحرمان الرئيسية التي يعاني منها الأفراد ذوي الإعاقة والتي تتطلب اهتماماً فورياً.

ويهدف البحث إلى رسم خريطة الإعاقة في مصر من خلال التعرف على معدلات انتشار الإعاقة في مصر وتبايناتها حسب محل الإقامة (حضر/ريف)، وحسب الأقاليم الجغرافية، وحسب المحافظات المختلفة، وحسب النوع، وحسب الفئات العمرية، بالإضافة إلى التعرف على شدة الإعاقة (بسيطة/ شديدة/ تامة)، وأكثر أنواع الإعاقات انتشاراً في مصر وفقاً للفئات المشار إليها حسب تعريف مجموعة عمل واشنطن المعنى بإحصاءات الإعاقة (الرؤية، والسمع، والحركة، والتذكر أو التركيز، والرعاية الشخصية، والتواصل مع الآخرين). كما يهدف البحث إلى التعرف على الملامح الأساسية للمعاقين بحسب الحالة التعليمية والحالة العملية والحالة الزوجية، وكذلك حسب المستوى المعيشي للأسرة.

1- مراجعة الأدبيات

تنعدد المفاهيم الخاصة بالإعاقة والتي تؤدي إلى اختلاف أسلوب القياس ومن ثم مستويات مختلفة للتقدير. ونوضح فيما يلي أكثر مفاهيم الإعاقة انتشاراً، وأساليب قياسها:

1-1 مفهوم الإعاقة

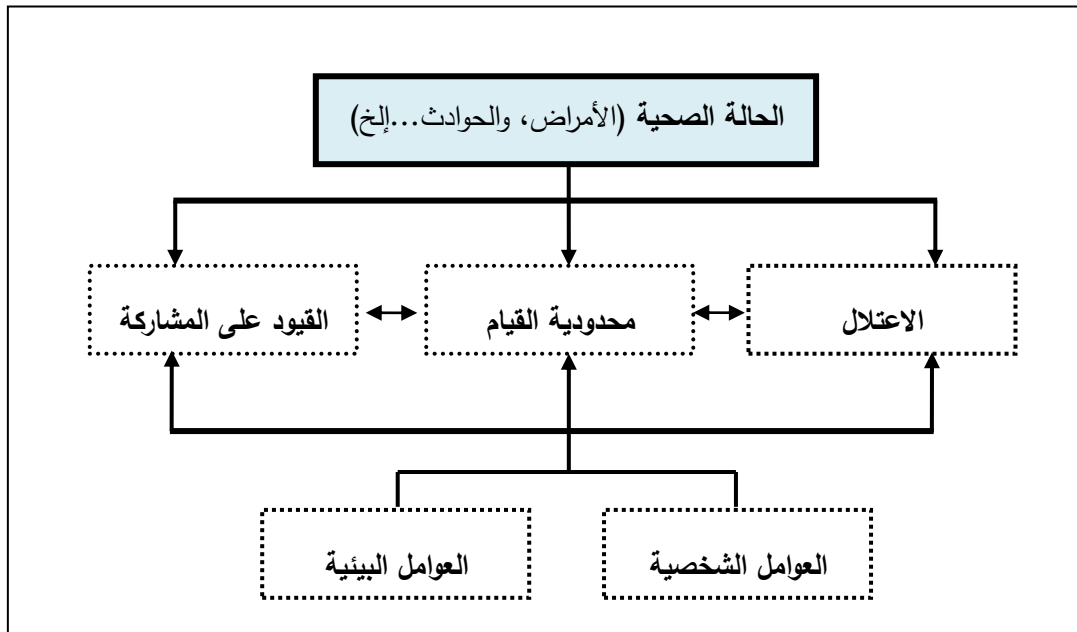
يعتبر مفهوم الإعاقة من المفاهيم المعقدة، وما زال قيد التطور (الأمم المتحدة 2006)، ومع ذلك فإن الإعاقة عادة ما تشير إلى حالة جسدية أو عقلية أو نفسية تؤثر وتحد من أنشطة الأفراد (Mont 2007). وقد عرفت منظمة الصحة العالمية (World Health Organization 2011) الإعاقة بوصفها مصطلح شامل يضم كل من:

- الاعترال "Impairments": وهي مشاكل في وظائف الجسم أو بنية الجسم مثل الانحراف عن الوضع الطبيعي.

- محدودية القيام بالأنشطة "Activity Limitations": الأنشطة هي المهام أو الأعمال التي يقوم الفرد بتنفيذها، وتعتبر محدودية القيام بالأنشطة عن الصعوبات التي تواجه الفرد في تنفيذ الأنشطة المختلفة، مثل إرتداء الملابس، أو تناول الطعام.
- القيود على المشاركة "Participation Restrictions": يقصد بالمشاركة إندماج الفرد في الحياة الاجتماعية والاقتصادية، مثل عدم القدرة على الذهاب إلى المدرسة، أو الحصول على فرصة عمل.

وتتأثر هذه الأبعاد الثلاث (الاعتلال، ومحدودية القيام بالأنشطة، والقيود على المشاركة) بكل من عوامل البيئة وعوامل الشخصية، وتضم الأولى البيئة الاجتماعية، والسلوكية التي يعيش فيها الأفراد. فالاعتلال قد لا يؤدي بالضرورة إلى محدودية القيام بالأنشطة والمشاركة في المجتمع، إذا كانت البيئة المحيطة مهيئة لاستيعاب الأفراد ذوي الإعاقة. أما الثانية فتضم الخصائص الديموجرافية والاجتماعية والاقتصادية للأفراد، مثل النوع والعمر والحالة التعليمية والمستوى الاقتصادي. فعلى سبيل المثال الإناث ذوي الإعاقة قد يتضاعف التمييز ضدهن، كذلك ذوي الإعاقة في الطبقات الفقيرة قد لا يستطيعون الاندماج في المجتمع بشكل كامل مقارنة بذوي الإعاقة في الطبقات الغنية.

ويوضح الشكل رقم (1) التصنيف الدولي لتأدية الوظائف والاعاقة والصحة، ويبدأ هذا النموذج مع حالة صحية "A Health Condition" (على سبيل المثال: الأمراض، والإصابات والحوادث، والظروف الصحية الأخرى ذات الصلة) التي تتفاعل مع العوامل المحيطة (العوامل الشخصية والاجتماعية) فينتج عنها كل من الاعتلال، ومحدودية النشاط، وصعوبة المشاركة.



Source: World Health Organization and The World Bank. 2011. World report on disability 2011. Malta: World Health Organization.

1- 2 أساليب قياس الإعاقة

تختلف الأساليب المستخدمة في قياس الإعاقة في التعدادات والمسوح بين الدول وبعضها البعض، وأحياناً داخل الدولة الواحدة، ونستعرض فيما يلي أهم هذه الأساليب:

- **الإدلاء الذاتي بالإعاقة "Self-identification as disabled"**: حيث يتم سؤال الأفراد بشكل مباشر "هل أنت أو أي أفراد الأسرة معاق؟".
- **حالات التشخيص "Diagnosable conditions"**: حيث يتم سؤال الأفراد عن ما إذا كانوا يعانون من عدد من الحالات الصحية مثل شلل الأطفال، أو الصرع...إلخ.
- **أنشطة الحياة اليومية "Activities of Daily Living (ADL)"**: حيث يتم تصنيف الأفراد على أنهم ذوي إعاقة إذا كانوا يعانون من صعوبة في أداء أياً من أنشطة الحياة اليومية، مثل ارتداء الملابس، والاستحمام، وتناول الطعام...إلخ.
- **أنشطة فعالة في الحياة اليومية "Instrumental Activities of Daily Living (IADL)"**: وهو مشابه للأسلوب السابق ولكنه يتضمن أنشطة على نطاق أوسع مثل إذا كان الفرد لديه مشكلة في إدارة الأموال، أو التسوق في المحلات...إلخ.
- **المشاركة**: يتم سؤال الأفراد وفقاً لهذا الأسلوب إذا كان لديهم أي حالة صحية تؤثر على مشاركتهم في المجتمع، مثل الذهاب للمدرسة أو الحصول على وظيفة.

وجدير بالذكر أن بعض هذه الأساليب يعاني من انتقادات، حيث يؤدي الأسلوب الأول "الإدلاء الذاتي بالإعاقة" إلى تقدير لحجم ظاهرة الإعاقة أقل من الفعلي، حيث تتراوح عادة نسبة الإعاقة وفقاً لهذا الأسلوب بين 1% إلى 3% (Mont 2007)، ويرجع ذلك إلى العديد من الأسباب منها أن كلمة "إعاقة" لها كثير من الدلالات السلبية لذا قد يلجأ الأفراد إلى إخفاء هذه الإعاقة، هذا علاوة على اختلاف مفهوم الإعاقة من شخص لآخر ومن مجتمع لآخر. ونستعرض فيما يلي الأساليب المتبعة في قياس الإعاقة في بعض الدول، وهي مصر، والاتحاد الأوروبي، والولايات المتحدة الأمريكية.

1-2-1 أسلوب قياس الإعاقة في مصر

تم قياس الإعاقة في التعدادات السابقة في مصر وفقاً لأسلوب الإدلاء الذاتي بالإعاقة (بالاعتماد على سؤال واحد فقط)، حيث يتم سؤال الأفراد بشكل مباشر "هل أنت أو أي أفراد الأسرة معاق؟"، ثم يتم السؤال عن نوع هذه الإعاقة. وقد عرف قانون رقم 39 لعام 1975 الشخص ذو الإعاقة، بأنه "كل شخص أصبح غير قادر على الاعتماد على نفسه ومزاولة عمله أو القيام بعمل والاستقرار فيه أو نقصت قدرته على ذلك نتيجة لقصور عضوي أو عقلي أو حسي أو نتيجة عجز خلقي منذ الولادة". وتصل نسبة الإعاقة في مصر إلى 0.7% وفقاً للتعداد العام للسكان لعام 2006 (الجهاز المركزي للتعبئة العامة والإحصاء 2006).

وجدير بالذكر أن مصر قد غيرت أسلوب قياس الإعاقة في تعدادها الأخير لعام 2017، حيث تم الالتزام بمجموعة الأسئلة الموجزة التي أعدها فريق واشنطن المعني بإحصاءات الإعاقة، وتشمل ستة أسئلة حول ما إذا كان الفرد يعاني من صعوبات خاصة بكل من: الرؤية، والسمع، والحركة، والتذكر أو التركيز، والرعاية الشخصية، والتواصل مع الآخرين، ووفقاً لهذا الأسلوب في القياس فقد تم تقدير نسبة الإعاقة (من الدرجة البسيطة إلى المطلقة) بحوالي 10.7% وذلك بين المصريين 5 سنوات فأكثر (الجهاز المركزي للتعبئة العامة والإحصاء 2017).

1-2-2 أساليب قياس الإعاقة في دول الاتحاد الأوروبي

تتعدد الأساليب المستخدمة لقياس الإعاقة في دول الاتحاد الأوروبي، وبالتالي تتعدد نسب الإعاقة، حيث تتراوح بين 11.0% و14.0% (Eurostat 2012)، وأبرز هذه الأساليب هي:

الأسلوب الأول: يتم قياس الإعاقة وفقاً لعشرة أسئلة، حيث يتم سؤال الأفراد إذا كانوا يعانون من صعوبة في أداء أيّاً من الأنشطة الأساسية العشرة وهي: (1) الرؤية، (2) السمع، (3) المشي أو صعود السلالم، (4) الجلوس أو الوقوف، (5) التذكر أو التركيز، (6) التواصل مع الآخرين (مثل فهمه للغير أو فهم الغير له)، (7) قدرة الذراعين على الحركة إلى الأمام، فوق الرأس، أو وراء الظهر (Reaching or stretching)، (8) رفع وحمل الأشياء، (9) الانحناء، (10) استخدام اليد بشكل جيد. ويُعرف هذا الأسلوب لقياس الإعاقة بصعوبات الأنشطة الأساسية "Basic Activity Difficulty". وتبلغ نسبة الإعاقة وفقاً لهذا الأسلوب في الاتحاد الأوروبي 14.0%.

الأسلوب الثاني يعتمد على تصنيف الشخص على أنه ذو إعاقة إذا كان يعاني من عدم القدرة على العمل بسبب مشكلة صحية طويلة الأمد و/أو صعوبة في الأنشطة الأساسية العشرة السابق ذكرها. ويعرف هذا الأسلوب بالإعاقة في العمل "Disability in Employment". وتبلغ نسبة الإعاقة وفقاً لهذا الأسلوب في الاتحاد الأوروبي 11.0%.

الأسلوب الثالث يعتمد على تصنيف الشخص على أنه ذو إعاقة إذا كان يواجه قيود في المشاركة في مجالات الحياة المختلفة - مثل التنقل والحركة، والتعليم، والعمل - بسبب مشكلة صحية طويلة الأمد و/أو صعوبة في الأنشطة الأساسية السابق ذكرها، وتبلغ نسبة الإعاقة وفقاً لهذا الأسلوب في الاتحاد الأوروبي 12.8%، وتصل في الدنمارك إلى 18.2%، وفي بريطانيا إلى 16.9%، وفي أسبانيا إلى 11.3%، وفي فرنسا إلى 11.1%، وفي قبرص إلى 10.2%.

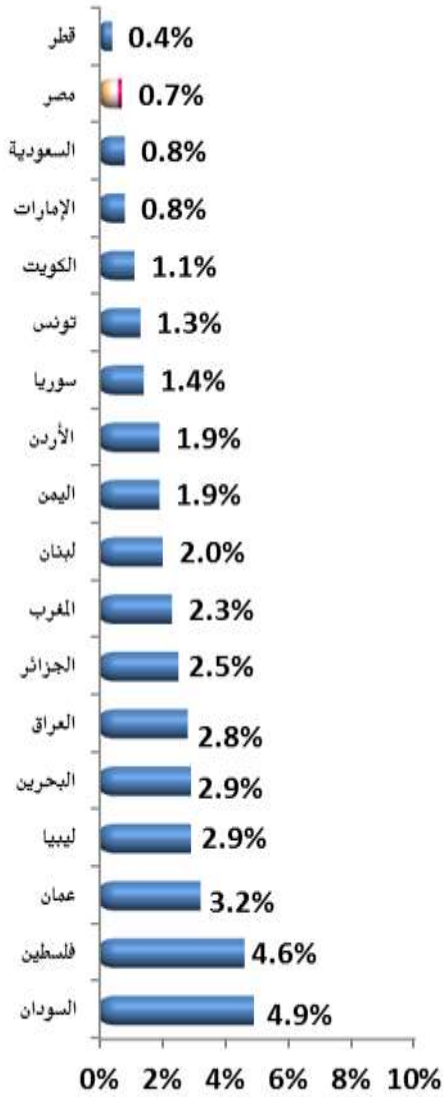
1-2-3 أساليب قياس الإعاقة في الولايات المتحدة الأمريكية

يتم قياس الإعاقة في الولايات المتحدة الأمريكية وفقاً لستة أسئلة عن المعاناة فى كل من: الرؤية، والسمع، والحركة، والتذكر أو التركيز أو القدرة على اتخاذ القرار، والرعاية الشخصية (الاستحمام، اللبس،...)، والقيام ببعض الأنشطة بمفرده مثل زيارة عيادة الطبيب أو التسوق (الحياة المستقلة). ويتم تصنيف الشخص على أنه معاق إذا كان يعاني من أي صعوبة من هذه الصعوبات الستة. وفيما يتعلق بانتشار الإعاقة في الولايات المتحدة الأمريكية، فقد بلغت نسبة الإعاقة 12.6% (Institute on Disability 2016).

هناك أسلوب آخر لقياس الإعاقة في الولايات المتحدة الأمريكية، حيث يتم تصنيف الشخص على أنه ذو إعاقة وفقاً لسؤالين: (1) هل تعاني بأي شكل من الأشكال من محدودية القيام بأي أنشطة بسبب مشاكل جسدية أو عقلية أو نفسية؟، (2) هل لديك الآن أي مشكلة صحية تتطلب منك أن تستخدم معدات خاصة، مثل كرسي متحرك، عكاز، سرير خاص، أو هاتف خاص (بما في ذلك استخدامها في بعض الأحيان أو استخدامها في ظروف معينة)؟، وتصل نسبة الأفراد ذوي الإعاقة إلى 21.6% بين الأفراد في الفئة العمرية 18 سنة فأكثر (Centers for Disease Control and Prevention (CDC) 2015).

1- 3 انتشار الإعاقة

شكل (2): معدل انتشار الإعاقة في عدد من الدول العربية



المصدر: اللجنة الاقتصادية والاجتماعية لغربي آسيا (الإسكوا)، جامعة الدول العربية، (2014)، "الإعاقة في المنطقة العربية - لمحة عامة".

ملحوظة: ترجع البيانات المعروضة إلى أحدث بيان متوفر لكل دولة وفقاً للإحصاءات الرسمية للدولة. وبالنسبة لمصر يرجع البيان إلى عام 2006، وفقاً للتعداد العام للسكان.

على الرغم من الاهتمام المتزايد بقضايا الإعاقة مازال توفر إحصاءات دقيقة وموثوقة يشكل تحدياً كبيراً يواجه معظم الدول بصفة عامة والنامية بصفة خاصة (Mont 2007).

وعلى الرغم من ذلك سعت منظمة الصحة العالمية بالتعاون مع البنك الدولي إلى توفير تقديرات عن انتشار الإعاقة، ويُقدَّر عدد الأفراد ذوي الإعاقة في العالم بحوالي مليار شخص - أي ما يُعادل 15% من سَكَّان العالم - (World Health Organization) (2011).

وفي مقابل ذلك، تتراوح نسبة الإعاقة في الدول العربية بين 0.4% في قطر و 4.9% في السودان (اللجنة الاقتصادية والاجتماعية لغربي آسيا (الإسكوا) 2014).

وتصل نسبة الإعاقة في مصر إلى 0.7% وفقاً للتعداد العام للسكان لعام 2006.

هذا التفاوت في نسب الإعاقة بين الدول المختلفة يرجع إلى الاختلاف في التعريفات وأساليب القياس، وجودة البيانات، علاوة على ثقافة عدم الإدلاء بالإعاقة.

كما قدرت نسبة الإعاقة في مصر بحوالي 1.6% عام 1991 وفقاً لمسح صحة الأم والطفل، و 0.5% وفقاً لتعداد 1996 (الديب، 2007).

وتشير البيانات المنشورة إلى ارتفاع نسب الإعاقة بين الذكور مقارنة بالإناث في معظم الدول العربية. تقدر هذه النسبة في الأردن بحوالي 2.2% و 1.6% بين الذكور والإناث على التوالي، بينما تقدر هذه النسبة في تونس بحوالي 1.7% بين الذكور، وتخفض لتصل إلى 1.0% بين الإناث، وفي مصر بلغت نسبة الإعاقة بين الذكور 0.8% بينما بلغت 0.5% بين الإناث من واقع بيانات التعداد العام لعام 2006 (اللجنة الاقتصادية والاجتماعية لغربي آسيا (الإسكوا) 2014).

ويختلف هذا النمط في دول الإتحاد الأوروبي وفي الولايات المتحدة الأمريكية، حيث ترتفع نسبة الإعاقة بين الإناث عن الذكور في الإتحاد الأوروبي، حيث تصل بين الإناث إلى 15.1% في الفئة العمرية 15-64 سنة، وتخفض بين الذكور إلى 12.9%، وذلك وفقاً لأسلوب القياس السابق ذكره، والذي يُعرف بصعوبات الأنشطة الأساسية (Eurostat 2012).

كذلك ترتفع نسبة الإعاقة بين الإناث في الولايات المتحدة الأمريكية عن الذكور، حيث تصل بين الإناث إلى 12.8%، وبين الذكور إلى 12.5%، وذلك لإجمالي السكان (U.S. Census Bureau 2016).

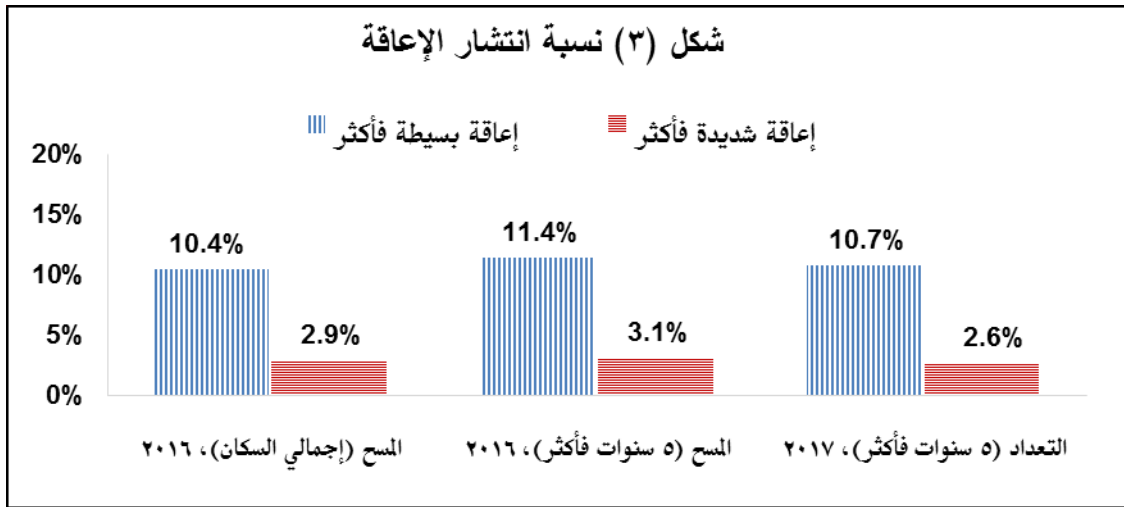
2- مصدر البيانات

يعتمد البحث على بيانات المسح القومي "مسح أحوال الأسرة المصرية لعام 2016-الدورة 13"، والذي نفذه مركز المعلومات ودعم اتخاذ القرار، وقد تم جمع بيانات عن الإعاقة استناداً إلى التعريفات الموصى بها دولياً، وتم الالتزام بمجموعة الأسئلة الموجزة التي أعدها فريق واشنطن المعنيّ بإحصاءات الإعاقة (United Nations, ESCAP 2014)، ويوضح جدول (أ) في الملحق هذه الأسئلة، ويستهدف هذا المسح تمثيل الأسر في المجتمع المصري على مستوى الجمهورية، وعلى مستوى المناطق الجغرافية الستة (المحافظات الحضرية، وحضر الوجه البحري، وريف الوجه البحري، وحضر الوجه القبلي، وريف الوجه القبلي، ومحافظات الحدود)، وعلى مستوى المحافظات، وقد تم استيفاء بيانات عن 49431 فرد من خلال 11592 أسرة على مستوى الجمهورية، خلال الفترة من 4 مارس 2016 حتى 20 مايو 2016. وقد تم السؤال عن الإعاقة بين أفراد الأسر من خلال ستة أسئلة حول ما إذا كان الفرد يعاني من صعوبات خاصة بكل من: الرؤية، والسمع، والحركة، والتذكر أو التركيز، والرعاية الشخصية (الاستحمام، اللبس،....)، والتواصل مع الآخرين (مثل فهمه للغير أو فهم الغير له)، كما تم السؤال عن شدة هذه الإعاقة، حيث يتم تصنيف الإعاقة إلى إعاقة بسيطة، أو إعاقة شديدة، أو إعاقة تامة.

3- النتائج

3-1 انتشار الإعاقة

تُقدَّر نسبة ذوي الإعاقة (بسيطة فأكثر³) في مصر في عام 2016 بحوالي 10.4% (5156 فرد) من إجمالي حجم العينة في عام 2016، وذلك كما يتضح من شكل (3). ويتم اعتبار الفرد لديه إعاقة إذا كان يعاني من صعوبة واحدة على الأقل من الصعوبات الستة التي أعدها فريق واشنطن المعني بإحصاءات الإعاقة والتي سبق الإشارة إليها، أيًا كانت درجة صعوبة الإعاقة (بسيطة أو شديدة أو تامة). ويعرض جدول (ب) في الملحق تعريفات أنواع الإعاقات ودرجات شدة الإعاقة. وقد استخدم التعداد العام لمصر 2017 نفس أسلوب التقدير المستخدم في البحث الحالي، ويشير شكل (3) إلى أن نسبة الإعاقة (بسيطة فأكثر) بين الأفراد في الفئة العمرية 5 سنوات فأكثر من واقع بيانات تعداد 2017 بلغت 10.7%، وقد بلغت هذه النسبة وفقاً لنتائج المسح المستخدم في هذا البحث 11.4% لنفس الفئة العمرية، وهو ما يشير إلى اتساق النتائج بينهما، كما بلغت نسبة الإعاقة شديدة فأكثر للأفراد 5 سنوات فأكثر 2.6% و3.1% من واقع بيانات التعداد 2017 وبيانات المسح 2016 على التوالي.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

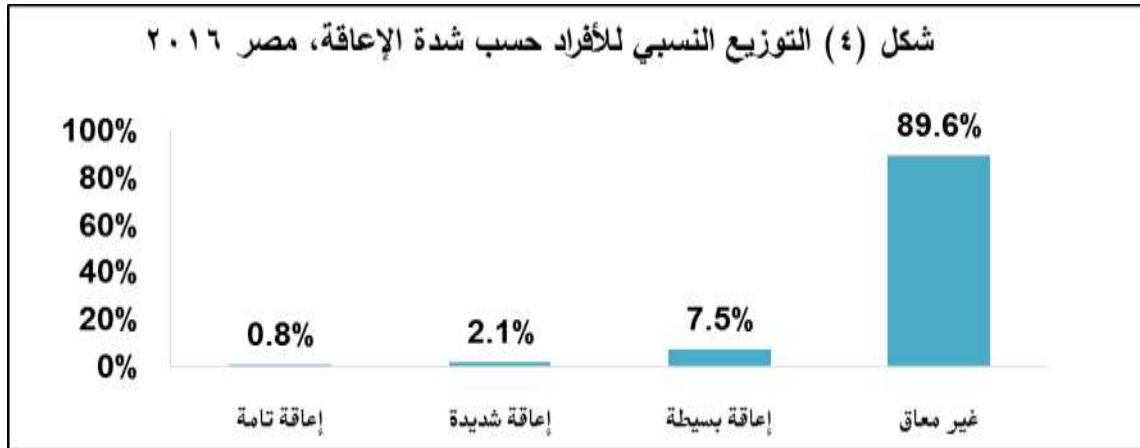
الجهاز المركزي للتعبئة العامة والإحصاء، (2017)، "النتائج النهائية للتعداد العام للسكان والإسكان والمنشآت لعام 2017".

³ تم اعتبار الفرد معاق إعاقة بسيطة فأكثر عند حساب معدل الانتشار إذا كان يعاني من صعوبة واحدة على الأقل من الصعوبات الستة التي سبق الإشارة إليها، أيًا كانت درجة صعوبة الإعاقة، أما المعاق إعاقة شديدة فأكثر فقد تم تعريفها هنا إذا كان الفرد يعاني من صعوبة شديدة واحدة على الأقل من الصعوبات الستة.

أما بالنسبة لمستويات شدة الإعاقة، فسوف يتم تحديد الحد الفاصل بين كل مستوى على النحو التالى(بالشكل الذى يضمن أن يجمع توزيع مجتمع الدراسة 100%)

- **إعاقة تامة:** يتم اعتبار الفرد لديه إعاقة تامة إذا كان يعاني من صعوبة تامة في أحد الأنواع الستة للصعوبات، وجدير بالذكر أن الفرد يمكن أن يعاني من أكثر من صعوبة في نفس الوقت، فقد يعاني من صعوبة تامة في الحركة وصعوبة شديدة في الرؤية، وفي هذه الحالة يتم تصنيف الفرد على أنه يعاني من إعاقة تامة.
- **إعاقة شديدة:** يتم اعتبار الفرد لديه إعاقة شديدة إذا كان يعاني من صعوبة شديدة على الأكثر في أحد الأنواع الستة للصعوبات، وجدير بالذكر أن الفرد يمكن أن يعاني من أكثر من صعوبة في نفس الوقت، فقد يعاني من صعوبة شديدة في الحركة وصعوبة بسيطة في الرؤية، وفي هذه الحالة يتم تصنيف الفرد على أنه يعاني من إعاقة شديدة.
- **إعاقة بسيطة:** يتم اعتبار الفرد لديه إعاقة بسيطة إذا كان يعاني من صعوبة بسيطة على الأكثر في أحد الأنواع الستة للصعوبات.

تشكل الإعاقات البسيطة النصيب الأكبر من الإعاقات كما يتضح من الشكل رقم (4)، فحوالى 7.5% من الأفراد المعاقين لديهم إعاقة بسيطة، بينما يعانى 2.1% و 0.8% من الأفراد إعاقة شديدة وإعاقة تامة على التوالي. جدير بالذكر أنه تم تحديد الحد الفاصل بين كل مستوى من مستويات شدة الإعاقة عند حساب التوزيع حسب شدة الإعاقة على النحو التالى:



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

ويعرض جدول (1) نسبة الأسر التي يوجد بها أفراد ذوي إعاقة، ويتضح من هذا الجدول أن حوالى 21% من الأسر لديها فرد واحد ذو إعاقة بسيطة، تنخفض هذه النسبة لتصل إلى 2.6% من الأسر لديها فرد ذو إعاقة تامة، ونحو 0.2% من الأسر لديها فردين ذوي إعاقة

تامة، و0.08% من الأسر لديها 3 أفراد فأكثر ذوي إعاقة تامة. وبالنسبة للأسر التي يوجد بها أفراد ذوي إعاقة شديدة، نحو 7.9% من الأسر في مصر لديها فرد ذو إعاقة شديدة، ونحو 0.8% من الأسر لديها فردين ذوي إعاقة شديدة، و0.03% من الأسر لديها 3 أفراد فأكثر ذوي إعاقة شديدة.

وبصفة عامة فإن نحو ربع الأسر المصرية لديها فرد معاق سواء إعاقة بسيطة أو تامة أو شديدة، وحوالي 7.0% من الأسر لديها فردين ذوي إعاقة، وحوالي 1.6% من الأسر لديها 3 أفراد فأكثر ذوي إعاقة.

جدول (1) نسبة الأسر حسب عدد الأفراد المعاقين بها وحسب نوع الإعاقة، مصر 2016

الإجمالي	عدد الأفراد ذوي الإعاقة			نوع الإعاقة
	ثلاثة أفراد فأكثر	فردين	فرد معاق	
27.8%	1.1%	5.8%	20.9%	إعاقة بسيطة
8.73%	0.03%	0.8%	7.9%	إعاقة شديدة
2.88%	0.08%	0.2%	2.6%	إعاقة تامة
32.5%	1.6%	7.0%	23.9%	أي إعاقة

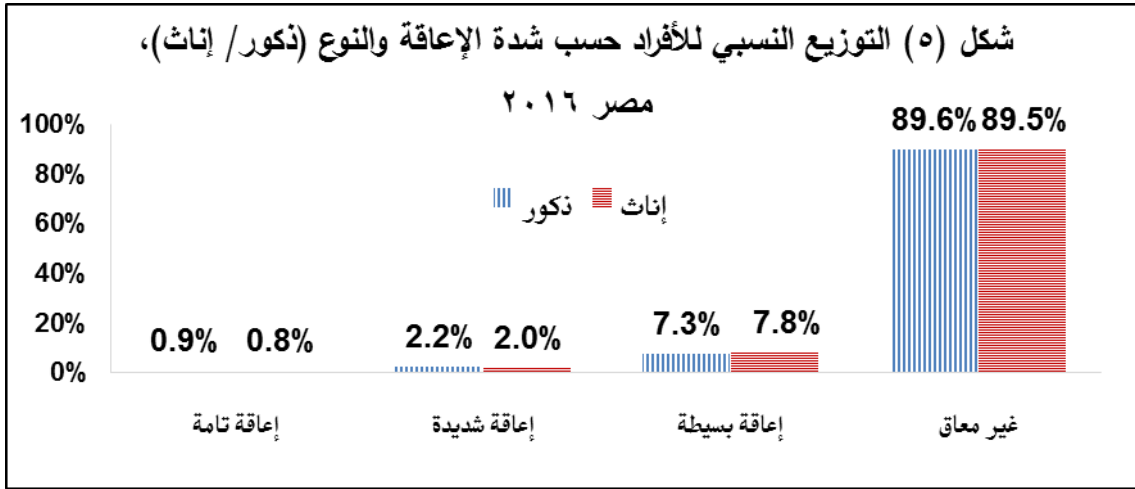
المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

3-1-1 انتشار الإعاقة وشدتها حسب النوع

يشير الشكل رقم (5) إلى عدم وجود اختلاف بين الذكور والإناث من حيث نسبة الإعاقة وشدتها، حيث تبلغ نسبة الإعاقة بين الذكور 10.4% وبين الإناث 10.5%. أما بالنسبة لشدة الإعاقة بين الذكور والإناث، فحوالي 0.9% من الذكور لديهم إعاقة تامة، وتصل هذه النسبة بين الإناث إلى 0.8%، وفيما يتعلق بالإعاقات الشديدة، حوالي 2.2% من الذكور لديهم إعاقات شديدة، مقارنة بـ 2.0% بين الإناث. وتعدُّ الإعاقات البسيطة هي الأكثر انتشاراً بين الذكور والإناث على حد سواء، حيث تصل النسبة بينهم إلى 7.3%، و 7.8% على التوالي.

وعند إجراء الاختبارات الإحصائية تبين أن العلاقة بين الإعاقة والنوع غير معنوية ($p > 0.05$) وذلك باستخدام كل من اختبار كاي² وكذلك (Goodman and Kruskal tau⁴).

⁴ يدرس هذا الاختبار العلاقة بين متغيرين اسميين، ويتميز عن اختبار كاي² انه يساعد على التعرف على قوة العلاقة.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

3- 2-1 انتشار الإعاقة وشدتها حسب فئات العمر

ترتفع نسبة ذوي الإعاقة كلما ارتفع العمر، ويعرض شكل (6) نسبة الأفراد ذوي الإعاقة (بسيطة فأكثر) وفقاً لفئات العمر، وتظهر النتائج أن حوالي 1.4% من الأطفال أقل من 5 سنوات يعانون من الإعاقة، وترتفع هذه النسبة لتصل إلى 3.1% بين الأطفال في الفئة العمرية 5-9 سنوات، وتبلغ 4.0% بين الأطفال في الفئة العمرية 10-14 سنة. فيما يتعلق بالشباب، تصل نسبة الإعاقة بين الأفراد في الفئة العمرية 15-19 سنة إلى 4.5%، وتصل إلى 4.6% بين الأفراد في الفئة العمرية 20-24 سنة.

تبدأ نسبة الإعاقة في الارتفاع بوتيرة أسرع مع ارتفاع العمر، فمن بين كل 100 فرد في الفئة العمرية 50 - 54 سنة هناك 18 فرد معاق، ويرتفع هذا العدد ليصل إلى 39 فرد معاق بين كبار السن للفئة العمرية (65- 69 سنة) من بين كل 100 فرد، ثم إلى حوالي 58 فرد معاق في الفئة العمرية 70 سنة فأكثر من بين كل 100 فرد. وهو نفس النمط في دول الاتحاد الأوروبي (Eurostat 2012)، وفي الولايات المتحدة الأمريكية (Institute on Disability 2016).



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

وعند استخدام اختبار كاي² كانت العلاقة بين الإعاقة وفئات العمر معنوية ($p < 0.01$)، كما كانت العلاقة معنوية وطردية ($p < 0.01$) عند حساب معامل جاما، فكلما ارتفع العمر ارتفعت نسبة الإعاقة.

وقد تم بناء نموذج إنحدار لوجستي ثنائي لدراسة العلاقة بين الإعاقة (متغير تابع Dummy variable له وجهين أحدهما معاق والآخر غير معاق) وفئات العمر (المتغير المستقل)، ويعرض جدول (2) نسب الأرجحية (Odds Ratio) ويتضح من هذا الجدول أن نسب الأرجحية ترتفع بشكل كبير مع ارتفاع العمر، فمفاضلة أن يصاب الإنسان بالإعاقة في الفئة العمرية من 24-20 سنة يصل إلى أكثر من 3 أضعاف أن يصاب الإنسان بالإعاقة في الفئة العمرية أقل من 5 سنوات، ويصل إلى 15 ضعف في الفئة العمرية 54-50 سنة، وإلى 100 ضعف في الفئة العمرية 70 سنة فأكثر مقارنة بالأطفال أقل من 5 سنوات.

جدول (2) نسب الأرجحية (Odds Ratio) للعلاقة بين الإعاقة (بسيطة فأكثر) وفئات العمر،

مصر 2016

95% C.I. For EXP (B)		Exp (B) نسبة الأرجحية	فئات العمر
Upper	Lower		
-	-	-	أقل من 5 سنوات (الفئة المرجعية)
3.042	1.695	***2.271	5-9 سنين
3.982	2.258	***2.999	10-14 سنة
4.459	2.543	***3.367	15-19 سنة

جدول (2) نسب الأرجحية (Odds Ratio) للعلاقة بين الإعاقة (بسيطة فأكثر) وفئات العمر،

مصر 2016

95% C.I. For EXP (B)		Exp (B) نسبة الأرجحية	فئات العمر
Upper	Lower		
4.593	2.566	***3.433	20-24 سنة
5.770	3.236	***4.321	25-29 سنة
7.008	3.967	***5.273	30-34 سنة
8.969	5.135	***6.786	35-39 سنة
11.628	6.696	***8.824	40-44 سنة
14.282	8.278	***10.873	45-49 سنة
20.378	11.909	***15.579	50-54 سنة
29.319	17.109	***22.397	55-59 سنة
45.702	26.774	***34.981	60-64 سنة
61.366	35.577	***46.725	65-69 سنة
130.920	76.288	***99.938	70 سنة فأكثر

*** p-value ≤ 0.001

المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

يختلف مستوى شدة الإعاقة حسب فئات العمر، وتشير النتائج بجدول (3) إلى عدد من النقاط الأساسية:

- ✓ تزداد نسبة الأفراد الذين يعانون من إعاقات سواء بسيطة أو شديدة أو تامة مع ارتفاع العمر.
- ✓ تشكل الإعاقات البسيطة النسبة الأكبر من إجمالي الإعاقات في جميع فئات العمر، يليها الإعاقات الشديدة، بينما تنخفض نسبة الإعاقات التامة لفئات السن المختلفة.
- ✓ بالنسبة للأفراد في الفئة العمرية 70 سنة فأكثر، تصل نسبة الأفراد ذوي الإعاقة إلى 58.2%، وتتوزع هذه النسبة بين 36.8% من الأفراد لديهم إعاقة بسيطة، و16.3% لديهم إعاقة شديدة، و5.1% لديهم إعاقة تامة.

جدير بالذكر أن عدد الأفراد ذوي الإعاقة التامة في العينة منخفض وقد يتأثر توزيعهم على الفئات العمرية المختلفة، بما يتطلب الحذر أثناء قراءة النتائج.

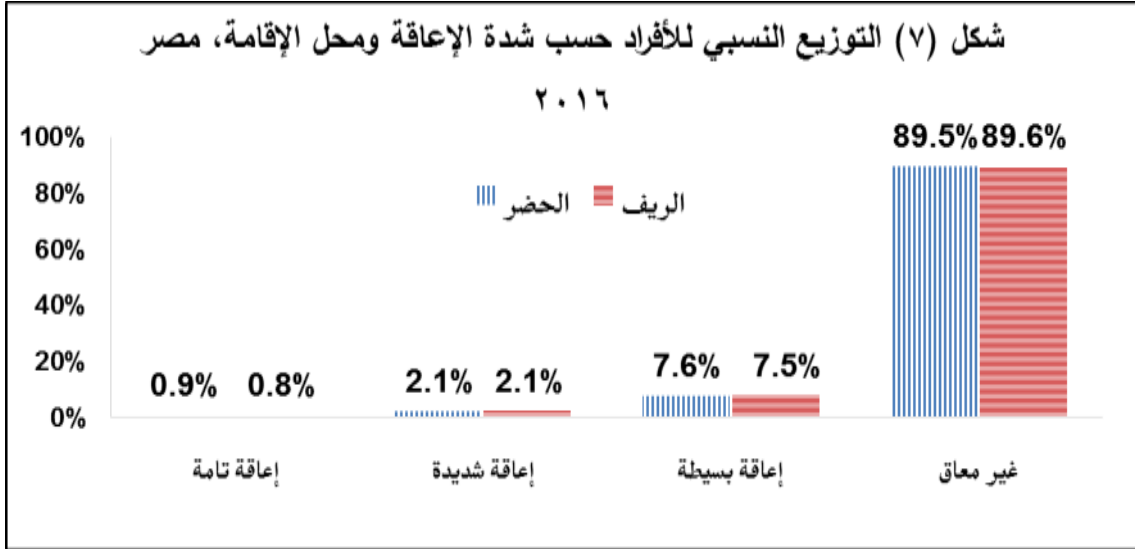
جدول (3) التوزيع النسبي للأفراد حسب شدة الإعاقة وفئات العمر، مصر 2016

الإجمالي	لديه إعاقة			غير معاق	فئات العمر
	إعاقة تامة	إعاقة شديدة	إعاقة بسيطة		
%100	%0.3	%0.3	%0.8	%98.6	أقل من 5 سنوات
%100	%0.6	%0.6	%1.9	%96.9	5- 9 سنين
%100	%0.6	%0.6	%2.8	%96.0	10- 14 سنة
%100	%0.8	%0.8	%2.9	%95.5	15- 19 سنة
%100	%0.7	%0.9	%2.9	%95.4	20- 24 سنة
%100	%0.9	%1.3	%3.5	%94.3	25- 29 سنة
%100	%0.9	%1.2	%4.7	%93.2	30- 34 سنة
%100	%0.3	%1.9	%6.4	%91.4	35- 39 سنة
%100	%0.6	%1.8	%8.6	%89.1	40- 44 سنة
%100	%0.5	%2.1	%10.5	%86.9	45- 49 سنة
%100	%0.5	%2.5	%14.8	%82.2	50- 54 سنة
%100	%0.9	%5.2	%17.6	%76.2	55- 59 سنة
%100	%1.3	%5.8	%25.6	%67.3	60- 64 سنة
%100	%1.8	%8.1	%29.5	%60.6	65- 69 سنة
%100	%5.1	%16.3	%36.8	%41.8	70 سنة فأكثر
%100	%0.8	%2.1	%7.5	%89.6	الإجمالي

المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

3-1-3 انتشار الإعاقة وشدتها حسب محل الإقامة

لا تختلف نسبة الإعاقة بين الحضر والريف، حيث تبلغ نسبة الإعاقة في الحضر 10.5% وفي الريف 10.4%. كما لا تختلف درجة شدة الإعاقة بين الحضر والريف، فحوالي 2.1% من الأفراد في كل من الحضر والريف يعانون من إعاقات شديدة، بينما يعاني 0.9% و 0.8% من الأفراد في الحضر والريف من إعاقات تامة على التوالي. وعند إجراء الاختبارات الإحصائية تبين أن العلاقة بين الإعاقة ومحل الإقامة غير معنوية ($p > 0.05$) وذلك باستخدام كل من اختبار كاي² وكذلك (Goodman and Kruskal tau).



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

تتفاوت نسبة ذوي الإعاقة وفقاً للمناطق الجغرافية المختلفة، حيث تصل إلى أقصاها (12.6%) في حضر الوجه البحري، يليه ريف الوجه البحري بنسبة 11.9%، ثم حضر الوجه القبلي بنسبة 10.8%، وتُقدَّر بـ 9.2% في المحافظات الحضرية، و 8.7% في ريف الوجه القبلي، وتصل إلى أدناها في محافظات الحدود بنسبة 7.0%، وذلك كما يتضح من شكل (8).

وعند إجراء الاختبارات الإحصائية تبين أن العلاقة بين الإعاقة والمناطق الجغرافية معنوية ($p < 0.01$) وذلك باستخدام كل من اختبار كاي 2 وكذلك (Goodman and Kruskal tau) وإن كانت قوة العلاقة ضعيفة للغاية. وجددير بالذكر أن بعض هذه الاختلافات بين المناطق الجغرافية قد يرجع إلى عدم الرغبة بالإدلاء بالإعاقة في بعض المناطق نظراً لسيطرة العادات والتقاليد التي تميل إلى عدم الرغبة في الإدلاء بنوع الإعاقة لأحد أفراد الأسرة.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

وتشير النتائج إلى أن الإعاقات البسيطة هي الأكثر انتشاراً في جميع المناطق الجغرافية، يليها الإعاقات الشديدة، ثم الإعاقات التامة، كما يتضح من الشكل (9). فيما يتعلق بالإعاقات التامة، تبلغ نسبة الأفراد الذين يعانون من إعاقات تامة حوالي 0.9% في كل من المحافظات الحضرية، وحضر الوجه البحري، وحضر الوجه القبلي، وتصل إلى 0.8% في ريف الوجه البحري، وإلى 0.7% في ريف الوجه القبلي، بينما تصل إلى أدناها في المحافظات الحدودية حيث تُقدر نسبة ذوي الإعاقات التامة فيها بحوالي 0.3%.

تتفاوت نسبة ذوي الإعاقة بين المحافظات في مصر بشكل ملحوظ وذلك كما يتضح من شكل (10)، وترتفع هذه النسبة في معظم محافظات الوجه البحري مقارنة بباقي المحافظات، حيث تبلغ في الغربية 16.9%، وفي كفر الشيخ 14.1%، وفي دمياط 13.7%، وفي البحيرة 13.1%، وفي الشرقية 12.6%. (ترتفع نسب الإعاقة بمحافظات الوجه البحري مقارنة بمحافظات الوجه القبلي).

بالنسبة للمحافظات الحضرية، تتفاوت نسبة ذوي الإعاقة في هذه المحافظات، حيث تبلغ في الإسكندرية 12.2%، وتصل في بورسعيد إلى 11.0%، وفي السويس إلى 9.7%، وفي القاهرة إلى 7.6%. أما فيما يتعلق بمحافظات الوجه القبلي، تتراوح نسبة ذوي الإعاقة في هذه المحافظات بين 6.6% و 10.8%، حيث تبلغ في أسوان 10.8%، وفي سوهاج 10.3%، وفي

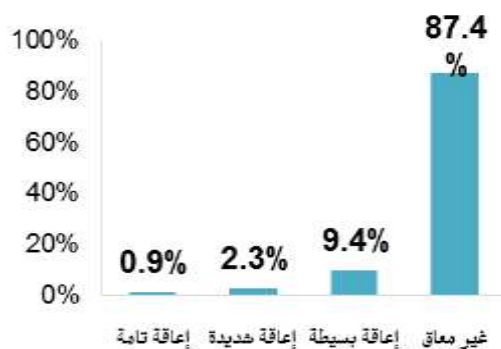
بنى سويف 9.9%، وفى كل من الفيوم وإمبنا 9.7%، وفى أسبوط 9.5%، وفى الأقصر 9.3%، وفى الجيزة 8.8%، وفى قنا 6.6%.

بالنسبة لمحافظة الحدود، تصل نسبة الإعاقة إلى أقصاها فى جنوب سبنا حيث تبلغ 12.2%، وتصل فى مرسى مطروح إلى 5.0%.

شكل (9) التوزيع النسبي للأفراد في المناطق الجغرافية المختلفة حسب شدة الإعاقة،

مصر 2016

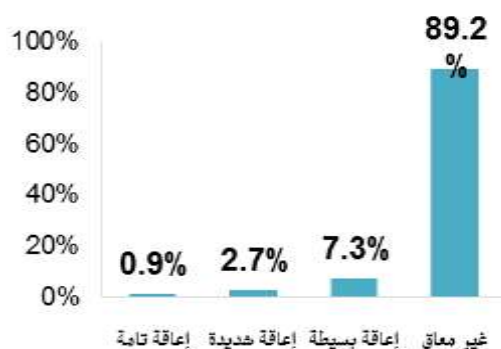
حضر الوجه البحري



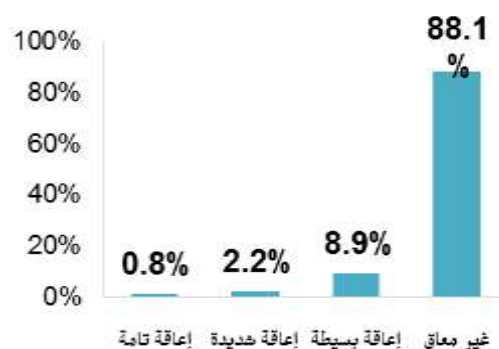
المحافظات الحضرية



حضر الوجه القبلي



ريف الوجه البحري



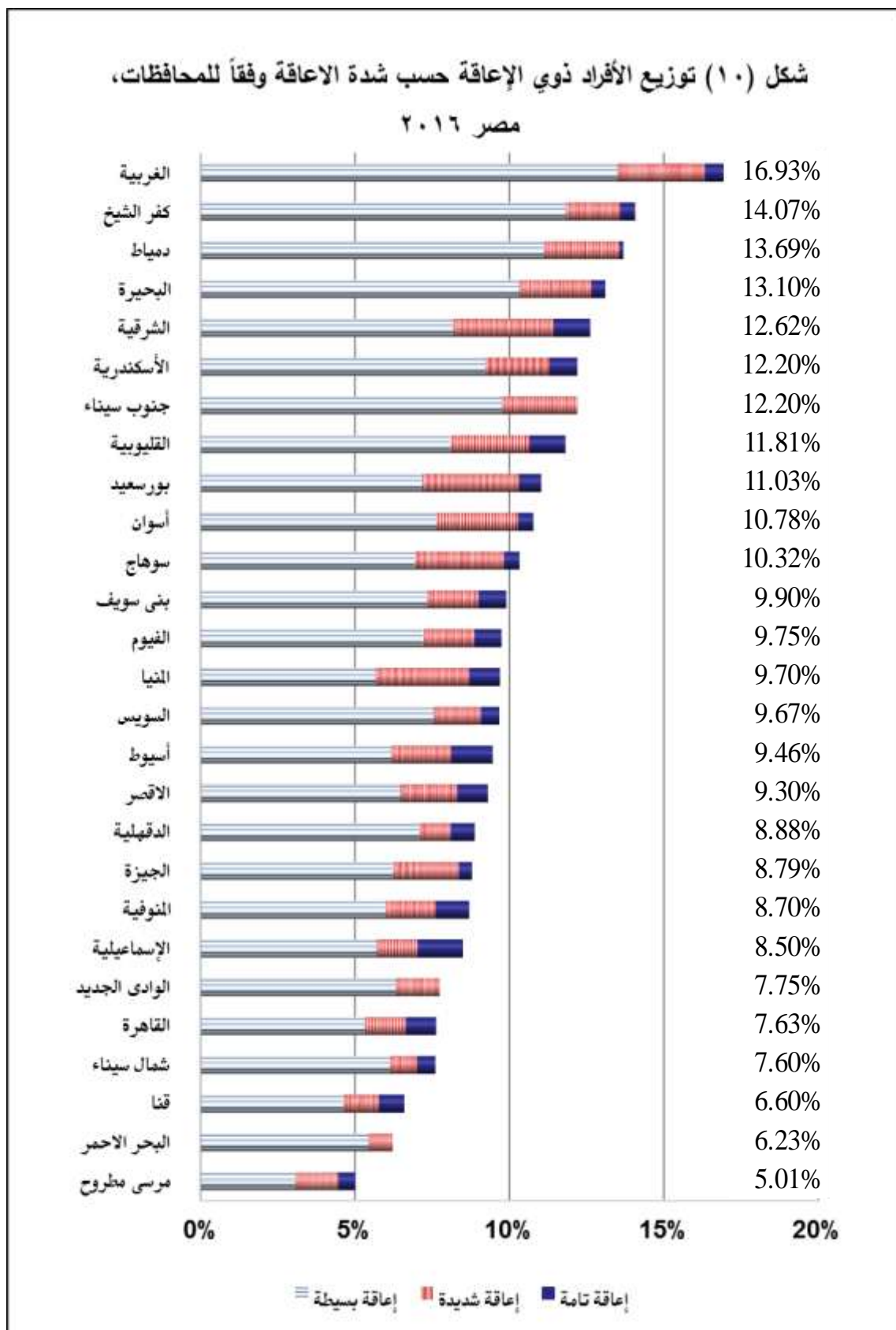
محافظات الحدود



ريف الوجه القبلي



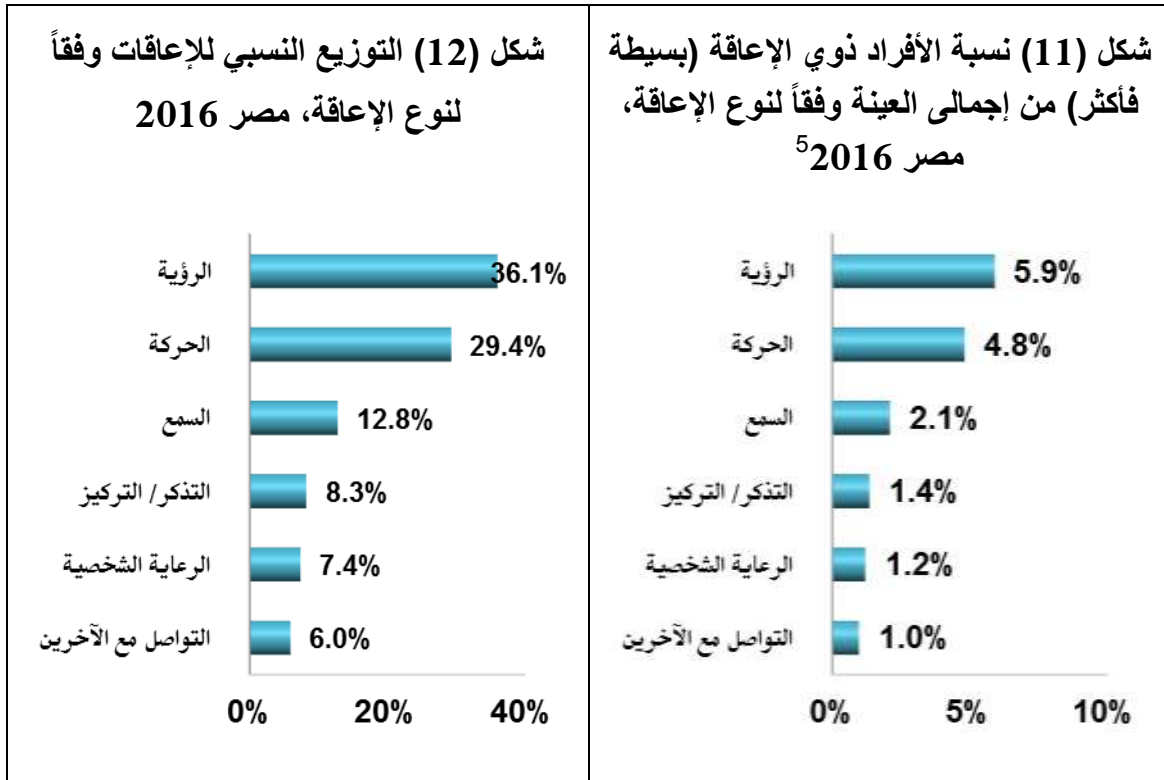
المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

2-3 أنواع الإعاقات في مصر

تتعدد أنواع الإعاقات المختلفة التي يتعرض لها الافراد، وتشير النتائج إلى أن أكثر الإعاقات التي يعاني منها الأفراد في مصر تتركز في صعوبات متعلقة بالرؤية، يليها الحركة ثم السمع، بنسبة 5.9%، 4.8%، 2.1% على التوالي، في حين تنخفض نسبة الأفراد الذين يعانون من صعوبات تتعلق بالتذكر/ التركيز، والرعاية الشخصية، والتواصل مع الآخرين، كما يتضح من الشكل رقم (11)، وتجدر الإشارة إلى هناك بعض الأفراد يعانون من أكثر من إعاقة في نفس الوقت. ويعرض الشكل رقم (12) التوزيع النسبي للأفراد المعاقين حسب نوع الإعاقة، وتعد إعاقات الرؤية هي الأكثر انتشاراً فحوالي 36.1% من الإعاقات في مصر هي إعاقات رؤية، يليها الحركة بنسبة 29.4%.

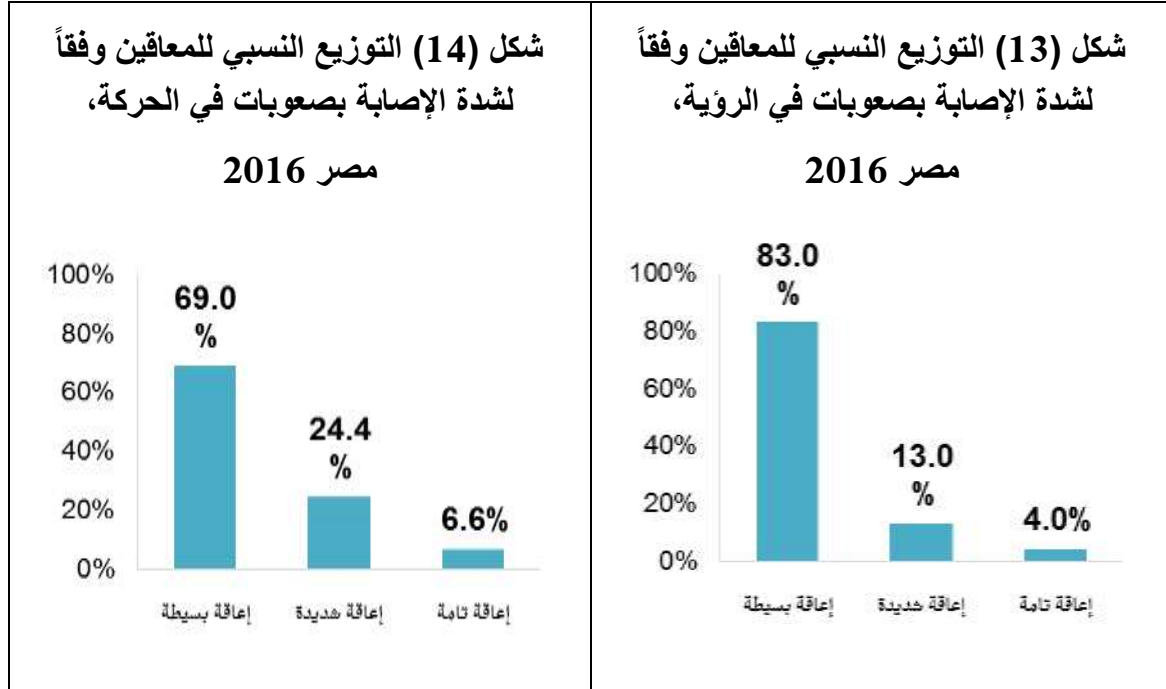


المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

وتختلف درجة شدة كل نوع من الإعاقات حيث من بين 5.9% من الأفراد الذين يعانون من صعوبات الرؤية، هناك 83.0%، 13% منهم يعانون من صعوبات بسيطة و شديدة على التوالي، وحوالي 4.0% لا يستطيعون الرؤية تماماً، وذلك كما يتضح من شكل (13). وكما

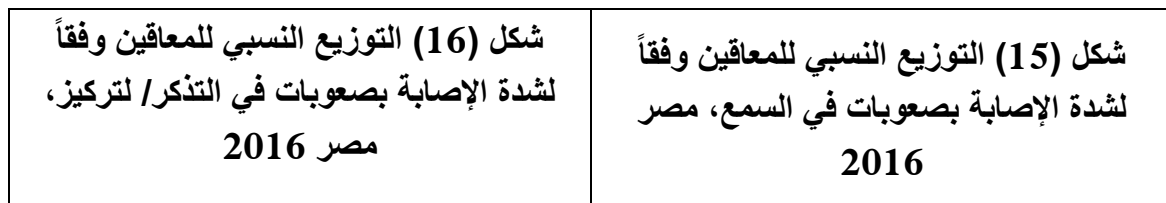
⁵ مجموع الاعددة يساوى 16.4% وهو أكبر من 10.4% حيث يمكن أن يعاني الفرد من أكثر من إعاقة.

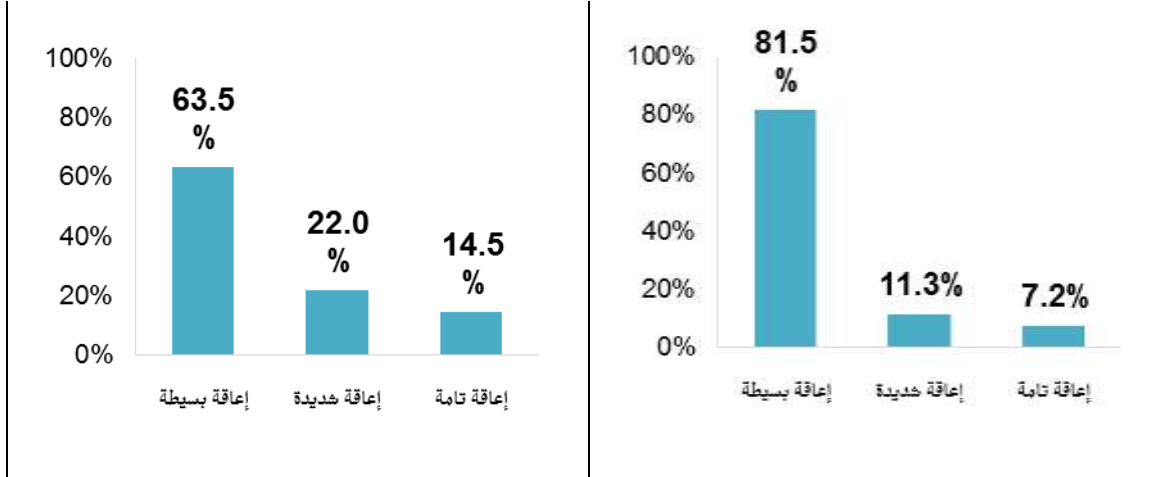
يتضح من الشكل رقم (14) انخفاض نسبة هؤلاء الأفراد يعانون من صعوبات بسيطة في الحركة (المشي/ صعود السلالم) إلى 69% مقارنة بالرؤية، بينما يعاني ربع الأفراد المعاقين بالحركة (24%) من صعوبات شديدة، وحوالي 6.6% لا يستطيعون الحركة تماماً.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

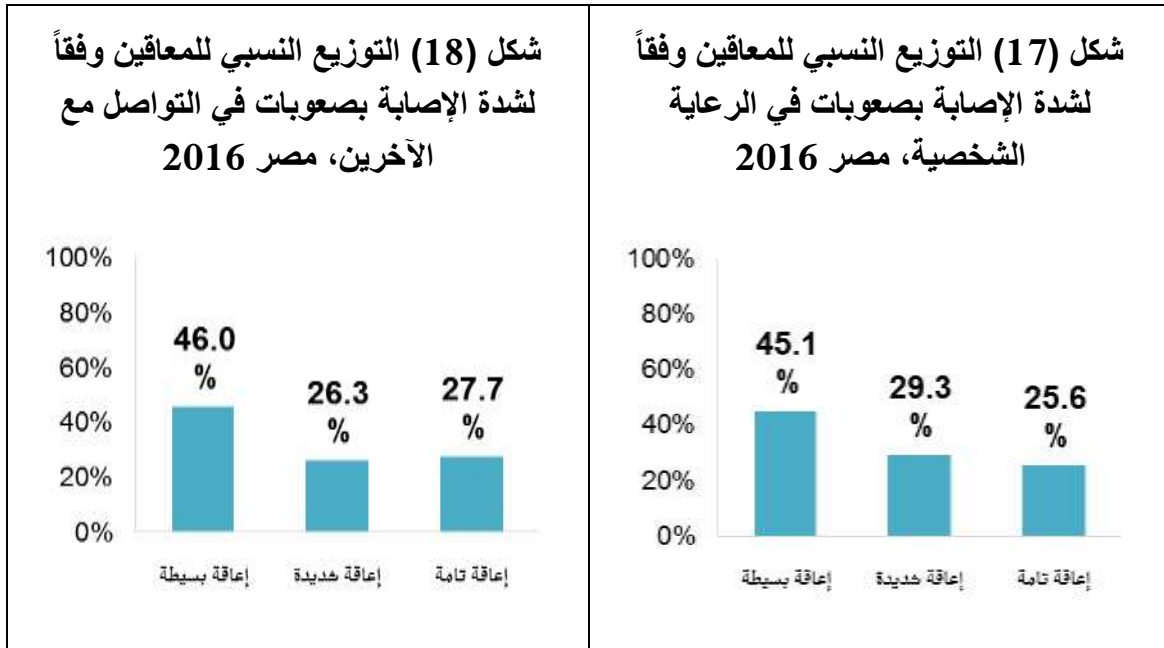
بينما يعاني 2.1% من الأفراد في مصر من صعوبات متعلقة بالسمع، يظهر الشكل رقم (15) تفاوت شدة هذه الصعوبات، فحوالي 81.5% من هؤلاء الأفراد يعانون من صعوبات بسيطة في السمع، وحوالي 11.3% يعانون من صعوبات شديدة في السمع، وحوالي 7.2% لا يستطيعون السمع تماماً. وتأتي صعوبة التذكر/ التركيز في المركز الثالث بين الإعاقات (1.4% من إجمالي السكان)، ويشير الشكل رقم (16) أن حوالي 63.5% من هؤلاء الأفراد يعانون من صعوبات بسيطة، بينما يعاني 22.0% يعانون من صعوبات شديدة، وحوالي 14.5% لا يستطيعون التذكر أو التركيز بشكل تام.





المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

وتعد الإعاقة بعدم القدرة على الرعاية الشخصية والتواصل مع الآخرين من أقل الإعاقات انتشاراً فى المجتمع المصرى (1.2% و 1% على التوالى)، ويعرض الشكل رقم (17) والشكل رقم (18) التفاوتات حسب شدة هذه الصعوبات. من بين الإعاقات بعدم القدرة على الرعاية الشخصية يوجد حوالي 45.1% يعانون من صعوبات بسيطة، وحوالى 29.3% يعانون من صعوبات شديدة، وحوالى 25.6% لا يستطيعون القيام تماماً ببعض الأنشطة المتعلقة برعاية أنفسهم. ولا تختلف درجات شدة الإعاقة بالنسبة للتواصل مع الآخرين، ما يقرب من نصف هؤلاء الأفراد يعانون من صعوبات بسيطة (46%)، وحوالى الربع (26.3%) يعانون من صعوبات شديدة، وحوالى 27.7% لا يستطيعون التواصل مع الآخرين تماماً.

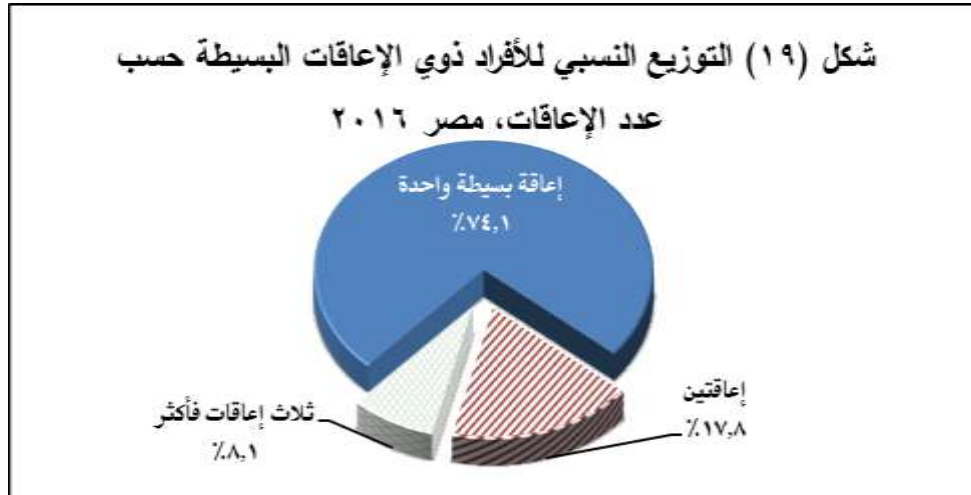


المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

3-3 الإصابة بأكثر من إعاقة Double Burden

قد يعاني الأفراد ذوي الإعاقة من أكثر من إعاقة في نفس الوقت، مما قد يزيد من وطأة الإعاقة ويحد من قدرته على التمتع باحتياجاته الأساسية. وتقدم الأشكال (19)، (20)، (21) نسب انتشار إصابة الفرد بأكثر من إعاقة حسب مستويات شدة الإعاقة المختلفة (بسيطة، شديدة، إعاقات تامة) على التوالي.

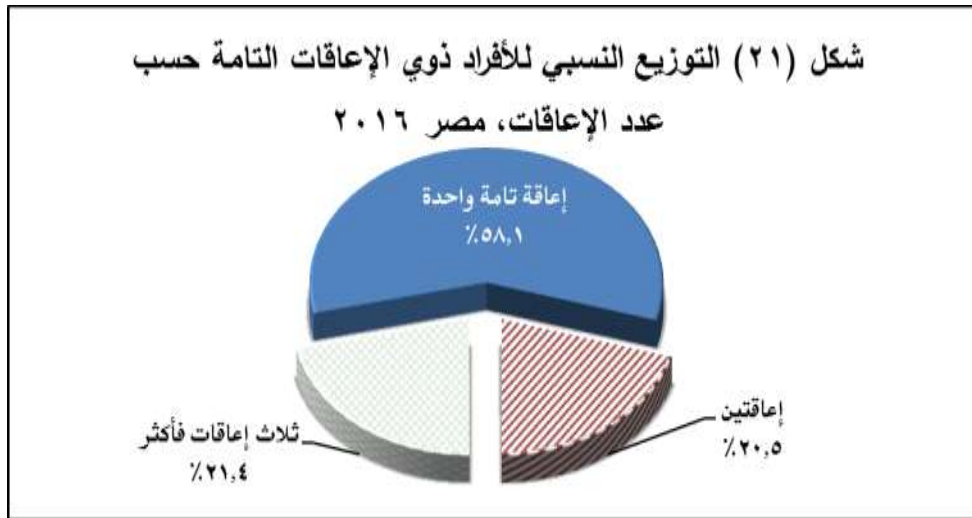
فيما يتعلق بالأفراد ذوي الإعاقة البسيطة، يتضح من شكل (19) أن 74.1% من إجمالي الأفراد الذين يعانون من إعاقات بسيطة لديهم إعاقة واحدة، وحوالي 17.8% لديهم إعقتين بسيطتين، و8.1% لديهم ثلاث إعاقات بسيطة أو أكثر. بينما نجد أن 77% من إجمالي الأفراد الذين لديهم إعاقات شديدة يعانون بإعاقة واحدة شديدة فقط، وحوالي 16.5% لديهم إعقتين، و6.5% لديهم ثلاث إعاقات أو أكثر، الشكل (20). أما بالنسبة للأفراد ذوي الإعاقة التامة، فيشير الشكل (21) إلى أن نسبة كبيرة منهم يكون لديهم أكثر من إعاقة في نفس الوقت، حيث أن 58.1% من الأفراد ذوي الإعاقة التامة لديهم إعاقة واحدة، وحوالي 20.5% لديهم إعقتين، و21.4% لديهم ثلاث إعاقات أو أكثر.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

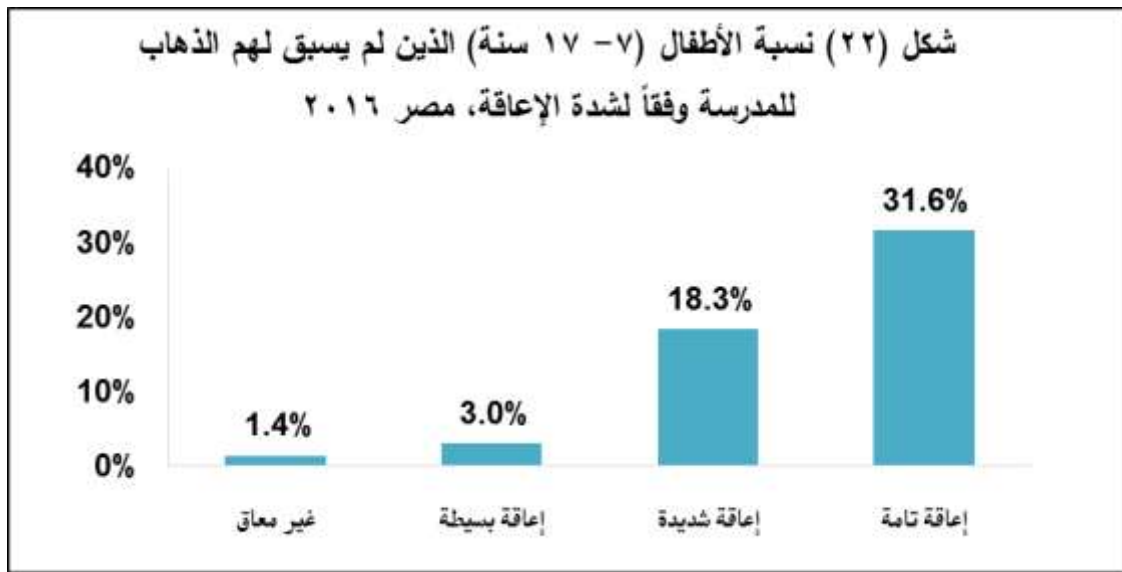
4-3 الإعاقة والتعليم

يهتم هذا الجزء بالتعرف على الاختلافات بين الأفراد ذوي الإعاقة مقارنة بالأفراد غير المعاقين وفقاً للحالة التعليمية، وسوف يتم التركيز على كل من تعليم الأطفال في الفئة العمرية (7-17 سنة)، بالإضافة إلى تعليم البالغين في الفئة العمرية (18 سنة فأكثر).

3-4-1 الإعاقة وتعليم الأطفال

يعرض شكل (22) نسبة الأطفال في الفئة العمرية (7 – 17 سنة) الذين لم يسبق لهم الذهاب للمدرسة، ويتضح من هذا الشكل أن الأطفال ذوي الإعاقة التامة يعانون من الحرمان من التعليم بشكل واضح مقارنة بغير المعاقين، فحوالي ثلث الأطفال ذوي الإعاقات التامة لم يلتحقوا بالمدرسة، وحوالي 18.3% من الأطفال ذوي الإعاقات الشديدة لم يلتحقوا بالمدرسة، بينما 1.4% فقط من الأطفال غير المعاقين لم يلتحقوا بالمدرسة.

وباختبار العلاقة بين الإعاقة والحالة التعليمية للأطفال باستخدام اختبار كاي²، فقد أظهر الاختبار معنوية العلاقة عند مستوى معنوية ($p < 0.01$).



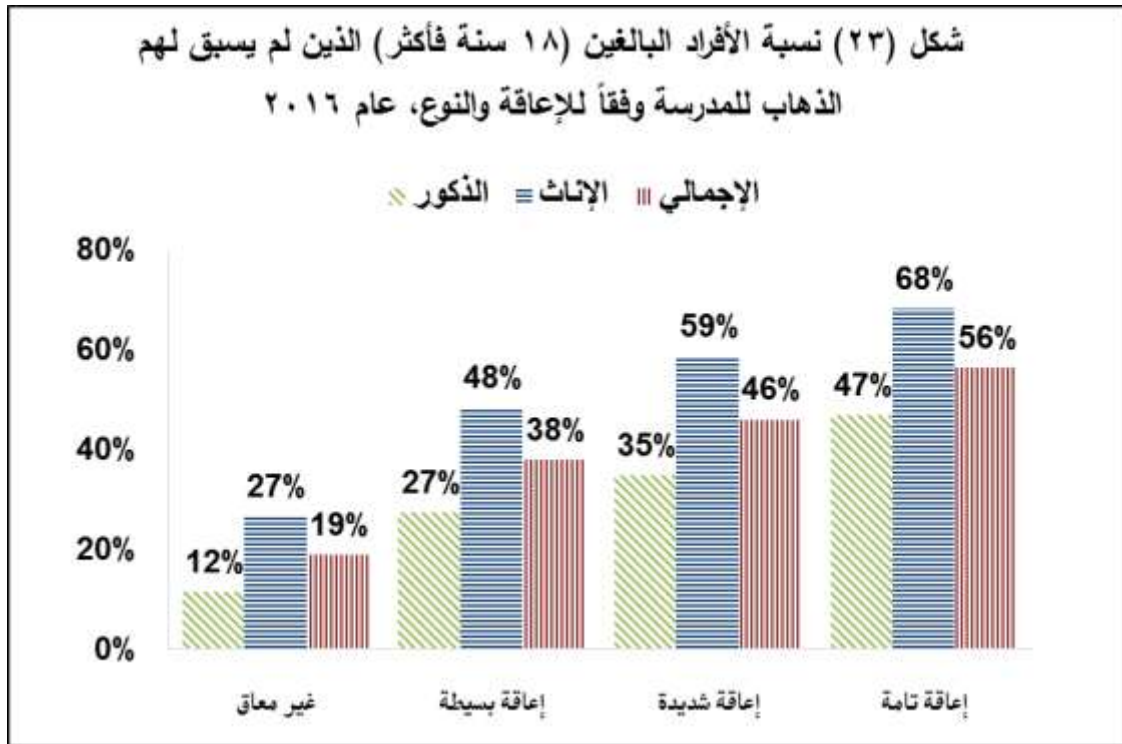
المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

3-4-2 الإعاقة وتعليم الأفراد البالغين

يعرض شكل (23) نسبة الأفراد البالغين (18 سنة فأكثر) الذين لم يسبق لهم الذهاب للمدرسة، ويتضح من هذا الشكل أن أكثر من نصف الأفراد البالغين ذوي الإعاقات التامة لم يلتحقوا بالمدرسة، وحوالي 46.0% من الأفراد ذوي الإعاقات الشديدة لم يلتحقوا بالمدرسة، بينما 19.0% من الأفراد غير المعاقين لم يلتحقوا بالمدرسة، وهذا يشير إلى وجود علاقة عكسية بين الإعاقة والالتحاق بالتعليم، فكلما زادت شدة الإعاقة كلما انخفضت نسبة الالتحاق بالتعليم. وعند إجراء الاختبارات الإحصائية تبين أن العلاقة بين الإعاقة والحالة التعليمية للأفراد البالغين معنوية ($p < 0.01$) وذلك باستخدام اختبار كاي².

أما عن العلاقة بين الإعاقة والتعليم حسب النوع، فتشير النتائج إلى أن الأفراد ذوي الإعاقة سواء ذكور أو إناث يعانون بشكل واضح من تدني المستوى التعليمي لهم مقارنة بالأفراد

غير المعاقين، وتزداد المعاناة مع زيادة شدة الإعاقة، فحوالي 47.0% من الذكور الذين يعانون من إعاقة تامة لم يسبق لهم الذهاب للمدرسة، وذلك مقارنة بحوالي 12.0% من بين الذكور غير المعاقين، كذلك فإن 68.0% من الإناث اللاتي يعانين من الإعاقة التامة لم يسبق لهن الذهاب للمدرسة، مقارنة بحوالي 27.0% بين الإناث غير المعاقات. أما بالنسبة للإناث المعاقات يزداد حرمانهم من التعليم مقارنة بالذكور المعاقين، حيث تتضاعف معاناتهن سواء بسبب النوع أو بسبب الإعاقة.



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

وقد تم دراسة احتمالات الانتقال من مرحلة تعليمية لأخرى، للتعرف على مدى التسرب من التعليم بين الأفراد ذوي الإعاقة التامة مقارنة بغير المعاقين، ويتضح من جدول (4) النقاط التالية:

- احتمالات الانتقال من مرحلة تعليمية لأخرى بين الأفراد ذوي الإعاقة التامة منخفضة مقارنة بالأفراد غير المعاقين، فإحتمال الالتحاق بالمرحلة الابتدائية بين الأفراد ذوي الإعاقة التامة يصل إلى 43.0% مقارنة بـ 80.0% بين الأفراد غير المعاقين.
- بالنسبة للأفراد الذين التحقوا بالمرحلة الابتدائية فإن احتمال الانتقال إلى المرحلة الإعدادية يصل إلى 53.0% بين الأفراد ذوي الإعاقة التامة، مقارنة بـ 83.0% بين الأفراد غير المعاقين.

- يُعد الانتقال إلى المرحلة الثانوية هو أفضل وضع بالنسبة للأفراد ذوي الإعاقة التامة، حيث يلاحظ أن الأفراد الذين التحقوا بالمرحلة الإعدادية يصل احتمال انتقالهم إلى المرحلة الثانوية إلى 85.0%، وذلك مقارنة بـ 86.0% بين الأفراد غير المعاقين. بينما تتسع الفجوة مجدداً بين الأفراد المعاقين وغير المعاقين عند الالتحاق بالمرحلة الجامعية/ المعاهد فوق المتوسطة.
- إذا نظرنا إلى توزيع 1000 من الأفراد ذوي الإعاقة التامة سوف نلاحظ التالي: أن 433 فقط منهم سوف يلتحقوا بالمرحلة الابتدائية، و 230 سوف يلتحقوا بالمرحلة الإعدادية، و 196 سوف يلتحقوا بالمرحلة الثانوية، و 76 سوف يلتحقوا بالمرحلة الجامعية/ المعاهد فوق المتوسطة.
- إذا نظرنا إلى توزيع 1000 من الأفراد غير المعاقين سوف نلاحظ التالي: أن 804 منهم سوف يلتحقوا بالمرحلة الابتدائية، و 667 سوف يلتحقوا بالمرحلة الإعدادية، و 572 سوف يلتحقوا بالمرحلة الثانوية، و 249 سوف يلتحقوا بالمرحلة الجامعية/ المعاهد فوق المتوسطة.

جدول (4) احتمالات الانتقال من مرحلة تعليمية لأخرى (Education Progression)

(Ratio) للأفراد (أكثر من 18 سنة)، مصر 2016

لديه إعاقة تامة		غير معاق		
توزيع 1000 من الأفراد ذوي الإعاقة التامة	احتمالات الانتقال من مرحلة تعليمية لأخرى	توزيع 1000 من الأفراد غير المعاقين	احتمالات الانتقال من مرحلة تعليمية لأخرى	
1000		1000		لم يسبق له الذهاب للمدرسة
433	%43	804	%80	إبتدائي
230	%53	667	%83	إعدادي
196	%85	572	%86	ثانوي
76	%39	249	%44	أعلى من ثانوي

المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

3-5 الإعاقة والمستوى الاقتصادي للأسرة

تم قياس المستوى الاقتصادي من خلال مؤشر الثروة والذي تم تكوينه باستخدام التحليل العاملي "Factor analysis" وذلك بالاعتماد على ممتلكات الأسرة وخصائص المسكن. وتشير النتائج أن الأفراد ذوي الإعاقة يعيشون في أسر أفقر مقارنة بباقي الأفراد كما يتضح من جدول (5)، فإذا نظرنا إلى الأفراد ذوي الإعاقة التامة سنلاحظ أن حوالي نصف هؤلاء الأفراد يعيشون في أسر فقيرة تنتمي للمستوى الأدنى والمستوى الثاني لمؤشر الثروة، بينما ينتمي 7.0% فقط منهم لأعلى مستوى، وهذا هو الوضع كذلك بالنسبة للأشخاص ذوي الإعاقة الشديدة. وعند إجراء الاختبارات الإحصائية تبين أن العلاقة بين الإعاقة ومؤشر الثروة معنوية ($p < 0.01$) وذلك باستخدام اختبار كاي²، كما تبين أن العلاقة عكسية ومعنوية من خلال حساب معامل جاما، فكلما زادت شدة الإعاقة كلما انخفض المستوى الاقتصادي.

جدول (5) التوزيع النسبي للأفراد وفقاً للإعاقة ومؤشر الثروة، مصر 2016

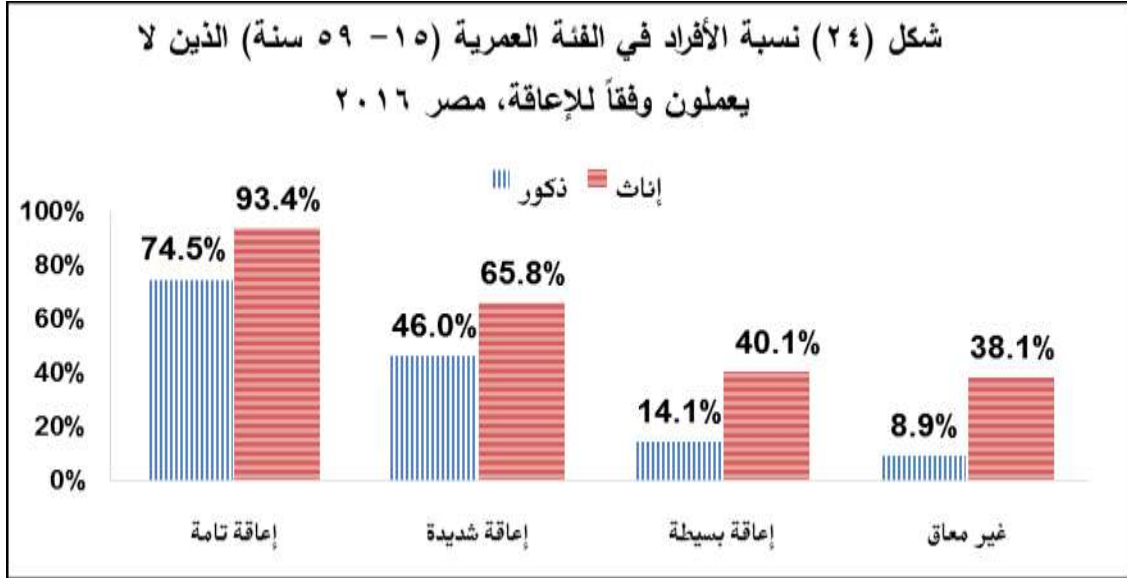
مؤشر الثروة	غير معاق	لديه إعاقة		
		إعاقة بسيطة	إعاقة شديدة	إعاقة تامة
أدنى مستوى	19.7%	23.9%	29.9%	25.3%
المستوى الثاني	20.5%	21.1%	20.5%	24.6%
المستوى المتوسط	19.7%	20.7%	20.1%	23.1%
المستوى الرابع	19.5%	19.5%	17.2%	19.9%
أعلى مستوى	20.6%	14.8%	12.3%	7.1%
الإجمالي	100.0%	100.0%	100.0%	100.0%

المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

3- 6 الإعاقة والحالة العملية

تؤثر الإعاقة على قدرة الأفراد على العمل، خاصة مع وجود بيئة عمل محيطية غير مهيأة لاستيعاب الأفراد ذوي الإعاقة، ويتضاعف هذا التأثير بالنسبة للإناث ذوي الإعاقة، ويتضح من شكل (24) أن 74.5% من الذكور ذوي الإعاقة التامة في الفئة العمرية (15- 59 سنة) لا يعملوا، وذلك مقارنة بحوالي 8.9% بين الذكور غير المعاقين.

بالنسبة للإناث ذوي الإعاقة التامة حوالي 93.4% منهن لا يعملن، وذلك مقارنة بحوالي 38.1% بين الإناث غير المعاقات. وعند استخدام اختبار كاي² كانت العلاقة بين الإعاقة والحالة العملية معنوية ($p < 0.01$) وذلك لكل من الذكور والإناث.



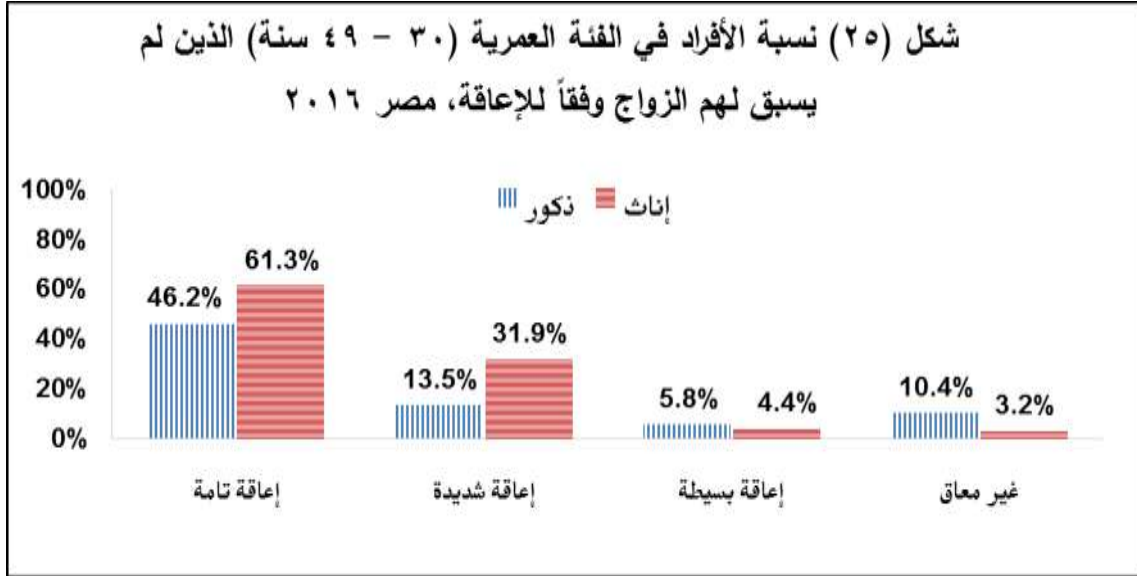
ملحوظة: تم استبعاد كل من الطلبة والمجندين والمسجونين والزاهدين عن العمل وربات المنازل.

المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

3-7 الإعاقة والحالة الزوجية

يعاني الأفراد ذوي الإعاقة من الحرمان من بعض الحقوق الإنسانية كالحق في الزواج، وقد يزداد هذا الحرمان بالنسبة للإناث ذوي الإعاقة، ويتضح من شكل (25) أن حوالي 46.2% من الذكور ذوي الإعاقة التامة في الفئة العمرية (30 - 49 سنة)⁶ لم يسبق لهم الزواج، وذلك مقارنة بحوالي 10.4% بين الذكور غير المعاقين. بالنسبة للإناث ذوي الإعاقة التامة حوالي 61.3% منهم لم يسبق لهم الزواج، وذلك مقارنة بحوالي 3.2% بين الإناث غير المعاقات. وعند استخدام اختبار كاي² كانت العلاقة بين الإعاقة والحالة الزوجية معنوية ($p < 0.01$) وذلك لكل من الذكور والإناث.

⁶ تم الاقتصار على هذه الفئة العمرية لأن من هم أقل من 30 سنة مازالت احتمالات زواجهم مرتفعة، وبالتالي تم الاقتصار على من يعانون من تأخر في سن الزواج، ولم يتم اختيار من هم أكبر من 49 سنة حتى لا يدخل في العينة الإعاقات الناتجة عن كبر السن (على سبيل المثال شخص أصبح معاق بعد سن الستين ولكنه متزوج منذ شبابه، فسوف يظهر هذا الفرد في العينة كمعاق ومتزوج، وبالتالي يقلل من تأثير الإعاقة على الزواج (under estimation).



المصدر: محسوب من قبل الباحثين من واقع بيانات مسح أحوال الأسرة المصرية لعام 2016.

4- الخلاصة والتوصيات

يعرف البحث الحالى الإعاقة بأنها تشير إلى حالة جسدية أو عقلية أو نفسية تؤثر وتحد من أنشطة الأفراد، ويتم اعتبار الفرد لديه إعاقة إذا كان يعاني من صعوبة واحدة على الأقل من الصعوبات الستة التي أعدها فريق واشنطن المعنى بإحصاءات الإعاقة والتي تشمل كل من: الرؤية، والسمع، والحركة، والتذكر أو التركيز، والرعاية الشخصية (الاستحمام، اللبس،....)، والتواصل مع الآخرين (مثل فهمه للغير أو فهم الغير له)، أيا كانت درجة صعوبة الإعاقة (بسيطة أو شديدة أو تامة).

وتساهم نتائج البحث الحالى فى رسم ملامح خريطة الإعاقة فى مصر، فوفقاً للتعريف الذى تبنته الدراسة تبلغ نسبة الإعاقة فى مصر 10.4% فى عام 2016 من إجمالى السكان، وبينما تعتبر هذه النسبة مرتفعه إذا ما قورنت بالنسبة التى تم حسابها من تعداد 2006 (تم تقديرها بحوالى 0.7%)، إلا أنها لا تعتبر مرتفعة مقارنة بالمستويات العالمية، حيث تصل نسبة الإعاقة فى العالم إلى 15%، كما تتراوح بين 11.0% و 14.0% فى الاتحاد الأوروبى، وتصل فى الولايات المتحدة الأمريكية إلى 12.6%.

لا يختلف انتشار الإعاقة حسب النوع أو حسب محل الإقامة حضر/ ريف، بينما تتباين حسب المناطق الجغرافية، حيث تصل إلى أقصاها فى حضر الوجه البحرى وأدناها فى ريف الوجه القبلى وفى محافظات الحدود، وقد يرجع ارتفاع نسب انتشار الإعاقة فى المناطق الأكثر تحضراً لارتفاع الوعى فى هذه المناطق وبالتالي القبول بالإدلاء عن وجود إعاقة بين الأفراد. كما تتباين نسب الإعاقة حسب فئات العمر حيث ترتفع بين فئات العمر المتقدمة.

وتشكل الإعاقات البسيطة النسبة الأكبر من الإعاقات، فحوالي 7.5% من الأفراد لديهم إعاقة بسيطة، و 2.1% من الأفراد يعانون من إعاقة شديدة، و 0.8% يعانون من إعاقة تامة، وتعد الإعاقات الخاصة بالرؤية أكثر الإعاقات التي يعاني منها الأفراد في مصر، يليها الحركة ثم السمع، بنسبة 5.9%، 4.8%، 2.1% على التوالي.

وتجدر الإشارة إلى أن تفاوت نسب الإعاقة بشكل عام قد يرجع إلى العديد من الأسباب والتي منها أسلوب القياس والذي يتأثر بأسلوب الإدراك الذاتي للأشخاص حيث يكون الشخص هو المسؤول عن الإدلاء بالمرض/ الإعاقة ويعاني هذا الأسلوب من مشاكل في المصادقية والاتساق، ويتأثر باللغة المستخدمة، وموقع الأسئلة في الاستمارة، والنسيان، ووقت إلقاء السؤال، ومن هو المجيب على الاستمارة. وعلى الجانب الآخر تلعب التقاليد والمورثات الاجتماعية دورا في الاعتراف بالإعاقة، فالمرض/الإعاقة هو ظاهرة اجتماعية وثقافية تتشكل بواسطة السياق الثقافي للمجتمع، فاستعداد الشخص للاعتراف بالمرض/الإعاقة يعتمد على القبول الاجتماعي لأنواع معينة من الأمراض والإعاقة، فعلى سبيل المثال يعد الصرع في مجتمع وصمة عار بينما هو في مجتمع آخر مرض عادي، وبعض الأفراد يرفضوا الإفصاح عن المرض/الإعاقة ويتعمدوا الإدلاء بمعلومات غير صحيحة لأسباب مرتبطة بهم. وبالتالي فإن الإدراك الذاتي للمرض الإعاقة هو دالة في كل من المرض/ الإعاقة من ناحية والسياق الاجتماعي والثقافي للفرد من ناحية أخرى.

واستكمالا لمعالم خريطة المعاقين في مصر، فقد أشارت النتائج إلى أن الإعاقة تلعب دورا في عدم تمتع الأفراد المعاقين بحقوقهم، فقد كان لها تأثير سلبي على الالتحاق بالتعليم سواء للأطفال أو البالغين، فحوالي ثلث الأطفال ذوي الإعاقات التامة لم يلتحقوا بالمدرسة، بينما 1.4% فقط من الأطفال غير المعاقين لم يلتحقوا بالمدرسة، كما ينخفض المستوى التعليمي للبالغين ذوي الإعاقات، وتتنخفض احتمالات انتقالهم من مرحلة تعليمية لأخرى، وتزداد وطأة التأثير بين الإناث مقارنة بالذكور. كما يعيش الأفراد ذوي الإعاقة في أسر فقيرة مقارنة بباقي أسر الأفراد غير المعاقين. وتعتبر الإعاقة مانع من الالتحاق بسوق العمل حيث أشارت النتائج أن حوالي ثلاثة أرباع الذكور ذوي الإعاقة التامة في الفئة العمرية (15- 59 سنة) لا يعملوا، وذلك مقارنة بحوالي 8.9% بين الذكور غير المعاقين، وحوالي 93.4% من الإناث ذوي الإعاقة لا يعملن، وذلك مقارنة بحوالي 38.1% بين الإناث غير المعاقات، كما تؤثر الإعاقة على الحالة الزوجية للأفراد، فحوالي 46.2% من الذكور ذوي الإعاقة التامة في الفئة العمرية (30 – 49 سنة) لم يسبق لهم الزواج، وذلك مقارنة بحوالي 10.4% بين الذكور غير المعاقين، وحوالي 61.3% من الإناث ذوي الإعاقة لم يسبق لهم الزواج، وذلك مقارنة بحوالي 3.2% بين الإناث غير المعاقات.

وجدير بالذكر أن هذا النمط لا يختلف كثيرا مقارنة بالوضع في الدول العربية، فقط أشارت الدراسة التي قامت بها (اللجنة الاقتصادية والاجتماعية لغربي آسيا (الإسكوا) 2014) عن الاعاقة في الدول العربية، أن الأفراد المعاقين يعانون من انخفاض مستويات التشغيل مقارنة بغير المعاقين، كما أوضحت هذه الدراسة تدني مستويات التعليم بين الأشخاص ذوي الإعاقة مقارنة بغير المعاقين، كما لا يختلف ذلك عن الوضع في بعض الدول الأوروبية (Hosain et al. 2002, Filmer 2008 and Mitran et al 2011).

وتوصى الدراسة بإجراء مسوح قومية تتضمن بيانات أكثر تفصيلاً عن المعاقين وظروف إعاقته، مثل تاريخ الإعاقة، ومن ثم العمر عند الإصابة بهذه الإعاقة، أوضاع العمل بشكل أكثر تفصيلاً، بيانات عن الوالدين من حيث التعليم والعمل والإعاقة وهو ما يسمح برسم خريطة أكثر تفصيلاً عن أوضاع ذوي الإعاقة في مصر. كما يمكن الاستعانة بمجموعة اسئلة مجموعة واشنطن المفصلة فى المسوح المتخصصة.

كما توصى الدراسة بمزيد من البرامج الثقافية والتوعوية لإدماج ذوي الإعاقة في المجتمع، مما يمكن أن يكون له أثر إيجابى على ثقافة عدم الإدلاء بالإعاقة، وهو ما انعكس في التفاوتات الكبيرة في نسب انتشار الإعاقة بين المحافظات وفقاً لدرجة التحضر المختلفة، بالإضافة إلى أهمية توجيه برامج صحية خاصة بذوى الإعاقة وخاصة فى أكثر أنواع الإعاقة انتشاراً (الرؤية والحركة) مما قد يساعد على تمتع المعاق بالخدمات المختلفة كالتعليم والاندماج بسوق العمل ويرفع من جودة الحياة لهم.

المراجع

المراجع باللغة العربية:

1. الأمم المتحدة، (2006)، "اتفاقية حقوق الأفراد ذوي الإعاقة، والبروتوكول الاختياري".
2. الأمم المتحدة، (2015)، "تحويل عالمنا: خطة التنمية المستدامة لعام ٢٠٣٠". الجمعية العامة، الدورة السبعون، البنود ١٥ و ١١٦ من جدول الأعمال، قرار اتخذته الجمعية العامة في ٢٥ أيلول/سبتمبر ٢٠١٥.
3. الجهاز المركزي للتعبئة العامة والإحصاء، (2006)، "النتائج النهائية لتعداد العام للسكان والإسكان والمنشآت لعام 2006".
4. الجهاز المركزي للتعبئة العامة والإحصاء، (2017)، "النتائج النهائية لتعداد العام للسكان والإسكان والمنشآت لعام 2017".
5. الديب، بثينه (2007). "التقرير الوطني عن إحصاءات الإعاقة في مصر".
6. اللجنة الاقتصادية والاجتماعية لغربي آسيا (الإسكوا)، جامعة الدول العربية، (2014)، "الإعاقة في المنطقة العربية – لمحة عامة".
7. مركز المعلومات ودعم اتخاذ القرار، (2016)، "مسح أحوال الأسرة المصرية لعام 2016".

المراجع باللغة الانجليزية:

1. Filmer, D. 2008. "Disability, Poverty, and Schooling in Developing Countries: Results from 14 Household Surveys". *The World Bank Economic Review*, Vol. 22, No. 1, pp. 141–163. Oxford University Press.
2. Hosain, G.M., Atkinson, D. and Underwood, P. 2002. Impact of disability on quality of life of rural disable people in Bangladesh. *Journal of Health Population and Nutrition*, Vol.20(4), pp. 297-305
3. Institute on Disability 2016. "Disability Statistics Annual Report 2015". University of New Hampshire.
4. Mitra, S., Posarac, A., & Vic, B. 2011. "Disability and Poverty in Developing Countries: A Snapshot from The World Health Survey". The World Bank Social Protection and Labor Unit.
5. Mont, D. 2007. "Measuring Disability Prevalence". The World Bank Social Protection and Labor Unit.
6. United Nations, Economic and Social Commission for Asia and the Pacific. 2014. "ESCAP Guide on Disability Indicators for the Incheon Strategy". Bangkok: United Nations, Economic and Social Commission for Asia and the Pacific (ESCAP).

7. United Nations, Economic and Social Commission for Asia and the Pacific (ESCAP). 2012. "**Disability at a Glance 2012: Strengthening the Evidence Base in Asia and The Pacific**". Bangkok: United Nations, Economic and Social Commission for Asia and the Pacific.
8. U.S. Census Bureau. 2016. "**2014 American Community Survey 1-Year Estimates**".
9. World Health Organization and The World Bank. 2011. "**World report on disability 2011**". Malta: World Health Organization.

المواقع الإلكترونية:

1. Eurostat: online data codes: hlth_dp010, hlth_dp060 and hlth_dpeh005. Accessed at October 15, 2017.
2. Eurostat:
http://ec.europa.eu/eurostat/statistics-explained/index.php/Disability_statistics.
Accessed at October 15, 2017.
3. Centers for Disease Control and Prevention:
<https://www.cdc.gov/media/releases/2015/p0730-us-disability.html>. Accessed at November 1, 2017

الملاحق

جدول (أ)

مجموعة الأسئلة الموجزة التي أعدها فريق واشنطن المعني بإحصاءات الإعاقة

يتم سؤال جميع أفراد الأسرة هذه الأسئلة الستة:

- 1- هل يواجه الفرد صعوبات في الرؤية (حتى لو كان مرتدياً نظارة) ؟
- 2- هل يواجه الفرد صعوبات في السمع (حتى ولو كان يستخدم وسيلة سمع مساعدة) ؟
- 3- هل يواجه الفرد صعوبات في التذكر أو التركيز ؟
- 4- هل يواجه الفرد صعوبات في المشي أو صعود السلالم ؟
- 5- هل يواجه الفرد صعوبات في رعاية نفسه (الاستحمام اللبس،....) ؟
- 6- هل يواجه الفرد صعوبات في التواصل مع الآخرين باستخدام لغته (مثل فهمه للغير أو فهم الغير له) ؟

فئات الإجابة لهذه الأسئلة الستة هي:

- 1- لا يواجه صعوبة
- 2- نعم بعض الصعوبة (صعوبة بسيطة)
- 3- نعم صعوبة كبيرة (صعوبة شديدة)
- 4- لا يستطيع مطلقاً.

جدول (ب) تعريفات الإعاقات وشدتها (بسيطة/ شديدة/ تامة)

أنواع الإعاقات	إعاقة بسيطة	إعاقة شديدة	إعاقة تامة
<p>1- الرؤية: الأشخاص الذين لديهم بعض المشاكل في الرؤية والتي قد تحد من قدرتهم على أداء واجباتهم اليومية</p>	<ul style="list-style-type: none"> • يمكن معالجتها بالنظارة، ويمكنه التحرك بدون صعوبة 	<ul style="list-style-type: none"> • لا يمكن معالجتها بالنظارة 	<ul style="list-style-type: none"> • لا يستطيع رؤية أي شيء عدا الضوء، يحتاج إلى عصا للمشي
<p>2- الحركة: الأشخاص الذين لديهم بعض المشاكل في التنقل والتجول على الأقدام والتي قد تحد من أدائهم للأنشطة اليومية</p>	<ul style="list-style-type: none"> • بعض الصعوبات في الجلوس الطويل • لا يحتاج إلى عون أو مساعدة • يستطيع المشي بدون مساعدة، وقد يعرج أو يمشى بطرف صناعي • قد يعاني من مشاكل في صعود السلالم العالية أو جر قدميه 	<ul style="list-style-type: none"> • يحتاج إلى عون/ مساعدة للجلوس منتصبًا • يحتاج مساعدة قصوى للمشي، لكنه يستطيع التحرك 	<ul style="list-style-type: none"> • غير قادر على المحافظة على توازنه بدون عون. • يحتاج مساعدة تامة عندما يجلس منتصبًا • لا يستطيع المشي مطلقًا
<p>3- السمع: الأشخاص الذين لديهم بعض المشاكل في السمع والتي تساهم في الحد من قدرتهم على أدائهم للأنشطة اليومية</p>	<ul style="list-style-type: none"> • بعض الصعوبات في السمع • يتكلم بصورة طبيعية • قد يستخدم أداة تساعده على السمع (سماعة) 	<ul style="list-style-type: none"> • صعوبة في السمع حتى مع استخدام أداة مساعدة (سماعة) • تؤثر على طريقة تعبيره بالكلام 	<ul style="list-style-type: none"> • لا يوجد دليل/ مؤشر على السمع
<p>4- التذكر أو التركيز: الأفراد الذين يجدون صعوبة في الفهم والإدراك للقيام بالأنشطة اليومية مثل اتخاذ القرارات، فهم الكلام، قراءة شيء مكتوب، التعرف على الأفراد، الاستدلال على المناطق الجغرافية، العمليات الحسابية</p>	<ul style="list-style-type: none"> • جدير بالذكر أن قياس هذه الصعوبة يرتبط بقياس معدل الذكاء، ولكن نظرًا لصعوبة القياس يتم سؤال الأفراد مباشرة عن شدة الصعوبة. 		

جدول (ب) تعريفات الإعاقات وشدتها (بسيطة/ شديدة/ تامة)

أنواع الإعاقات	إعاقة بسيطة	إعاقة شديدة	إعاقة تامة
<p>5- الرعاية الشخصية: الأفراد الذين لديهم بعض المشاكل في استخدام الأيدي والأصابع لمسك الأدوات أو الكتابة وغيرها، وكذلك عدم القدرة على العناية الشخصية بنفسه مثل الاستحمام واللبس</p>	<ul style="list-style-type: none"> • الإمساك بالأشياء من الممكن أن يكون ضعيفاً أو غير مُحكم، لكن يستطيع إمساك معظم الأدوات (قلم، سكين، كوب....) • يستطيع ارتداء ملابسه بنفسه • قد لا يستطيع رفع زراعته فوق مستوى رأسه بصورة كاملة 	<ul style="list-style-type: none"> • صعوبة في الإمساك بالأدوات واستخدام الزراعين لارتداء الملابس • عدم القدرة على التقاط شيء في حجم حبة الدواء الصغيرة • في حالة الشلل النصفي، تكون إحدى اليدين غير مستعملة تماماً 	<ul style="list-style-type: none"> • لا يستطيع استخدام أي من اليدين في أي حركة إلا في مجرد الوصول إلى شيء أو الإشارة باليد
<p>6- التواصل مع الآخرين: عدم قدرة الشخص على تبادل المعلومات والأفكار مع الآخرين والتعامل معهم من خلال استخدام النطق، أو الإشارة أو الحركة أو الكتابة للمعلومات التي يرغبون بتبادلها مع الآخرين، وقد يكون ذلك ناتجاً عن عجز في السمع أو الكلام، أو عدم القدرة على تفسير وإدراك ما يقوله الآخرون من إشارات وكلمات وحركات</p>	<ul style="list-style-type: none"> • يتكلم جملاً كاملة (حسب العمر) • يمكن فهمه، لكن كلامه ليس مطولاً • لدرجة أن تستطيع استخلاص أكثر من فكرة بسيطة منه 	<ul style="list-style-type: none"> • يتكلم كلمات منفردة تعبر عن الاحتياجات الأساسية 	<ul style="list-style-type: none"> • لا يتكلم مطلقاً

العوامل المؤثرة على نية ريادة الأعمال بين الشباب المصرى إيمان أحمد على إسماعيل*

1- مقدمة

يشير مصطلح الشباب الى تلك المرحلة العمرية التى ينتقل فيها الانسان من مرحلة الطفولة الى النضج ، ومع هذا فان بعض الادبيات ما توسع المدى العمرى لمرحلة الشباب لتبدأ من سن الخامسة عشرة وتنتهى مع سن الخامسة والثلاثون . وسوف يستخدم هذا البحث هذا التعريف الموسع للشباب . هذا ويشكل الشباب طبقا لهذا التعريف نسبة 37.38 % من جملة سكان جمهورية مصر العربية طبقا لبيانات الجهاز المركزى للتعبئة العامة والاحصاء (2016) .

ولعل من احد اسباب الأخذ بهذا التعريف الموسع هو ان كثير من الأدوار والمسئوليات المتعلقة بسن النضوج عادة مايقوم بها الأفراد مبكرا فى افريقيا ومصر فى سن اصغر كثيرا من سن الخامسة عشرة سواء كان العمل او الزواج المبكر

وتكمن أهمية الشباب فى مصر فى أنه أكثر ما تملكه مصر من حيث القيمة والذى يمكن توجيهه إلى مصلحة الإقتصاد والخروج من الأزمة الإقتصادية التى تعاني منها أغلب دول المنطقة فى شمال افريقيا والشرق الأوسط ، وبالفعل فهؤلاء الشباب يمكن أن يكونوا مصدر قلق وخطر إذا لم تستوعبهم فرص العمل

1-1 تعريف ريادة الأعمال والمشروعات الصغيرة

Entrepreneurship and Small and Medium Enterprises

عادة ما يتم استخدام كلا من اللفظين كبديلين لنفس المعنى ومع ذلك هناك بعض الإختلافات التى يجب أخذها فى الإعتبار، فالإثنان يقعان تحت نفس تصنيف العمل لحساب النفس أو أصحاب المهن الحرة، ولكن رائد الأعمال يريد أن يخلق أشياء جديدة ويطورها مع مرور الوقت ولكن صاحب المشروع غالبا مايرضى بحال مشروعه ويريد استمراره فقط ، كذلك فان رائد الأعمال غالبا ما يميل إلى إستخدام التكنولوجيا ويميل إلى الإبتكار أما صاحب المشروع فهو يعمل ما يعمل ككثيرون أمثاله كإفتتاح محل تجارى أو صالون تجميل أو مطعم ، ورائد الأعمال يسعى كذلك لتحقيق شىء أكبر ثم ينتقل إلى شىء آخر ولكن صاحب المشروع سعيد بالبقاء كما هو عليه، كذلك رائد الأعمال يحب المخاطرة ويريد تغيير العالم أما صاحب المشروع يريد العمل الآمن فقط وتحقيق مكسب.

*استاذ مساعد بالأكاديمية العربية للعلوم والتكنولوجيا والنقل البحرى.

جامعة القاهرة – معهد الدراسات والبحوث الإحصائية

2-1 أمثلة على مشروعات ريادة الأعمال

من امثلة ريادة الأعمال عالميا الآتى:

- Mark Zuckerberg (Facebook)
- Reed Hastings (Netflix)
- Jeff Bezos (Amazon)
- Sergey Brin and Carry Page (Google)
- Bill Gates (Microsoft)

3-1 أهمية البحث

- 1 – يمثل الشباب الشريحة الأكبر من السكان فى مصر بنسبة 37.38% حسب بيانات الجهاز المركزى للتعبئة العامة والإحصاء .
- 2 – عدم قدرة القطاع الخاص المصرى على إستيعاب الشباب الداخل سنويا سوق العمل ، بالإضافة إلى ترهل القطاع الحكومى وانخفاض انتاجيته ، مما إنعكس على معدلات البطالة المرتفعة بين الشباب المصرى.
- 3 – الدور الذى يلعبه ريادة الأعمال والتوسع فى البرامج التعليمية خصوصا الجامعية المتناوله له فى إقتصاديات الكثير من الدول - بعيدا عن الدراسات التقليدية ككليات الإدارة – حيث أنه أصبح مسارا إختياريا للعديد من الشباب كمسار وظيفى ..

4-1 أهداف البحث

- 1 – التعرف على انتشار النية على ريادة الأعمال بالإعتماد على بيانات مسح الشباب المصرى 2014.
- 2 – التعرف على معوقات والعوامل المساعدة والمؤثرة على وجود النية لريادة الأعمال بين الشباب المصرى.

5-1 الدراسات السابقة

- 1 – يرى (Chigunta) وآخرون 2005 ان ريادة الأعمال فى دول شرق وجنوب قارة افريقيا هى احد الأساليب التى تحسن من اساليب المعيشة للشباب فى الدول النامية وان هذه الفكرة تحظى الآن باهتمام الكثير من واضعى السياسات والأدبيات العالمية خصوصا فى قارة افريقيا التى تتميز بزيادة أعداد الشباب بها وتعانى من مشكلة البطالة بينهم . وقد وجدت هذه الدراسة انه لا بد من تضافر جهود الدولة مع الجهود المجتمعية والنظمات التى تعمل لصالح الشباب مع الشباب انفسهم (ILO, 2005).

- 2 – يهدف (Talas) وآخرون 2013 إلى تحديد العوامل الديموجرافية وأثرها على نية زيادة الأعمال بين شباب خريجي أحد الجامعات التركية كإختيار وظيفي . وقد أسفرت الدراسة إلى أن نوع الكلية والدراسة الثانوية ودخل الأسره كان من أكثر العوامل المعنوية إحصائيا لتلك الدراسة (Talas et al, 2013).
- 3 – يرى (Khuong & Nguyen) 2016 أثر الصفات الشخصية والبيئة الخارجية والعادات الاجتماعية على زيادة الأعمال من خلال ردود الأفعال الإيجابية والسلبية من خلال دراسة على طلبة أحد الجامعات في فيتنام . وتظهر أهمية الدراسة في مدى انتشار مشاريع زيادة الأعمال في فيتنام إلى أنها أصبحت تساهم بنسبة 40% في الناتج المحلي الأجمالي لفيتنام . وهنا يظهر الأختلاف بين هذا البحث وبين أبحاث أخرى حيث يتسائل هذا البحث ما الذي يدفع الشباب إلى أن يتركوا الوظائف ذات الراتب الأعلى ويفضلون عليها الأخذ بمخاطر التمويل والقيام بعملهم الخاص (Khuong & Nguyen, 2016).
- 4 – يرى (Basu) وآخرون 2004 أن المؤسسات الصغيرة في دول إفريقيا جنوب الصحراء الكبرى مثل غانا وتنزانيا لديها مشكلات كثيرة في التمويل خصوصا من البنوك الرسمية ، وأن هناك فقط من خمسة إلى ستة في المائة فقط تستطيع الوصول إلى التمويل البنكي ، ونظرا لإزدياد الفقر وانتشاره ولزيادة الطلبات لإنشاء المشروعات الصغيرة فإنه تم إنشاء مؤسسات للتمويل لهذه الأغراض في إفريقيا ، وتعتمد هذه المؤسسات في تمويلها على الجهات الحكومية والجهات المانحة ولكن مواردها أيضا محدودة . ويرى البحث ضرورة استقلال هذه المؤسسات في عملية اتخاذ قراراتها بما في ذلك تحديد معدلات الإقراض بنسبة ربح مناسبة. وقد أوضحت الدراسة أن هناك تعاون متزايد بين هذه المؤسسات والنظام البنكي الرسمي في إفريقيا وهذا من شأنه أن يقوى الصلة بين الأنشطة الاقتصادية التي تتم في القطاعين الرسمي وغير الرسمي ، ولكن لا يزال هناك العديد من العقبات أمام هذه المؤسسات مثل نقص الخبراء والفنيين العاملين في مثل هذه المؤسسات (Basu et al, 2004).
- 5 – وتعرض منظمة العمل الدولية 2012 المشكلة العالمية التي تواجه الشباب حول العالم وتتمثل في إيجادهم فرص عمل محترمه وأنه خلال السنوات الأخيرة فإن الثورتين التونسية والمصرية قد نقلت إلى الشباب الأوروبي ووصولاً إلى الشباب الأمريكي مايسمى بالجرعة من أجل الوظيفة والتي من أهدافها الحرية السياسية وإقتصاد أكثر أماناً وإستقراراً . وأشارت هذه الدراسة أنه حتى قبل الأزمة المالية العالمية فإن هناك دولاً مثل الصين والهند أصبح لديها مشكلة في إيجاد فرص عمل جيدة لشبابها والذي أصبح يقبل بنوعية أقل كثيراً من الوظائف في الإقتصاد غير الرسمي . واقترحت الدراسة إلى ضرورة إشراك الشباب في عملية اتخاذ القرارات المتعلقة بهم وضرورة الأهتمام بالتدريب الفنى والمهني لإيجاد المزيد من فرص العمل للشباب (ILO, 2012).

6 – تؤكد هالة السعيد وآخرون في هذا البحث (2014) على أهمية المشروعات الصغيرة والمتوسطة على نمو الاقتصاد المصري وإستدامته. وعلى الرغم من ان الإصلاحات البنكية كانت قد بدأت من عام 2004 إلا أن الوصول إلى التمويل المناسب لا يزال يعتبر من أكبر المشاكل التي تواجه المشروعات المتوسطة والصغيرة في مصر، ومن أجل هذا السبب وحتى يمكن زيادة الخدمات التمويلية المقدمة لتلك المشروعات ولزيادة مزايا الإصلاحات البنكية فإن البنك المركزي المصري قد قام بمبادرة عام 2008 كجزء تكميلي للإصلاح البنكي ومرحلة ثانية له لتعزيز نصف المشروعات المتوسطة والصغيرة قد تركزت جغرافيا في ثلاث محافظات فقط وهي بالترتيب الشرقيه والقاهره والغربية، كما ان الأنشطة الاقتصادية قد تركزت في نشاطين على وجه الخصوص وهم التصنيع والتجارة، واخيرا وجدت الدراسة انه حوالي 50% فقط من المشروعات المتوسطة والصغيرة هي فقط التي تتعامل وتستفيد من الخدمات التمويلية البنكية (El- Said et al, 2014).

7 – ويرى Dobbs & Maddgavkar, 2014 أنه في قارة اسيا في العقود الثلاثة الماضية فإن العولمة والتكنولوجيا قد اعادت تشكيل الأقتصاديات حول العالم وهذا ادى الى خلق حوالي 900 مليون فرصة عمل غير زراعي في الدول النامية وهذا ساعد مئات الملايين من الخروج من الفقر، ومع هذا فإنه لا يزال هناك الكثير من الشباب الغير عامل وبلغت معدلات البطالة حد الخطر في العديد من الدول، وقد وجدت هذه الدراسة انه بحلول العام 2030 سوف تزيد قوة العمل العالمية الى 3.5 بليون. وقد استنتجت هذه الدراسة ان عدم التوافق الذي عانت منه الأقتصاديات المتقدمة سوف يزيد من مدها في الدول النامية بسبب الفجوة بين مهارات العاملين وبين ما يريد اصحاب الأعمال (Dobbs & Maddgavkar, 2014).

8 – وقد اوضحت Adly 2014 ان الوصول الى التمويل في مصر هو السبب الثالث الذي يقف عقبة في سبيل نمو مشروعات ريادة الأعمال التي شملتها هذه الدراسة. وان التمويل هو أحد أهم المشاكل خصوصا وان التمويل البنكي هو التمويل الرسمي المتاح بالإضافة الى التمويل الشخصي، وان 37% من العينة قد قدمت طلبا للحصول على قرض بنكي، وان 42% فقط حصلوا على هذا التمويل، كما اوضحت الدراسة ان هؤلاء الذين لم يتقدموا للحصول على قرض بنكي كان السبب وراء ذلك هو ارتفاع معدلات الفائدة البنكية بالإضافة الى عدم تمكنهم من الحصول على الضمانات المطلوبة كما أشاروا الى تعقد الإجراءات وما تتطلبه من وقت وجهد في الإعداد والتقدم (Adly, 2014)

9 – وقد تناول المركز السويدي للتجارة والإستثمار 2015 مشروعات ريادة الأعمال في مصر بالمقارنه بمثيلاتها في السويد، وقد أوضحت هذه الدراسة المعوقات التي تواجهها هذه المشروعات في مصر بداية من

صعوبة الحصول على تمويل بنكى وصعوبة الإجراءات القانونية وانخفاض مشاركة الإناث وتركز مشروعات ريادة الأعمال فى مجالات التجاره والصناعة والسياحة والمطاعم وانخفاض مشروعات التكنولوجيا والمعلومات حيث انها اقل من 2% (The Swedish Trade and Invest Council,) (2015).

10 – وتلقى نانسى واخرون 2015 فى هذه الدراسة الضوء على ريادة الأعمال على أنها وسيلة لمساعدة من يفتقرون إلى إمكانيات تطوير المهارات المطلوبة للعثور على الوظائف لذا فإنها أصبحت من أولويات المنظمات الحكومية ، وتؤكد هذه الدراسة اعتمادا على دراسات حالة من شرق أفريقيا انه على الرغم من أهمية برامج ريادة الأعمال لتمكين الشباب من الوصول إلى فرص العمل إلا أنها ليست بما يكفى للتحسن المستدام لمستوى معيشة المشاركين فيها نظرا لأن العوامل البيئية التى تلعب دورا محوريا وبدون الإهتمام بهذه العوامل فإن ريادة الأعمال قد لاتستطيع تخفيف حدة الفقر اذا لم تؤخذ هذه العوامل فى الإهتمام. وتتلخص هذه العوامل فى عوامل خاصة بالدولة مثل اللوائح والقوانين والتشريعات المختلفة واخرى خاصة ببيئة المجتمع المحلى واخرى خاصة بالإقتصاد (نانسى و اخرون، 2015)

ونلاحظ على هذه الدراسات مايلى

تتشابه ظروف مشروعات ريادة الأعمال فى كثير من الدول النامية فى افريقيا والشرق الاوسط كما يلى:

- 1 – وجود شريحة كبيرة من الشباب بالإضافة الى معدلات البطالة العالية لهؤلاء الشباب.
- 2 – انخفاض جودة التعليم بين هؤلاء الشباب.
- 3 – عدم ادراج برامج ريادة التعليم فى المسارات التعليمية فى المرحلة السابقة للتعليم الجامعى او مرحلة التعليم الجامعى.
- 4 – اتجاه الشباب نحو مشاريع ريادة الأعمال ليس اختياريا او بسبب وجود الأفكار الإبتكارية لدى هؤلاء الشباب وانما هو مسار يتم اختياره لرغبتهم فى التوظيف.
- 5 – عدم الإهتمام بمجالات التكنولوجيا والمعلومات واغلب مشروعات ريادة الأعمال هى فى مجالات التجارة (دون الإهتمام بالتصدير) او السباحة والتصنيع والفنادق والمطاعم.
- 6 – انخفاض مشاركة الإناث عموما.
- 7 – صعوبات التمويل خصوصا البنكية وكثره الضمانات المطلوبة والإجراءات الورقية.

6-1 مصادر البيانات

تم الإعتماد على بيانات المسح التتبعي للنشء والشباب في مصر 2014 والذي قام به مجلس السكان الدولي. وتأتى الدورة الثانية من المسح والتي سبقها دورة أولى في عام 2009 لتتبع نفس الشباب لمعرفة التغييرات التي طرأت عليهم. كما تم الإعتماد على بيانات أولية لتعداد مصر 2016 بالإضافة إلى بيانات البنك الدولي

7-1 منهجية الدراسة

تم استخدام التحليل الوصفي تمثلا في الجداول والنسب ومعامل الإقتران بالإضافة إلى استخدام الاسلوب الإحصائي اللوجستي.

8-1 تنظيم الدراسة

تتكون هذه الدراسة من عرض للمقدمة وتعريف كل من ريادة الأعمال والعمل الحر ، ثم أهمية البحث وأهدافه والدراسات السابقة ، ومنهجية الدراسة والإسلوب الإحصائي المستخدم ، ثم متغيرات النموذج الإحصائي والإطار التحليلي المقترح في الدراسة ، ثم بعد ذلك يأتي عرض للتعليم والبطالة في مصر وانتشار المشاريع الصغيرة وريادة الأعمال في مصر ، ثم تأتي الخصائص الخلفية لشباب المسح وعواملهم الاسرية واخيرا نتائج معامل التوافق ونتائج نموذج الانحدار اللوجستي ثم النتائج والتوصيات واخيرا المراجع وملحق للجداول المستخدمة في الدراسة .

2-أسلوب التحليل الإحصائي:

1-2 الانحدار اللوجستي Logistic Regression

الانحدار اللوجستي يعتبر من الأساليب الشائعة الاستخدام في العلوم الاجتماعية. ويطبق أسلوب الانحدار اللوجستي عندما يكون المتغير التابع متغير وصفي ثنائي (Y) والذي يأخذ القيمة (صفر ، 1) ويتبع المنحني اللوجستي.

في هذا النموذج نقوم بتقدير احتمال ظهور الحدث (Y=1)، في حالة وجود المتغيرات المستقلة X's وهذا الاحتمال يأخذ

الصورة الآتية:

$$E(y | x_i) = P(y = 1) = \frac{e^{\sum_{i=0}^K \beta_i X_i}}{1 + e^{\sum_{i=0}^K \beta_i X_i}}$$

ويمكن التغلب على مشكلة الدالة الغير خطية بتحويلها إلى دالة خطية عن طريق استخدام تحويله اللوجيت Logit Transformation وذلك بالمعادلة الآتية :

$$\text{Logit}(P) = \ln \left[\frac{p(y=1)}{1-P(y=1)} \right] = \ln \left(\frac{p}{1-p} \right)$$

حيث تقع هذه القيمة بين $-\infty$ و ∞

$$-\infty \leq \ln \left(\frac{p}{1-p} \right) \leq \infty$$

وهذا يتفق مع الانحدار الخطي، ومن هنا يمكن التعامل مع هذا النموذج في الصورة التالية:

$$\ln \left(\frac{p}{1-p} \right) = \hat{B}_0 + \hat{B}_1 X_1 + \hat{B}_2 X_2 + \dots + \hat{B}_k X_k$$

حيث:

P : تشير إلى احتمال حدوث الحدث أي أن $P(y=1)$

$\left(\frac{p}{1-p} \right)$: تشير إلى نسب الترجيح * Odds Ratio

\hat{B}_k : معاملات الانحدار اللوجستي، وتشير إلى التغير الذي يحدث في Log odds نتيجة لتغير المتغير

المستقل بمقدار الوحدة حيث $K=1,2,\dots,k$

X_k : المتغيرات المستقلة.

\hat{B}_0 : ثابت الانحدار.

تقدير المعاملات:

يتم تقدير المعاملات في النموذج باستخدام طريق الإمكان الأكبر Maximum Likelihood التي تمكنا من تعظيم القدرة التنبؤية للنموذج في ظل الفرض العدمي $(H_0: \hat{B}_i = 0)$. ويمكن اختبار معنوية معامل الانحدار عن طريق الاختبار الإحصائي Wald Statistic الذي يعرف بأنه مربع النسبة بين المعامل للمتغير (i) إلى الخطأ المعياري له

$$\text{Wald_Statistic} = \left(\frac{\hat{B}_i}{SE(\hat{B}_i)} \right)^2$$

يتبع هذا الاختبار توزيع Chi-Square بدرجات حرية مساوية للواحد الصحيح في حالة المتغيرات المستمرة، أما في حالة المتغيرات Categorical Variables فإن درجات الحرية تساوي (عدد الطبقات-1).

ولتقييم أداء النموذج نحسب التغير في نسبة الإمكان الأكبر بإجراء اختبار لمقارنة نماذج مختلفة أو لتقييم أداء نموذج واحد فقط. وذلك لرفض أو قبول النموذج (غير معنوي أو معنوي على التوالي) ويعرف إحصائي الاختبار كالتالي :

$$LR(i) = -2 \left[\log Likelihood(\hat{B}_0) - \log Likelihood(\hat{B}_0, \hat{B}_1) \right]$$

حيث $\log Likelihood(\hat{B}_0)$ خاص بالنموذج بدون متغيرات، $\log Likelihood(\hat{B}_1)$ خاصة بالمتغير X_1 بعد إدخاله للنموذج، وأما في حالة مقارنة نموذجين فإن إحصائي الاختبار يتبع توزيع χ^2 بدرجة حرية تساوي الفرق بين درجات الحرية بين النموذجين وتكون كالتالي:

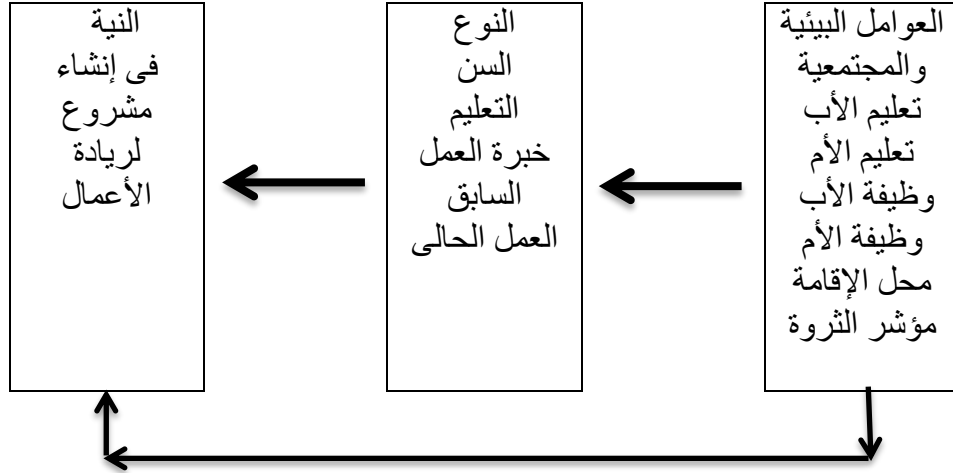
$$LR(i) = -2 \left[\log Likelihood(1^{st} \text{ model}) - \log Likelihood(2^{nd} \text{ model}) \right]$$

2-2 متغيرات النموذج

هذه الدراسة استخدمت النية لريادة الأعمال كمغير تابع ، وتم استخدام سبعة متغيرات مستقلة فقط ، وكانت الدراسة تهدف إلى استخدام مزيدا من المتغيرات ولكن لم تسمح البيانات الناقصة لكثير من استثمارات الشباب بهذا الأمر ، والمتغيرات المستخدمة في النموذج الإحصائي اللوجيستي هي العمر والجنس وخبرة العمل السابقة والعمل الحالي ومؤشر الثروة ومحل الإقامة والتعليم .

3-2 الإطار التحليلي للدراسة

ويشتمل هذا الإطار التحليلي على المحددات او العوامل المقترحة والتي تؤثر على رغبة الشباب المصري في إقامة مشروع ريادة أعمال، وبالطبع فإن تلك العوامل تنقسم على عوامل غير مباشرة واخرى مباشرة. وفيما يلي الإطار التحليلي المقترح للدراسة. ولكن نظرا لوجود العديد من مشاكل البيانات الناقصة فقد إقتصرت النموذج المستخدم على العوامل المباشرة فقط دون العوامل البيئية فيما عدا محل الإقامة والثروة.



4-2 التعليم في مصر

يعتبر التعليم طبقاً للدستور المصري حق مكفول لكل المصريين ، وخلال العقود الثلاثة السابقة فإن مصر قد حققت مزيداً من معدلات الالتحاق المرتفعة ، حيث انها زادت من 67 في عام 1980 الى 103.93 في 2014 وللذكور والإناث 104.076 و 103.774 بالترتيب (world bank, 2017)، ولكن التركيز على هذا الجانب فقط من التعليم – كما هو الحال في الكثير من الدول – يهمل عوامل أكثر أهمية من مجرد الالتحاق أو حتى الحصول على شهادات للمراحل الدراسية المختلفة، مثل جودة التعليم وجودة المدارس والمدرسين (Assaad, 2014) ليس هذا فحسب فقد وجدت الدراسات أن الأطفال في مصر أكثر عرضة للتسرب بسبب انخفاض جودة التعليم (Assaad and Krafft,2015).

ولانتزال الامية أحد أهم المشكلات التي تعاني منها مصر فيوجد حوالي 18.4 مليون امي في مصر وذلك بنسبة 25.8% من إجمالي السكان، ويبلغ عدد الذكور 7.8 مليون امي والإناث 10.6 ، وبالطبع فإن نصيب الريف أعلى من الحضر إذ يبلغ العدد بين الذكور في الريف 5.4 وعدد الإناث 7.5 ، في حين تبلغ الأعداد في الحضر 2.4 للذكور و3.1 للإناث وذلك طبقاً للنتائج الأولية لتعداد 2016 المصري.

إن تعليم ريادة الأعمال عادة ما يبدأ من المراحل الأولى من التعليم الابتدائي ويستمر ويتطور مع تقدم المراحل التعليمية حتى يصل إلى المستوى الجامعي حيث يشمل تعليم ريادة الأعمال فهم بعض الأسس الاقتصادية والأفكار التمويلية ومهارات حل المشكلات وبالطبع بالإضافة إلى مفهوم ريادة الأعمال وإضافتها للمجتمع (The Swedish Trade & Invest Council 2015).

وبالنسبة لحضانات الأفكار الإبداعية والإبتكارية والتي غالباً ماتكون موجودة في الجامعات والتي يتركز عملها في إعداد ورش العمل والتمويل والتدريب على إقامة المشاريع وتبنى الأفكار الإبتكارية

والإبداعية ورعايتها وتوفير المشوره المهنية والفنية ولا تزال عدد هذه الحضانات غير كافيا فى مصر ولعل أهمها تلك الموجودة فى الجامعة الأمريكية، ولكن هناك حضانات مستقلة موجودة بقوة فى مصر فى السنوات الأخيرة الماضية والتي تزيد من فرص إقامة مشاريع ريادة الأعمال فى مصر ومن ضمنها

Flat6Labs, Delta Inspire, Strat-up Egypt, Tamkeen, Endeavour Egypt, Intilaaqah Egypt, Nahdet Masr & Ayady

وتعانى مصر فى الآونة الأخيرة من ارتفاع معدلات البطالة بين حملة الدرجات الجامعية حتى ان هناك حوالى 70% من البطالة بين الشباب المصرى من خريجي الجامعات (Barsoum et al, 2015) وعلى الرغم من الإنتقادات الموجهه للتعليم المصرى ولكن تظل النسبة صارخه وتحتاج إلى مزيد من الأبحاث لحل هذه المشكلة

5-2 البطالة فى مصر

يواجه سوق العمل المصرى كمثلته فى دول الشرق الأوسط وشمال أفريقيا العديد من التحديات والتي تقيد من خلق المزيد من فرص العمل بين الشباب المصرى ، وبصفة عامة فإن بطالة الشباب هى الأعلى فى شمال أفريقيا والشرق الأوسط بمعدل 28.2 و 30.5 بالترتيب (ILO, 2015) ، ويبلغ معدل البطالة فى مصر 11.5 و للإناث 23.516 و للذكور (World Bank, 2017) ، وتكمن أحد المشاكل التى يعانى منها سوق العمل المصرى هو تلك الأعداد المتزايدة فى سن العمل الداخلة فى سوق العمل المصرى والتي تفوق محاولات سوق العمل لإستيعابها الأمر الذى يلقى بظلال ثقيلة على الأقتصاد المصرى بل والمجتمع كله، ومع الفرص القليلة التى يتيحها القطاع العام ، فإن الفجوة فى التوظيف لا يستطيع القطاع الخاص الرسمى سداها ، وبالتالي فإن نسبة العمالة الغير رسمية زادت وهذا أثر على نوعية الوظائف التى يلتحق بها هؤلاء الشباب ، هذا بالطبع بالإضافة إلى تلك المعضلة المتمثلة فى عدم التوافق بين مخرجات التعليم الرسمى وبين إحتياجات سوق العمل المصرى، بالإضافة إلى الفروق بين معدلات توظيف الذكور والإناث والفروق بين الريف والحضر، ولا يزال الشباب المصرى يفضل التوظيف فى القطاع العام وينظر للوظائف المهنية التقنية والمهنية بنظرة دونية (ILO, 2017).

ومن المعدلات التى استحدثتها منظمة العمل الدولية وهى تظهر حجم مشكلة العمالة فى مصر معدل الشباب فى المرحلة العمرية من 15 حتى 29 والغير مدرجين فى التعليم او التوظيف أو التدريب (NEET) وهى فى مصر 26.85 و للإناث والذكور على الترتيب 35.21 و 18.86 (World Bank, 2015) ، وتظهر أهمية تلك المعدلات فى أن معدلات البطالة تعبر عن الأشخاص الذين يبحثون عن عمل ولم يعملوا بعد ، أما معدلات (NEET) فإنها تعبر عن هؤلاء الغير باحثين عن عمل لأنهم يعتقدون أن فرص العمل

محدودة أو بسبب عدم قدرتهم على الحراك وراء فرص العمل أو بسبب تعرضهم للتمييز أو بسبب القيود الإجتماعية أو الثقافية ، وبالتالي هذه المعدلات تعكس بشكل أوسع طاقات الشباب الغير مستغلة (ILO, 2017)

2-6 وظائف المستقبل

حيث أن العالم يعيش الآن الثورة الصناعية الرابعة التى تتميز بالتشغيل الآلى وتبادل المعلومات، الأمر الذى يتطلب مزيدا من المهارات ويوفر مزيدا من فرص العمل الجديدة التى تتناسب مع هذه المرحلة ، وقد قدرت أن 49% من الأنشطة القائمة فى مصر سوف تتطور الكترونيا (Human Capital Index 2016). ومن ناحية اخرى ، فإن هذا التطور التكنولوجى الصاعد سوف يترتب عليه إزدياد معدلات البطالة خصوصا فى دول شمال أفريقيا والشرق الأوسط التى تعانى بالفعل من إرتفاع معدلات بطالة الشباب بها ، مما يستتبع معه تغييرات فى المهارات المطلوبة من الشباب الداخلى إلى سوق العمل وبالتحديد المهارات المتعلقة بالتكنولوجيا والمعلومات وهذا قد يسهم بالفعل فى زيادة فرص العمل فى التوظيف بالنسبة للإناث خصوصا إذا ماتم التوسع فى إستخدام الوظائف لإسلوب العمل من خلال الإنترنت حيث انه يوفر المرونة فى ساعات بل ومكان العمل وهو الأمر الذى أشارت إليه العديد من الأبحاث كأحد متطلبات الإناث للمشاركة فى سوق العمل.

ومع هذا فإن مشروعات ريادة الأعمال فى مجال التكنولوجيا فى مصر تحتل نسبة أقل من 2% فقط، ومن ضمن هذه المشروعات تطبيق بيكولك (Bey2ollak) والذى بدأ فى اكتوبر 2010 وكان وسيلة لتبادل المعلومات حول أماكن الإزدحام المرورى لتجنبها بل وإعطاء المسارات البديلة ، وأصبح مستخدمى هذا التطبيق الآن يفوق المليون شخص . كذلك فإن تطبيق دكتور بريدج (Dr Bridge) والذى بدأ فى إبريل 2012 والذى أصبح همزة وصل بين المريض والهيئات الطبية مجتمعة بما فى ذلك الأطباء ومعامل التحليل والأشعة .

ولكن هذه الوظائف لن تلغى بطبيعة الحال الوظائف التقليدية خصوصا مع الإحتياج الدائم إلى العديد من مشاريع البنية التحتية بالإضافة إلى الوظائف الخدمية مثل تلك التى فى قطاع التعليم والصحة.

3- ريادة الأعمال فى مصر

كما تم ذكره سابقا فان هناك العديد من نقاط الالتقاء بين المشروعات الصغيرة والمتوسطة وبين ريادة الأعمال بل ان هناك العديد من الأدبيات ما تجمع بينهما ، ولذا فى هذا الجزء سوف نتناول الدراسة كل منهما كالتالى :

1-3 أولا المشروعات الصغيرة والمتوسطة :

فى مصر حوالى 25 مليون مشروع صغير أو متوسط حسب بيانات المعهد المصرفى المصرى والبنك المركزى المصرى(2010) وهم يشكلون 75% من إجمالى قوة العمل العامل ، ويشكلون 99% من مؤسسات القطاع الخاص غير الزراعى ، وعلى الرغم من هذا فإن المشروعات الصغيرة لاتزال تعاني من العديد من المشاكل خصوصا الوصول إلى التمويل والذى يعتبر من السمات المشتركة بين الدول النامية. وفى الواقع فإنه حوالى 70% من الدول الغير أعضاء فى منظمة التعاون الإقتصادى – والتي تضم فى عضويتها الدول الأكثر تقدما وهم حاليا 35 دولة – تشير إلى فجوة عميقة فى التمويل مقارنة بنسبة 30% بين الدول الأعضاء، وبالطبع فإن تقليل تلك الفجوة التمويلية من شأنها ان تحسن من النمو الأقتصادى وتزيد من خلق المزيد من فرص العمل .

وترى البنوك المصرية أن تمويل المشروعات الصغيرة والمتوسطة هو أمر محفوف بالمخاطر ويتطلب تكاليف عالية وبالتالي يحقق ربحية أقل مقارنة بالكيانات الأكبر .

هذا وقد اجرى البنك المركزى المصرى والمعهد المصرفى المصرى تعدادا لحصر المشروعات الصغيرة والمتوسطة فى مصر(2010) والتي اظهرت بعض نتائج الأتى :

1 – ان المشروعات الصغيرة والمتوسطة يتركز حوالى نصفها تقريبا فى محافظات الشرقية والقاهرة والغربية (بنسبة 16% و15% و9.2% بالترتيب).

2 – ان تلك المشروعات تتركز فى قطاعين فقط وهما الصناعة والتجارة (51.1% و40.5% بالترتيب) .

3 – القليل جدا من هذه المشروعات يقوم بالتصدير.

4 – حوالى 50% من هذه المشروعات هم الذين لجئوا للتمويل البنكى كما أن خدمات التمويل قد تركزت فى الثلاث محافظات والقطاعات الاثنىن اللآتى تم ذكرهم .

4 – 66% من المشروعات المتوسطة والصغيرة قد أرجعت مشاكلها مع البنوك إلى إرتفاع معدلات الفائدة ورسوم التسجيل .

-هذا وقد تم إنشاء وحدة للمشروعات المتوسطة والصغيرة فى عام 2009 لتقديم الدعم الفنى والخدمات غير التمويلية للبنوك بالإضافة الى تقديم الخدمات التمويلية .

-الصندوق الإجتماعى للتنمية : والذى انشأ فى عام 1991 وهو كيان حكومى يعمل على تخطيط و تطوير وتنسيق وتشجيع توسع المنشآت الصغيرة والمتوسطة ومساعدتهم أيضا فى التمويل والخدمات الأخرى بالتنسيق مع الجهات الحكومية الأخرى ذات الصلة، ومع هذا فإن الصندوق غالبا ما ينصب تركيزه على

المشروعات الصغيرة جدا والمتناهية الصغر ذات الأنشطة التقليدية ، بالإضافة الى الإفتقار إلى عامل الابتكار .

3-2 ثانيا: مشروعات ريادة الأعمال في مصر

في أعقاب ثورة يناير 2011 أصبح هناك إهتمام زائد بين الشباب المصري في مجال ريادة الأعمال ولكنها واجهت العديد من المشكلات العديدة ولعل من أهمها هو أن برامج ريادة الأعمال غير مدرجة في المناهج الدراسية المدرسية باختلاف مراحلها وكذلك في التعليم الجامعي ، هذا بالإضافة إلى مشاكل التمويل كما تم ذكره سابقا ، ومن الجدير بالذكر أن أحد أهم أسباب لجوء الشباب المصري لمشروعات ريادة الأعمال هو رغبتهم الملحة في التوظيف أو بمعنى آخر أن ريادة الأعمال في مصر عادة ما لا يكون الدافع لها هو الأفكار الإبتكارية كما هو متوقع وإنما هو الرغبة في التوظيف والبحث عن مزيدا من الإستقلالية و المرونة في ظروف العمل .وبالفعل فقد وجدت دراسة عن ريادة العمل في مصر ان حوالى ثلث رواد الأعمال قد أقاموا مشروعاتهم الخاصة نظرا لأنهم لم يجدوا عمل بأجر بسبب معدلات البطالة المرتفعة بالإضافة إلى شروط سوق العمل الأمر الذى يدفع الشباب إلى تجربة ريادة الأعمال. بل وأكثر من هذا ،أكثر من نصف هؤلاء الرواد من الذكور قد ذكروا سبب الرغبة في الإستقلالية بنسبة 36% كأحد أسباب القيام بمشروعات ريادة الأعمال ، و9% من الإناث قد أشرن إلى سبب مرونة ساعات العمل والتي توفرها ريادة الأعمال ، وهذا يؤكد على أن إمتلاك العمل الخاص قد يكون له بريق أمام الإناث والذي يمكنهم من الموازنة بين إلتزاماتهم العائلية وبين العمل (Sieverding & Selwaness, 2012)

3-3 الخصائص الخلفية للشباب الذى لديه النية لريادة الأعمال

يوضح الجدول رقم (1) أن 9.9% من العينة لديهم النية للقيام بمشروعات لريادة الأعمال (عددهم 1008)، وذلك بنسبة 72.9% و27.1% من الذكور والإناث بالترتيب، وهذا أمر متوقع نظرا للإستقلالية والحراك الذى يتمتع به الذكور غالبا في مجتمعات الشرق الأوسط وشمال أفريقيا، وأن غالبية هؤلاء الشباب يتركزون في الفئتين العمريتين من 25 حتى 29 (حوالى الثلث) ومن 20 حتى 24 (حوالى الثلث)، وأن غالبية هؤلاء الشباب يقع ضمن الشريحة الأغنى (24%) والمرحلة الأقل (22.7%)، وقد يرجع هذا لإحتياجات تلك المشروعات للجانب التمويلي والذي غالبا ما يتم استيفائه فرديا أو اسريا بسبب صعوبات الحصول على التمويل البنكى، وبالنسبة لمحل الإقامة نجد ان 40% من الإناث الذين لديهم نية ريادة الأعمال هم من قاطنى وجه قبلى ، فى حين تأتى النسبة الأعلى من الذكور هؤلاء القاطنين فى وجه بحرى بنسبة 41.6%، وبالنسبة للتعليم تأتى النسبة الأكبر من الراغبين فى إقامة مشروع لريادة الأعمال من حملة المؤهل الفنى أو المهنى

بنسبة 40.8% (41.7% و 38.3% للذكور والإناث بالترتيب) مما يبين أهمية هذا المسار التعليمي بصفة عامه وبصفة خاصة في الخروج من الأفكار التقليدية للتوظيف .

أما بالنسبة للحالة الوظيفية الحالية لهؤلاء الشباب فنجد أن أكثر من ثلث العينة (35.4%) يعملون بأجر ولديهم النية لإقامة مشروع لريادة الأعمال والنسبة كبيرة بالنسبة للذكور بالمقارنة بالإناث (43.8% في مقابل 12.7%) وهذا الأمر قد يعكس عدم رضاء الذكور عن وظيفتهم الحالية، ولكن نسبة الإناث الغير عاملات وغير باحثات عن عمل نسبة صارخه وهي 49.4% لمن خارج التعليم ، وتحتاج هذه النسبة لمزيد من الدراسات المتعمقة للبحث وراء اسبابها.

4-3 العوامل الاسرية

على الرغم من أهمية العوامل الاسرية كأحد أهم العوامل الغير مباشرة في التأثير على المتغير التابع إلا انها لم يتم تضمينها في النموذج الإحصائي المستخدم وذلك بسبب عدم استكمال بعض بيانات الإستمارات من قبل المستجيبين ، فبالنسبة لسؤال الشباب عن تعليم الأب جدول رقم (2) فإن نصف العينة فقط من أجاب عن هذا السؤال (49.7%) وأشار 72.6% منهم أن الأب غير متعلم الأمر الذي يوضح مدى عمق مشكلة التعليم والامية في مصر وخصوصا مع تسارع معدلات التعليم واختفاء الامية في العالم ويوضح ضرورة إتخاذ إجراءات فورية لعلاج هذه المشكلة ، وبسؤال العينة عن تعليم الامهات كانت الإستجابة أقل بكثير حيث أنها كانت 37.6% ممن أجابوا عن هذا السؤال ، وبالمثل فإن غالبية الامهات غير متعلمة 87.1% ، وبالسؤال عن وظيفة الآباء استجاب 46.6% وكانت النسبة الأكبر للعمل الخاص بنسبة 65%، وبالسؤال عن وظيفة الأم فقد استجاب 3.9% فقط وكانت النسبة الأكبر تعمل في قطاع الأعمال 67.1% .

5-3 العاملون فعليا

ويوضح جدول (3) العاملون فعليا من العينة وهم 75.8% و 24.8% من الذكور والإناث بالترتيب بإجمالي 62.1% ، وبالسؤال عن مدى الإستقرار في العمل نجد أن النسبة الكلية تقريبا نصف العينة (49.6) وتتفوق نسبة الإناث عن نسبة الذكور 69.1 في مقابل 47.2 وهذا يؤكد على أهمية الإستقرار الوظيفي بالنسبة للإناث ، وبسؤال العاملين عن قطاعهم الوظيفي وجد أن النسبة الأعلى هي العمل في القطاع الخاص بنسبة 90.7% (92.2% و 78.4% للذكور والإناث بالترتيب) يليها بفارق كبير القطاع الحكومي بنسبة 6.8% (5.2% و 20.5% للذكور والإناث بالترتيب) مما يوضح انحسار مسار التوظيف التقليدي في القطاع العام وتدل النسبة الأكبر للإناث على تفضيلهم للعمل في القطاع العام نظرا لمزياءه العديدة بخلاف الأجر.

وبسؤال العاملون عن إستخدام الكمبيوتر فى الوظيفة الحالية وجد أن النسبة هى 8.5% (14.4% و7.8% للإناث والذكور بالترتيب) ومن ضمن هؤلاء 64.5% منهم يستخدمون الإنترنت (64.6% و63.8% للذكور والإناث بالترتيب) وقد يرجع هذا إلى ان معظم هؤلاء يعملون بالسكرتارية.

3-6 ريادة الأعمال

ويوضح الجدول رقم (4) بعض الخصائص للشباب الذين قاموا بالفعل بمشاريع لريادة الأعمال ونسبتهم 24.3% (27.8% و14.8% للذكور والإناث بالترتيب) ، وبسؤالهم عن السبب الذى دفعهم لى يقيموا عملهم الخاص بدلا من العمل لحساب الغير كانت النسبة الأكبر من الإجابة هو لمزيد من الإستقلالية 43.4% (45.4% و27.6% للذكور والإناث بالترتيب).

وبسؤال هؤلاء الشباب عن إذا ما كانوا تلقوا أى مساعدة فى نشاطهم الإقتصادى سواء بأجر أو بدون أجر وسواء كان من أفراد العائلة أو لا – سؤال متعدد الإجابة أى أن الشخص يمكن أن يختار أكثر من إجابة واحدة- وجد أن النسبة الأكبر لم تتلق أية مساعدة وإنما قاموا بالعمل كله بمفردهم 48.6% (49.1% و46.2% للذكور والإناث بالترتيب) ، وبسؤالهم عن إذا ما كان المشروع يحقق أرباح فكانت النسبة أكثر من النصف 58.6% (59.6% و53.6% للذكور والإناث بالترتيب) وحوالى الثلث منهم يحقق نقطة التعادل 33.1% (33.7% و30.1% للذكور والإناث بالترتيب) مما يؤكد على ضرورة تشجيع التوسع فى مشاريع ريادة الأعمال ، وبسؤالهم عن معرفتهم بأية جهات تقدم الدعم أو المشورة الغير تمويلية لمشاريعهم وجد أن حوالى نصفهم ليس لديه المعرفة 51.1% (51.7% و49.4% للذكور والإناث بالترتيب). وبسؤالهم عن الإنضمام لأية حملات ترويجية أو تعريفية لريادة الأعمال كان هناك إجماع تقريبا على عدم الإنضمام بنسبة 90% (91.3% و86.4% للذكور والإناث بالترتيب) مما يوضح ضرورة الإهتمام بهذا الجانب المعرفى والترويجى لتحقيق المزيد من إنتشار هذه المشاريع .

وبسؤالهم عن معوقات ريادة الأعمال كانت أعلى نسبة لغياب النصائح المتخصصة أو الإرشادية بنسبة 23.4% تليها الخدمات التسويقية بنسبة 18.6% ، وقد يعود انخفاض نسبة الحاجة إلى القروض إلى معرفة الشباب مسبقا بصعوبة الحصول على قرض ، حتى أنهم بسؤالهم عن قد حصل على قرض وجد انه شخص وحيد من حصل على هذا القرض، وبالسؤال عن مصادر التمويل التى إعتدوا عليها كانت النسبة الأكبر هى المدخرات الخاصة 53% أو المدخرات العائلية 46.6%.

7-3 معامل التوافق

ويوضح جدول (5) درجة العلاقة بين المتغير التابع والمتغيرات المستقلة ، وقد تم حساب معامل التوافق لقياس درجة العلاقة بينهم ، وقد وجد أن جميع المتغيرات المستقلة ذات علاقة معنوية بالمتغير التابع وخصوصا تلك العوامل الخاصة بالعمل الحالى أو الخبرة السابقة للعمل.

8-3 العوامل المؤثرة على النية لريادة الأعمال فى مصر

يوضح جدول رقم (6) المتغيرات التى اقترحتها الدراسة للتأثير على نية ريادة الأعمال، ومن الجدير بالذكر أن هذه العوامل قد تم اختيارها بناء على أسئلة المسح التتبعية للشباب (2014)، وعلى الرغم من أن هذا البحث كان يرغب من إدخال المزيد من العوامل ولكن كان هناك العديد من مشاكل البيانات الناقصة بسبب عدم إستكمال جميع بيانات الإستمارة.

9-3 تحليل نتائج الإنحدار اللوجستى للعوامل المؤثرة على النية لريادة الأعمال

تظهر النتائج فى جدول (7) نتائج تحليل الإنحدار اللوجستى لدراسة نية ريادة الأعمال بين الشباب فى مصر وقد كانت نسبة نجاح النموذج فى تصنيف البيانات بلغت 90% بناء على القيم المتوقعة ، وسوف يتم تفسير النموذج بالإعتماد على نسب الترجيح والتى تم حسابها حسب النموذج ، ونلاحظ الآتى:

- أن احتمال وجود النية تزيد تقريبا بمرتين (2.131) بين هؤلاء الذين ليس لهم تجربة مسبقة للعمل وإن كانت أثناء فترة الدراسة عن هؤلاء الذين سبق لهم العمل (الفئة المرجعية).
- بالنسبة لفئات العمر فإن احتمال وجود النية يزيد بين الشباب فى الفئة العمرية من 15 حتى 19 بنسبة أكبر من الفئتين العمريتين الأخريتين (من 20 حتى 24 ومن 30 حتى 35) عن الفئة المرجعية وهى من 25 حتى 29 ، ولعل هذا يرجع إلى أن هؤلاء الأكبر سنا قد استقروا بالفعل فى وظائفهم بالإضافة إلى أنهم فى الأغلب أصبح لديهم اسرهم الخاصة وبالتالي إمكانية تغيير مسارهم الوظيفى قد تكون مخاطرة كما أنه على الأغلب فإن الأصغر سنا قدرتهم على الحراك تكون أكبر من الأكبر سنا.
- وبالنسبة لمحل الإقامة نجد ان احتمال وجود النية يزيد بأكثر من مرتين ونصف بين قاطنى المحافظات الحضرية (2.632) ، كذلك يزيد الإحتمال بين المقيمين فى محافظات وجه قبلى بمرتين تقريبا (1.929) وأخيرا بين قاطنى وجه بحرى تقريبا بمره ونصف (1.574) عن هؤلاء المقيمين فى المحافظات الحدودية (الفئة المرجعية) ، وقد يرجع هذا لإنخفاض الوعى وعدم وجود حملات ترويجية لريادة الأعمال حتى أن نصف العينة لم يعلموا بوجود هذا النوع من الحملات (51.1%).

- أن الإحتمال بوجود النية يزيد بالنسبة للمستويات الثانية والثالثة من الثروة (1.393 و 1.1.3 بالترتيب) بالنسبة للفئة المرجعية وهى الفئة الأولى (الأقل فى مستويات الثروة) ، فى حين نجد العلاقة عكسية بين المستويين الأعلى من الثروة فى وجود النية لريادة الأعمال عن هؤلاء الموجودين فى أدنى مستويات الثروة (الفئة المرجعية) ، وقد يرجع ذلك الى ان هؤلاء الموجودين فى الشرائح الأغنى قد يكون لديهم تطلعات أعلى فيما يتعلق بالتوظيف أو أن أسرهم بالفعل لديهم مشروعات خاصة بهم وقد يلتحقوا بها مستقبلا.
- وبالنسبة للجنس فإن الذكور يزيد إحتمال وجود النية بينهم عن الإناث (1.479) أى تقريبا بمرة ونصف وبالطبع هذا يشير إلى الاختلافات بسبب النوع الإجتماعى والتي تعطى حراكا واستقلالية للذكور أكثر من الإناث بالإضافة إلى الإلتزامات الأسرية والتي عادة ما يتم حسابها بالنسبة للإناث من قبلهم أو من قبل أسرهم.
- وبالنسبة للحالة الوظيفية الحالية فنجد أن الإحتمال الأعلى لوجود النية لريادة الأعمال يكون بين هؤلاء الذين لا يعملون (عن الفئة المرجعية وهى تضم الغير عاملون ولكنهم يبحثون عن عمل) وهذا أمر بديهي (2.394) حيث أنهم قد يريدون التوظيف بأى شكل من الأشكال وليس رغبة فى القيام بمشروع لريادة الأعمال على وجه الخصوص، ويلى هذا هؤلاء الذين يعملون بأجر (2.254) وقد يعود ذلك إلى عدم رضاؤهم عن عملهم الحالى ، ويلى ذلك هؤلاء العاملون لحساب العائلة بدون أجر وهذا أيضا أمر بديهي (2.155) لرغبتهم فى الحصول على مقابل مادي نظير عملهم ، وأخيرا فإن هؤلاء الذين يعملون لحساب أنفسهم لديهم علاقة عكسية لوجود النية للريادة وقد يرجع هذا إلى أنهم بالفعل لديهم مشروعهم الخاص وهذا أمر مشترك بين ريادة الأعمال وبين العمل لحساب النفس كما تم ذكره سابقا .
- وأخيرا بالنسبة للتعليم فإن الإحتمال يزيد بين كل فئات التعليم (بالنسبة للفئة المرجعية وهم حملة الشهادات الإبتدائى والإعدادى) وخصوصا الجامعى فأكثر والثانوى عام (1.199 و 1.189 بالترتيب)، ومع ذلك فإن الإحتمال كبير أيضا بالنسبة لغير المتعلمين (1.137) وهذا قد يكون دليلا على مشكلة البطالة المنتشرة بين الشباب المصرى عموما والمنتشرة بين صفوفهم سواء متعلمون او غير متعلمون.

4- النتائج

تظهر هذه الدراسة أن انتشار فكرة ريادة الأعمال لاتزال ضعيفة جدا وهذا يظهر من إنخفاض نية الشباب المصرى لريادة الأعمال بنسبة 9.9% فقط ، وهذا يرجع لإنخفاض الوعى والمعرفة بريادة الأعمال سواء من ناحية البرامج الترويجية حيث اظهرت الدراسة ان نسبة (51.1%) لم تسمع عن أية برامج أو

حملات ترويجية لريادة الأعمال ، وأن 97.5% لم يتلقوا أية مساعدة من هذه الجهات ، وأن 90% منهم لم يحضروا أية ورش عمل أو ندوات بهذا الشأن .

تظهر النتائج أيضا إنخفاض الوعى العام بصفة عامة وهذا يظهر من الإستجابة الضعيفة للرد على كثير من أسئلة المسح ، يظهر هذا جليا عند سؤالهم عن خلفية الآباء والأمهات التعليمية والوظيفية . إن المقيمين فى المحافظات الحضرية هم الأكثر احتمالا لان يكون لديهم النية لريادة الأعمال وقد يرجع ذلك لإرتفاع مستوى الوعى بينهم ويليهم هؤلاء قاطنى محافظات وجه قبلى .

كذلك حالة العمل الحالية تعتبر من العوامل التى تزيد من إحتمال وجود النية بين الشباب المصرى على ريادة الأعمال خصوصا هؤلاء الذين لايعملون والعاملون باجر والعاملون لحساب العائلة بدون أجر . إنخفاض النية عند الفتيات مقارنة بالذكور حيث أن وجود النية عند الذكور أكثر بمره ونصف عن الإناث .

5- التوصيات

توصى الدراسة بما يلى:

- 1 – ضرورة زيادة الوعى بين الشباب المصرى لأهمية الدور الذى يمكن أن تلعبه ريادة الأعمال فى مستقبل هؤلاء الشباب ومن ثم مستقبل الإقتصاد المصرى ككل وذلك من خلال إقامة البرامج الترويجية والندوات وورش العمل لتعزيز فكرة ريادة الأعمال.
- 2 – لابد من تخفيض معدلات الأمية التى لاتزال أحد المعوقات الغير مقبولة فى ظل التسارع العالمى لتحسين نوعية الحياة للسكان وتحقيق أهداف التنمية المستدامة . ويمكن أن تصمم ورش عمل بحيث تتضمن محو الامية كهدف ثان .
- 3 – لابد من تضمين مناهج لريادة الأعمال منذ مراحل التعليم المختلفة بداية من التعليم الإبتدائى ومرورا بكافة المراحل التعليمية المختلفة ، ولابد من وجود مسارات تعليمية فى التعليم العالى لتخريج شباب لإقامة مشاريع ريادة الأعمال كما هو الحال فى العديد من دول العالم الآن بخلاف مسار كليات التجارة التقليدية.
- 4 – لابد من دعم الفئات المهمشة فى المجتمع وتحديد الأقل تعليما والأكثر فقرا والمقيمين فى المحافظات الحدودية والإناث .
- 5 – ضرورة توفير الدعم الفنى والمالى من خلال الدولة والمنظمات الأهلية غير الحكومية حيث أن الدعم الفنى والمالى يشكل أحد المعوقات الرئيسية .

6- References:

Adly, A. (April 2014). Reforming the Entrepreneurship Ecosystem in Post-Revolutionary Egypt and Tunisia. Center on Democracy, Development, and the Rule of Law - CDDRL Stanford.

Assaad, Ragui, (2014): Labour Markets in the Arab World. Palgrave Macmillan (ed.) The New Palgrave Dictionary of Economics. (pp. 1-11).

Assaad, R., & Krafft, C. (2015). Is free basic education in Egypt a reality or a myth?. *International Journal of Educational Development*, 45, 16-30.

Assaad, R., & Krafft, C. (2016). Inequality of Opportunity in Human Development. In *The Middle East Economies in Times of Transition* (pp. 131-158). Palgrave Macmillan UK.

Barsoum, G., Ramadan, M., & Mostafa, M. (2014). *Labour market transitions of young women and men in Egypt*. ILO.

Barsoum, G. (2016). 'Job opportunities for the youth': Competing and overlapping discourses on youth unemployment and work informality in Egypt. *Current Sociology*, 64(3), 430-446

Basu, A., & Yulek, M. (2004). Microfinance in Africa: Experience and lessons from selected African countries.

Chigunta, F., Schnurr, J., James-Wilson, D., Torres, V., & Creation, J. (2005). Being "real" about youth entrepreneurship in Eastern and Southern Africa. *SEED working paper*, 72.

Crockett, Teddy. (2014). Rethinking Arab employment: a systemic approach for resource-endowed economies. World Economic Forum, Geneva, Switzerland.

Dobbs, R., Madgavkar, A., Barton, D., Labaye, E., Manyika, J., Roxburgh, C., ... & Madhav, S. (2012). *The world at work: Jobs, pay, and skills for 3.5 billion people* (Vol. 28). Greater Los Angeles: McKinsey Global Institute.

Dobbs, R., & Madgavkar, A. (2014). The world at work: Matching skills and jobs in Asia. *Prospects*, 44(2), 197-210.

Dougherty, C. (2014). *The labour market for youth in Egypt: Evidence from the 2012 school to work transition survey* (Vol. 2). Silatech Working Paper No. 14
El-Said, H., Al-Said, M., & Zaki, C. (2013, May). What determines the access to finance of SMEs? Evidence from the Egyptian case. In *The Economic Research Forum (ERF)*

El-Said, H., Al-Said, M., & Zaki, C. (2014). Small and medium enterprises landscape in Egypt: New facts from a new dataset. *International Journal of Entrepreneurship and Small Business*, 20(3), 286-309.

World Economic Forum(2016). The Human Capital Report 2016.
International Labor Office. The YOUTH Employment Crisis (2012): Highlights of the 2012 ILC report. The youth Employment Crisis: Call for Action.

International Labor Organization (ILO) (2015): Skills Mismatch and Underemployment. How to Boost Employability of Young Women and Men in Egypt? Success Stories & Lessons Learned From ILO Interventions. Decent Work Team for North Africa Skills and Employability Branch.

International Labor Organization (ILO) (2016): Women Entrepreneurship Development Assessment Egypt.

International Labor Organization (ILO) (2017). Towards Evidence- Based Active Labour Market Programmes in Egypt: Challenges and way Forward.

International Labor Organization (ILO) (2017): Youth and Employment in North Africa: A regional Overview. A report Prepared for the Conference on Youth and Employment in North Africa, Geneva.

Khuong, M. N., & An, N. H. (2016). The factors affecting entrepreneurial intention of the students of Vietnam national university—a mediation analysis of perception toward entrepreneurship. *Journal of Economics, Business and Management*, 4(2), 104-111.

Krafft, C., & Assaad, R. (2016). Inequality of Opportunity in the Labor Market for Higher Education Graduates in Egypt and Jordan. In *The Middle East Economies in Times of Transition* (pp. 159-185). Palgrave Macmillan UK.

Morris, R. (2011). 2011 High-Impact Entrepreneurship Global Report. *Endeavor & GEM*, <http://www.gemconsortium.org/docs/download/295>.

Roushdy, R., & Sieverding, M. (2015). Panel Survey of Young People in Egypt (SYPE) 2014: Generating Evidence for Policy and Programs. *New York, NY: Population Council*.

Said, M. (2015). Policies and Interventions on Youth Employment in Egypt. *Torino: ETF*

UNICEF (2012). *The State of the World's Children 2012: Children in an Urban World*. United Nations Publications.

Sieverding, M., & Selwaness, I. (2012). Social Protection in Egypt: A Policy Overview. *Gender and Work in the MENA Region Working Paper Series*, (23)

Talaş, E., Çelik, A. K., & Oral, I. O. (2013). The influence of demographic factors on entrepreneurial intention among undergraduate students as a career choice: the case of a Turkish University. *American International Journal of Contemporary Research*, 3(12), 22-31.

The Swedish Trade & Invest Council (2015). Egypt Entrepreneurship Study. Business Sweden Dubai

World Economic Forum (2017): The Future of Jobs and Skills in the Middle East and North Africa, Preparing the Region for the Forth Industrial Revolution.

World Bank (2017). Retrieved from: data.worldbank.org.

نانسى بلوسكى ويجر، ديفيد وتشابمان ، ارين باكستر ، وجوان ديجيغير (2015) . أهمية السياق: نموذج للعوامل المرتبطة بفاعلية تدريب الشباب على أعمال الريادة . مجلة مستقبلات . العدد 176 . اليونسكو . جنيف . سويسرا .

الملاحق: (جميع الجداول تم إعدادها بواسطة الباحثه)

جدول رقم (1) خصائص الشباب الذي لديه نية لريادة الأعمال

السؤال	الإجابة	النسبة		
		ذكور	اناث	الاجمالي
هل لديك النية لريادة الأعمال	نعم	14.1	5.6	9.9
	لا	85.9	94.4	90.1
النوع	ذكر			72.9
	انثى			27.1
الفئة العمرية	15-19	16.6	18.1	17.0
	20-24	31.7	24.6	29.8
	25-29	30.3	30.8	30.4
	30-35	21.4	26.5	22.8
مؤشر الثروة	الأدنى	18.9	19.4	19.0
	المستوى الثاني	14.7	18	15.6
	المستوى الأوسط	19.3	17.1	18.7
	المستوى الرابع	24.1	18.9	22.7
محل الإقامة	المستوى الأغنى	23	26.6	24.0
	المحافظات الحضرية	23.6	32.2	18.7
	وجه بحرى	41.6	27.4	42.8
	وجه قبلى	33.4	40.0	36.7
مستوى التعليم	محافظات الحدود	1.4	0.4	1.8
	غير متعلم	7.3	12	8.5
	ابتدائي + اعدادي	23.8	18	22.2
	ثانوي عام +معهد سنتين	7.4	7.7	7.5
	ثانوي فني	41.7	38.3	40.8
الحالة الوظيفية الحالية	جامعي فأكثر	19.8	24.1	21.0
	العمل بأجر	43.8	12.7	35.4
	عمل حر	28.4	8.8	23.1
	عمل بدون أجر للعائلة	3.6	3.3	3.6
	لايعمل ويبحث عن عمل	6.9	8.4	7.2
	لايعمل ويدرس	9.4	17.4	11.5
الحالة الوظيفية	لايعمل ولا يدرس	7.9	49.4	19.2
	لا	24.2	75.2	37.9
	نعم	75.8	24.8	62.1

جدول رقم (2) بعض الخصائص العائلية

اجمالى النسب	النسبة	الإجابة	السؤال
49.7%	72.6	غير متعلم	أعلى مستوى تعليمى أكمله الأب
	13.9	ابتدائي + اعدادي	
	9.0	ثانوي عام +معهد سنتين	
	4.5	جامعي فأكثر	
37.6%	87.1	غير متعلم	أعلى مستوى تعليمى أكملته الأم
	4.9	ابتدائي + اعدادي	
	5.8	ثانوي عام +معهد سنتين	
	2.2	جامعي فأكثر	
46.6%	31.4	قطاع حكومى	قطاع عمل الوالد
	2.3	قطاع أعمال عام	
	65.0	قطاع خاص	
	1.3	اخرى	
3.9%	32.9	قطاع حكومى	قطاع عمل الوالده
	67.1	قطاع أعمال عام	
	-	قطاع خاص	
	-	اخرى	

جدول رقم (3) بعض خصائص العاملون فعلا

النسبة			الإجابية	الحالة العملية
الاجمالي	اناث	ذكور		
49.6	69.1	47.2	عمل منتظم	مدى إستقرارك في العمل
24.4	22.5	24.7	عمل مؤقت	
4.4	2.5	4.6	عمل موسمي	
21.6	5.9	23.5	عمل متقطع	
6.8	20.5	5.2	قطاع حكومي	ما هو القطاع الذي يتبعه عمالك
1.1	1.1	1.1	قطاع أعمال عام	
90.7	78.4	92.2	قطاع خاص	
1.4	-	1.5	قطاع استثماري	
34.5	23.5	35.9	نعم	هل يتطلب عمالك أية مهارات خاصة
65.5	76.5	64.1	لا	
8.5	14.4	7.8	نعم	هل تستخدم الكمبيوتر في عمالك
91.5	85.6	92.2	لا	
64.5	63.8	64.6	نعم	هل تستخدم الإنترنت
35.5	36.2	35.4	لا	

جدول رقم (4) بعض خصائص الشباب الذى لديهم بالفعل مشاريع ريادة أعمال

الاجمالي	اناث	ذكور	الإجابة	السؤال
24.3	14.8	27.8	نعم	انت فعلا عملت المشروع اللي فكرت فيه
75.7	85.2	72.2	لا	
33.9	33.6	33.9	لم أجد عملا بأجر	لماذا اخترت ريادة الأعمال بدلا من العمل لحساب شخص آخر
42.4	27.6	45.4	قدرا أكبر من الإستقلالية	
4.0	8.7	3.0	مرونة أكثر فى ساعات العمل	
11.9	17.3	10.8	مستوى أعلى للدخل	
5.1	6.0	4.9	لدى أهلى تجربة ناجحة فى العمل الحر	
2.7	6.6	1.9	اخرى	
10.7	3.6	12.1	أفراد من الاسره يعملون بأجر	هل ساعدك أحد فى نشاطك الإقتصادى
20.4	5.7	23.3	موظفون بأجر	
6.2	13.5	4.9	أفراد من الاسره يعملون بدون أجر	
8.0	9.1	7.7	شركاء يعملون بدون أجر	
48.6	46.2	49.1	لا يوجد أحد للمساعدة	
8.3	16.3	6.7	يحقق خسارة	هل مشروعك يحقق مكسب أم يخسر
58.6	53.6	59.6	يحقق مكسب	
33.1	30.1	33.7	نقطة التعادل	
48.9	50.6	48.3	نعم	هل لديك معرفة بمؤسسات تقدم خدمات غير تمويلية
51.1	49.4	51.7	لا	
2.5	2	2.7	نعم	هل تلقيت منهم أى مساعدة أو نقل للخبرات الفنية أو للمهارات المطلوبة للمشروع
97.5	98	97.3	لا	
10.0	13.6	8.7	نعم	هل تم دعوتك لحضور أى ندوات أو ورش عمل للتوعية بريادة الأعمال
90.0	86.4	91.3	لا	

تابع جدول (4):

السؤال	الإجابة	%
معلومات عن النشاط التجارى	معلومات عن النشاط التجارى	12.5
	خدمات تسويقية	18.6
	خدمات تمويلية	12.6
	خدمات محاسبية	1.7
	خدمات قانونية	2.4
	نصائح استرشادية	23.4
	تدريب فى مجال الأعمال	1.9
	تدريب للمهارات	3.3
	أدخال التكنولوجيا	1.5
	تطوير المنتج	4.4
	الحصول على تمويل	0.4
	إدخال المياه والكهرباء	1.2
	اخرى	16
معلومات ريادة الاعمال	مدخرات شخصية	53.0%
	مدخرات عائلية	46.6%
	قرض من العائلة أو الأصدقاء	16.5%
	قرض من البنك	0.4%
	قرض من شخص مقرض	0%
	قرض من مؤسسة حكومية	0%
	الصندوق الإجتماعى للتنمية	0%
	قرض من مؤسسة غير حكومية	0%
	لم يحتاج إلى إقتراض	1.3%
	اخرى	0.9%
ماهى مصادر تمويل مشروعك	مدخرات شخصية	53.0%
	مدخرات عائلية	46.6%
	قرض من العائلة أو الأصدقاء	16.5%
	قرض من البنك	0.4%
	قرض من شخص مقرض	0%
	قرض من مؤسسة حكومية	0%
	الصندوق الإجتماعى للتنمية	0%
	قرض من مؤسسة غير حكومية	0%
	لم يحتاج إلى إقتراض	1.3%
	اخرى	0.9%

جدول رقم (5) درجة العلاقة بين النية فى ريادة الأعمال والمتغيرات التابعة

المتغير	معامل التوافق	درجة المعنوية
الحالة الوظيفية الحالية	0.284	0.001
تجربة العمل السابقة حتى وإن كان أثناء الدراسة	0.198	0.001
النوع	0.141	0.001
الفئة العمرية	0.10	0.001
محل الإقامة	0.065	0.001
مستوى التعليم	0.045	0.001
مؤشر الثروة	0.036	0.001

جدول رقم (6) المتغيرات المستخدمة فى النموذج والتي تفسر نية الشباب للقيام بزيادة الأعمال

المتغير	الفئات
الجنس	ذكر انثى (الفئة المرجعية)
مستوى التعليم	غير متعلم ابتدائى + إعدادى (الفئة المرجعية) ثانوى عام + معهد سنتين ثانوى فنى جامعى فأكثر
مؤشر الثروة	المستوى الأول (الفئة المرجعية) المستوى الثانى المستوى الثالث المستوى الرابع المستوى الخامس (الأغنى)
فئات العمر	19-15 24-20 29-25 (الفئة المرجعية) 35-30
محل الإقامة	المحافظات الحضرية محافظات وجه بحرى محافظات وجه قبلى محافظات الحدود (الفئة المرجعية)
خبرة العمل السابقة حتى أثناء فترة الدراسة	نعم (الفئة المرجعية) لا
الحالة الوظيفية الحالية	العمل بأجر العمل الحر أو العمل لحسابه العمل بدون أجر للعائلة لا يعمل ويبحث عن عمل (الفئة المرجعية) لا يعمل

جدول رقم (7) نتائج الإنحدار اللوجستي لنية ريادة الأعمال بين الشباب المصري

prob.	95% C.I.for EXP(B)		Exp(B)	Sig.	S.E.	B	
	Upper	Lower					
68.1%	2.142	2.121	2.131	0.000	.002	.757	تجربة العمل السابق
				0.000			فئات العمر
57.3%	1.347	1.336	1.342	0.000	.002	.294	15-19
50.8%	1.036	1.029	1.032	0.000	.002	.032	20-24
51.2%	1.055	1.047	1.051	0.000	.002	.050	30-35
				0.000			محل الإقامة
72.5%	2.664	2.601	2.632	0.000	.006	.968	المحافظات الحضرية
61.1%	1.592	1.555	1.574	0.000	.006	.453	محافظات وجه بحرى
65.9%	1.952	1.907	1.929	0.000	.006	.657	محافظات وجه قبلى
				0.000			مؤشر الثروة
58.2%	1.399	1.387	1.393	0.000	.002	.332	المستوى الثانى
56.5%	1.305	1.294	1.300	0.000	.002	.262	المستوى الثالث
49.7%	.993	.985	.989	.000	.002	-.011	المستوى الرابع
49.3%	.977	.968	.973	.000	.002	-.028	المستوى الخامس(الأغنى)
59.7%	1.484	1.474	1.479	0.000	.002	.391	الجنس (ذكور)
				0.000			حالة العمل الحالى
69.3%	2.268	2.241	2.254	0.000	.003	.813	العمل بأجر
29.5%	.421	.416	.418	0.000	.003	-.871	العمل لحساب النفس
68.3%	2.174	2.137	2.155	0.000	.004	.768	العمل لحساب العائلة بدون أجر
70.5%	2.407	2.381	2.394	0.000	.003	.873	لا يعمل

تابع جدول رقم (7) نتائج الإنحدار اللوجستى لنية ريادة الأعمال بين الشباب المصرى

				0.000			المستوى التعليمى
53.2%	1.143	1.131	1.137	0.000	.003	.128	غير متعلم
54.3%	1.195	1.183	1.189	0.000	.003	.173	ثانوي عام +معهد سنئين
51.0%	1.043	1.036	1.039	.000	.002	.038	ثانوي فني
54.5%	1.204	1.194	1.199	0.000	.002	.182	جامعي فأكثر
			1.576	0.000	.004	.455	Constant
						57158.344	Model Chi-Square
						17569664.060 ^a	-2 Log Likelihood
						90%	Overall Percentage

ملحوظة:

جميع الجداول تم إعدادها بواسطة الباحثة.

دوافع المهاجرين الحاليين والعائدين فى مصر

محمد مصطفى حسن¹، ومحمد عبدالمنعم عبدالرؤف²، أيمن زهري³

المقدمة

تمثل الهجرة الدولية أحد أهم القضايا المثارة حالياً والتي تحتل صدارة الاهتمامات الوطنية والدولية خاصة فى ظل التوجه العالمى نحو تحرير قيود التجارة والعولمة والتي يتوقع نتيجة لها فتح الحدود وتخفيف القيود على السلع وحركة رؤوس الأموال. فالهجرة الدولية تعتبر ظاهرة ديموجرافية واجتماعية واقتصادية تتأثر بشكل مباشر بعوامل داخلية وخارجية كديناميكيات سوق العمل المحلية والدولية، كما تتأثر بالظروف السياسية سواء للدول المرسله أو المستقبلة.

وطبقاً لأحدث تقدير للمصريين المقيمين والعاملين بالخارج من واقع بيانات تعداد 2006، والبيانات الواردة من وزارة الخارجية ووزارة القوى العاملة والهجرة بلغ عددهم 3.9 مليون مواطن هجرة مؤقتة (عبدالكريم وآخرون، 2009). وبالتالي لم تعد الهجرة ظاهرة عرضية فى المجتمع المصرى بعد أن تغلغت وأصبحت مشروعاً قومياً مستمراً مما يستحق معه وضع إستراتيجية قومية للهجرة لا يمكن صياغتها وتنفيذها إلا كجزء من إستراتيجية قومية أعم وأشمل لتخطيط الموارد البشرية وتنميتها وإستخدامها الإستخدام الأمثل وذلك ما يستلزم توافر بيانات دقيقة وموثقة عن الظاهرة مما يشمل دوافع المغادرة والعودة. وادراكاً لهذه الأهمية قام الجهاز المركزى للتعبئة العامة والاحصاء بالتعاون مع الوزارات المعنية بالهجرة بتنفيذ مسح الهجرة الدولية 2013 من خلال التعاون مع المفوضية الاوروبية، وزارة التخطيط والتعاون الدولي، صندوق الأمم المتحدة للسكان، منظمة الصحة العالمية، المنظمة الدولية للهجرة، المفوضية العليا للاجئين، ويعتبر هذا المسح ثروة معلوماتية للباحثين والدارسين للتعرف على احوال الهجرة الخارجية وخصائصها بصفة عامة ودوافع المهاجرين الحاليين ومقارنتها بدوافع المهاجرين العائدين وهذه الدراسة تعتبر محاولة من الباحث فى التعرف على تلك النقطة

¹ أستاذ مساعد - معهد الإحصاء جامعة القاهرة

² إحصائي بالجهاز المركزى للتعبئة العامة والإحصاء، وطالب ماجستير بمعهد الدراسات والبحوث الإحصائية

³ رئيس الجمعية المصرية لدراسات الهجرة .

1. مشكلة الدراسة

تتمثل مشكلة الدراسة فى عدم معرفة الاختلافات النسبية والمعنوية بين دوافع المهاجرين الحاليين والعائدين من الهجرة ، بمعنى هناك اختلاف بين دوافع المهاجرين العائدين الى مصر قبل 1/1/ 2000 عن دوافع المهاجرين الحاليين حتى إجراء البحث عام 2013 ويمكن صياغة تلك المشكلة فى السؤال التالى :

- هل هناك فرق معنوى بين دوافع المهاجرين الحاليين والعائدين من الهجرة طبقاً لخصائصهم .

2. الدراسات السابقة

هناك الكثير من الدراسات الدولية التي تناولت دوافع الهجرة والعودة، كما تناولت في مصر دراسات أخرى لنفس القضية إلا أن بعضها اعتمد على بيانات ذات مستوى تجميعي توفرها أجهزة الاحصاء بشكل لا يسمح بكثير من التفصيل والتعمق في دراسة خصائص المهاجرين وتأثيرها على دوافع الهجرة أو العودة.

- دراسة (السعدني، سمية، 1993) "دوال للهجرة الدولية للعمل فى حالة مصر": طورت

الدراسة نموذجاً نظرياً لشرح عملية اتخاذ القرار بالهجرة على مستوى الفرد في وقت اهتمت فيه اغلب الدراسات الخاصة بالهجرة بتقدير اثارها على مستوى المجتمع المصري والافراد. وقامت الدراسة بتقدير المحددات الديموجرافية والاجتماعية لاتخاذ القرار بالهجرة من بين القوى العاملة في مصر خلال الفترة من 1974 إلى 1984. واستخلصت الدراسة مجموعة من النتائج منها أن احتمال هجرة الافراد يتناقص مع ارتفاع الدخل المتوقع الحصول عليه في مصر، إلا أن الدراسة استخلصت أيضاً ارتفاع احتمال هجرة الافراد يزيد بالرغم من تناقص الدخل المتوقع الحصول عليه وارجعت ذلك الى الدخل المتوقع هو نتاج لتفاعل عاملى الطلب والعرض ونتيجة لتناقص الطلب في فترة الدراسة على العمالة فأن الدخل المتوقع بالتالي منخفض. كما وجدت الدراسة أن الذكور أكثر ميلاً من الاناث إلى الهجرة الدولية، وتناقص احتمال الهجرة مع الاعمار الكبيرة، وترتفع احتمالات الهجرة بين الفئات العاملة في المهن العلمية والفنية والانتاج والنقل أكبر منه مقارنة بالفئات الأخرى.

– دراسة (إمام، مجدة، 2009) "هجرة الشباب الدولية والتنمية": ناقشت الدراسة العلاقة الجدلية

بين هجرة الشباب والتنمية، وملامح هذه الهجرة ومحدداتها لدى الشباب العربي، وقياس حجمها الحقيقي وعوامل الدفع والذب، كذلك العوامل والآليات الدافعة والجاذبة للهجرة الدولية والرؤى والتوقعات المستقبلية المختلفة لهجرة الشباب والتنمية. وهدفت من خلال مناقشة علاقة هجرة الشباب بالتنمية والتركيز على البعد التنموي للهجرة وبالتالي فهم علاقتها بالتنمية، وتوضيح الفرص والتحديات بما تتضمنه من آثار إيجابية وسلبية على المهاجر ومجتمعه ركزت المعالجة على دول مصر، الأردن، لبنان، سوريا، فلسطين، واعتمدت الدراسة على البيانات الإحصائية التجمعية على المستوى القطري لدول الدراسة المتاحة، وذلك من واقع قواعد البيانات الوطنية والدولية. خلصت الدراسة إلى أن الهجرة ظاهرة أكثر تنوعاً وتعقيداً وبالتالي ينبغي تناولها بطريقة شاملة ونهج ذو أبعاد شاملة على الصعيد الوطني والإقليمي والدولي.

– دراسة (زهري وأخرون، 2009) "إتجاهات الشباب المصري حول الهجرة لأوروبا": نتيجة

للتعاون الدولي قام قطاع الهجرة بوزارة القوى العاملة ممولا من قبل الحكومة الايطالية تم تنفيذ مسح ميداني للتعرف على خصائص المصريين المهاجرين هجرة غير منتظمة وقياس مدى وعيهم بحقائق ومخاطر الهجرة غير المنتظمة بما يشمل الإتجار بالمهاجرين، وذلك للمساعدة في إعداد وسائل إعلامية عن طريق التعاون بين الجهات الحكومية وغير الحكومية والمؤسسات الإعلامية المختلفة للوصول إلى الفئات المستهدفة والتأثير على ما لديها من المفاهيم الخاصة بالهجرة. ولتحقيق ذلك تم إجراء مسح ميداني في ثمانية محافظات ، القاهرة، والاسكندرية، والمنوفية، والغربية، والشرقية، والدقهلية، والفيوم، والاقصر وتم من الدراسة الميدانية مقابلة 1552 شابا فى الفئة العمرية من 18 الى 40 عاماً كما حددتها الدراسة . وكانت أهم نتائج الدراسة أن غالبية الشباب الذين يرغبون في الهجرة الى أوروبا ينوون العودة والاستقرار في مصر بعد هجرتهم الى اوروبا وان الدافع الاساسي لهجرة هؤلاء الشباب هو البطالة و إنخفاض الاجور في مصر ومحاكاة النماذج الناجحة . كما استخلصت الدراسة ان طابع الهجرة المصرية فردي ذكوري بعكس النمط العائلي للهجرة بمواطني افريقيا جنوب الصحراء.

- (زهري ايمن 2009) "اتجاهات هجرة المصريين فى ايطاليا وفرنسا": محاولة لفهم اتجاهات هجرة المصريين لاوروبا اختارت الورقة كلا من ايطاليا وفرنسا كنموذج لاستقراء نمط الهجرة المصرية، واعتمد الباحث فى دراسته على عينة غير احتمالية "كرة الثلج". وقد لاحظ الباحث ان تقديرات الهجرة فى مصر تفوق التقديرات المبلغة من الدول المستقبلية لهذه الهجرة ولعل واحدا من هذه الاسباب الهجرة غير الشرعية لبعض المصريين لدول اوروبا الغربية.
- (البرعي وأخرون، 2013) "الهجرة الدولية للشباب": هدفت الدراسة إلى التعرف على خصائص الشباب الراغب فى الهجرة، أسباب طموحاتهم إلى الهجرة، بلدان المقصد التى يطمح الشباب فى الهجرة إليها، رأى الشباب فى كل من الهجرة الشرعية وغير الشرعية، مصادر معلومات الطامحين إلى الهجرة حول بلدان المقصد وتجربة الشباب فى الهجرة الدولية. اعتمدت الدراسة فى اجابة الاسئلة السابقة على منهج التحليل الوصفى لعرض السمات البارزة لعينة من الشباب الطامحين للهجرة أو السابق لهم الهجرة فى الفئة العمرية (18-29) سنة، وذلك باستخدام بيانات مسح النشء والشباب فى مصر عام ٢٠٠٩ والذي أجراه مجلس السكان الدولى بالتعاون مع مركز المعلومات ودعم اتخاذ القرار التابع لمجلس الوزراء المصرى. واشتمل المسح على تحليل شامل لحالة الشباب بعينة ممثلة على المستوى القومى بلغ حجمها ١٥٠٢٩ شاب من الجنسين فى الفئة العمرية (18-29) سنة من مجموع ١١٣٧٢ أسرة معيشية. وخلصت الدراسة أن الاسباب الدافعة للهجرة لدى الشباب تأتى لعدم وجود فرصة عمل، قلة مستوى الدخل وسوء الاحوال المعيشية؛ فى حين أن الاسباب الجاذبة للهجرة إلى الخارج هي زيادة الدخل والاقامة بالخارج وهي نتائج تشير إلى توافق بين دوافع الخروج من مصر ورغبة الذهاب إلى الخارج. كما وجدت الدراسة أن اغلبية الشباب يطمح إلى الهجرة إلى البلدان العربية فى المقام الأول ثم أوروبا فالولايات المتحدة. وتم استيفاء الاسئلة الخاصة بدوافع الهجرة والعودة لمجموعة الشباب الذين سبق لهم الهجرة من عينة المسح.
- دراسة (متولى، سها، 2014) "تأثير الهجرة الداخلية على الفقر فى مصر": مثلت محاولة لاثراء المكتبة المصرية بدراسات لتفسير دوافع ومحددات الهجرة وانعكاساتها على المستويات

المختلفة. واعتمدت الدراسة على افتراض تمثيل الهجرة الداخلية أحد السبل المهمة أمام الافراد لرفع مستوى معيشتهم والخروج من دائرة الفقر. واعتمدت الدراسة على بيانات تتبعية لدراسة الظاهرة من مسح خصائص سوق العمل المصري.

– دراسة (حسن وآخرون 2016) "الهجرة الدولية وعلاقات الفقر والمكان في مصر: أدلة تطبيقية من مسح الدخل والانفاق والاستهلاك": اختبرت دراسة فرضية الارتباط بين الهجرة الدولية والفقر من خلال اختبار فرضية أن الاسر الفقيرة يكون أفرادها أكثر عرضة للهجرة الدولية مقارنة بالاسر غير الفقيرة. واستخدمت الدراسة 50% من بيانات مسح الدخل والانفاق والاستهلاك 2013/2012، وأظهرت النتائج النهائية وجود علاقة طردية بين نسب الفقر والهجرة الدولية على مستوى الجمهورية اجمالاً دون أن يكون ذلك نمط عام على كافة محافظات الجمهورية (27 محافظة) حيث لم تحصل الدراسة على دلائل على مستوى كل محافظة على حدة.

وبالإضافة إلى الدراسات المصرية فهناك العديد من الدراسات التي أجرت بالخارج لدراسة وتحليل الهجرة الدولية ودوافعها مثل:

– دراسة (Velizarova, M 2009) "دوافع الهجرة، العودة واستمرارية العودة: دراسة حالة الطلاب البلغار": ذهبت الدراسة في استقصاء دوافع الهجرة والعودة للمهاجرين البلغاريين ونتيجة لغياب بيانات تسمح بتحليل ومعالجة الظاهرة اعتمدت على اسلوب العينة غير الاحتمالية "كرة الثلج" في جمع البيانات الكترونياً وبلغ حجم العينة 70 مفردة اعتمدت عليهم الدراسة في التعرف على دوافع الهجرة والعودة. ووجدت الدراسة أن رغبة العائدين إلى الهجرة مرة أخرى، وأن أسباب عودة الغالبية منهم مرة أخرى للغالبية يتلخص في العائلة. واقترح Velizarova للتغلب على مشكلة البيانات التي واجهته في دراسته استعمال سجل لكل بلغاري يعيش في الخارج لفترة اطول من السنة، وينفذ السجل المقترح بمساعدة البلدان المعروفة بالإتجاهات المفضلة للمهاجرين البلغاريين. كما اقترح حلاً آخر يتمثل في اضافة اسئلة خاصة بالهجرة في تعداد السكان بشكل دقيق في بلغاريا.

– دراسة (Davids و Houte, 2014) "رجوع أو تقدم للأمام؟ هجرة عودة، التطوير وبناء

السلام": طبقاً لخصائص المهاجرين العائدين من أوروبا إلى بلدانهم الأصلية التي واجهت صراعات في الماضي، ومدى استعدادهم وقدرتهم على المساهمة في التنمية وبناء السلام ببلدانهم الأصلية. واعتمدت الدراسة على المنهج المقارن في ست دول على مستوى العالم ما بين عامي 2007 و2008، ودراسة حالة لأفغانستان في 2012. وخلصت الدراسة إلى عدم تجانس تجربة ما بعد العودة، ومعاني معقدة لدوافع العودة من الهجرة، واعتماد نجاح جهود عودة المهاجرين في بناء التنمية والسلام يتوقف على خصائص المهاجرين، وأخيراً استكشاف إمكانية عودة المهاجرين. ورأت الدراسة من أجل وضع سياسات ملائمة ضرورة تلبية احتياجات وإمكانات العائدين .

– دراسة" (Masja et al., 2014) "فهم التجارب المختلفة لما بعد العودة: دور الاستعداد،

دوافع العودة والتوقعات العائلية للمهاجرين العائدين في المغرب": استهدفت الدراسة فهم تجارب ما بعد العودة المختلفة في المغرب وأهمية فهم هذا الموضوع في سياق من التجارب ما بعد العودة. واعتمدت الدراسة على نتائج 44 مقابلة نوعية (كيفية) مع مجموعة متنوعة من المهاجرين العائدين. وأظهرت نتائج الدراسة أن العديد من المهاجرين يشاركون في دعم أقاربهم مع وجود مجموعة منهم في وضع لا يسمح لها بذلك ويتوقف هذا على البلد المهاجر إليها والاجور المقدمة، فالمهاجرين إلى البرتغال من أقل المهاجرين قدرة على تحقيق توقعات الأسرة نظراً لانخفاض الأجور هناك على سبيل المثال . وبالتالي فإن الدراسة أكدت على أن الظروف في بلد المقصد مهمة لفهم مختلف التجارب بعد العودة. وبالتالي فإن نتائج الدراسة ساهمن في مزيد من الفهم للتجارب المختلفة للعائدين في المغرب، وبالرغم من ذلك ونتيجة لصغر حجم العينة فإن الدراسة لم تكن قادرة على دراسة منهجية كيفية تأثير خصائص المهاجر مثل الجنس والعمر والمستوى التعليمي على نتائج أو تتقاطع مع عوامل أخرى لتشكيل نتائج في مسار تجربة العودة من الهجرة الدولية.

نقد الدراسات السابقة

على الرغم من وجود دراسات تناولت الهجرة الدولية فى مصر الا ان الكثير من تلك الدراسات اعتمدت فى بياناتها على بيانات تجميعية على المستوى المحلى ، الاقليمى او الدولى وان كانت بعض الابحاث قد استخدمت بيانات تحصل عليها من عينات غير احتمالية الا ان بيانات مسح الهجرة الدولية عام 2013 يوفر بيانات على مستوى الاسرة المعيشية والفرد المهاجر حالياً والعائد من الهجرة لعينة مماثلة بجمهورية مصر العربية وهذا ما قد يعطى لهذه الدراسة من موثوقية عالية للنتائج الاحصائية التى تتحصل عليها هذه الدراسة

3. أهداف الدراسة

تهدف الدراسة الى التعرف على الفروق والاختلافات النسبية والمعنوية بين دوافع المهاجرين الحاليين والعائدين من الهجرة طبقاً لخصائصهم فى مصر .

4. اطار الدراسة

يعتبر الاطار النظرى مفيداً فى فهم وتفسير دوافع الهجرة والعودة منها ومحدداتها ، ونظراً لأهمية تلك المحددات وتفاعلاتها ولما لها من اثر فى قراءات الهجرة والعودة منها ، فأنة يتعين على الدراسة ان ترصد تلك المحددات ودرجة تأثيرها مما يغدو من الضرورى وضعها امام صانعى السياسات الخاصة بالهجرة ، ليس بهدف التعظيم الايجابى منها وتدنية تأثيراتها السلبية فحسب بل تمتد الى وضع الخطط والاستراتيجيات الخاصة بالهجرة ، وعلى ذلك تم تقسيم الاطار النظرى إلى مجموعتين رئيسيتين نوضحهما على النحو التالى كما هو مبين بالشكل رقم (1) .

المجموعة الاولى :

وتمثلها دوافع المهاجرين الحاليين والعائدين (كمتغير تابع) وقد تم تقسيم تلك الدوافع الى اربع مجموعات نوضحهما كالاتي :

1- دوافع ظروف العمل فى مصر :-

وتضمن مجموعة من المتغيرات مثل - متعطل - هاجر بقرار من صاحب العمل - فرص العمل جيدة بدولة المهجر وغيرها

2- دوافع اقتصادية :- وتتضمن متغيرى - تحسين مستوى المعيشة - عدم كفاية الدخل

3- دوافع اجتماعية : وتتضمن - جمع شمل الاسرة (مرافقة) - للزواج - للهروب من المشاكل الاسرية وغيرها

4- دوافع اخرى : - تتضمن - ضعف الامن فى مصر - امكانية الحصول على اللجوء... وغيرها
المجموعة الثانية :

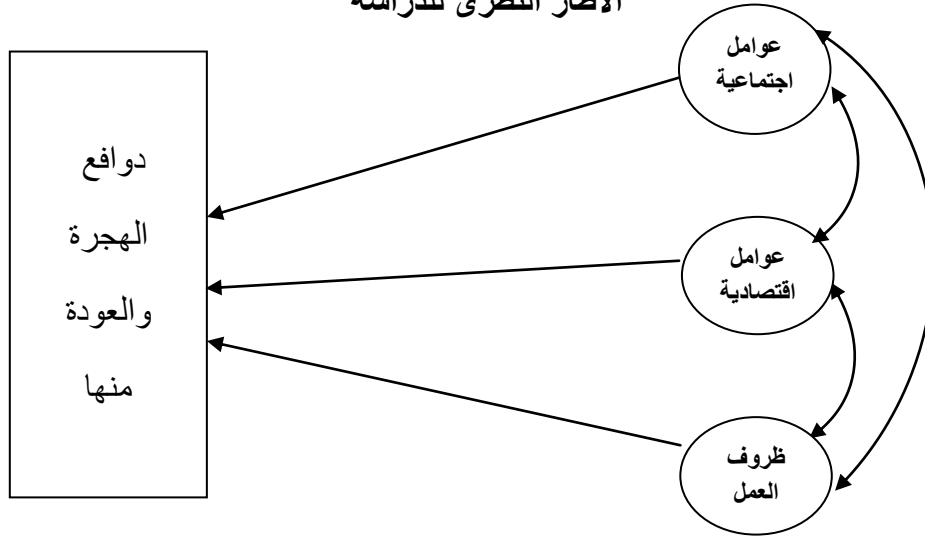
وتتمثل فى عاملين ممثلان الخصائص الخلفية المهاجرين الحاليين والعائدين والتي تؤثر (كمتغير مستقل) فى دوافع الهجرة والعودة منها وتتمثل تلك العوامل فى :

1- عوامل اجتماعية :- تمثل فى المتغيرات التالية : - النوع - العمر (الحالى - عند الهجرة (- مدة الهجرة - الحالة التعليمية عند الهجرة - صاحب قرار الهجرة - محل الإقامة .

2- عوامل اقتصادية وحالة العمل :- التمويل المالى للهجرة - نوع العمل قبل الهجرة - سابقة العمل قبل الهجرة ... وغيرها

شكل رقم (1)

الاطار النظرى للدراسة



مصدر الاطار : بمعرفة الباحث

5. فروض الدراسة

اعتمدت الدراسة على مجموعة من الفروض نبينها على النحو التالى :-

1- ان دوافع الهجرة والعودة منها تعتبر (متغير تابع) وهى دالة فى مجموعة من (المتغيرات المستقلة) والمتمثلة فى محددات ومتغيرات العوامل الاجتماعية والاقتصادية وظروف العمل بمصر .

2- تقوم الدراسة على فرض انه ليس هناك فرق معنوى بين دوافع المهاجرين الحاليين والعائدين من الهجرة تبعاً لخصائصهم .

6. أهمية الدراسة

تتركز الاهمية التطبيقية هذه الدراسة فى ، ان نتائجها سوف تكون عوناً للمخططين ومنفذى البرامج ومتخذى القرارات فى شأن الهجرة والمهاجرين وتمكنهم من التعرف على دوافع الهجرة لدى المصريين واتجاهاتها .

7. منهجية الدراسة

اعتمدت الدراسة على الاسلوب المكتبى فى توفير البيانات بغرض التعرف على المفاهيم والتعاريف النظرية لظاهرة الهجرة بالإضافة الى الابحاث والمسوح التى تناولت هذا الموضوع والمؤشرات التى تم التوصل اليها والتى نبينها على النحو التالى :

أ- التعداد العام للسكان والاسكان والمنشآت للأعوام 1986 ، 1996 ، 2006.

ب- مسح الهجرة الخارجية اعوام 1984 ، 1987 بالإضافة الى البحث التتبعى للهجرة الخارجية واستراتيجيتها عام 1997 .

ج- مسح اسباب الهجرة الدولية عام 1997 ، هذا بالإضافة الى مجموعة من المسوح الاسرية التى

تضمنت اسئلة عن الهجرة الدولية فى مصر اعوام 1996 ، 2006 ، وكذلك مسح الدخل

والانفاق والاستهلاك عن تلك الاعوام

وقد استخدمت الاساليب الاحصائية التالية فى تحليل البيانات

1-7 التحليل الإحصائي الوصفى

2-7 التحليل الإحصائي المتقدم متمثلا في

- تحليل التباين باستخدام مربع كا2 (x2) لقياس الفروق بين خصائص المهاجرين ودوافعهم من حيث الهجرة من مصر والعودة إليها .
- اختبار الفروض .

8. مصادر البيانات

اعتمدت الدراسة على البيانات الخام " لمسح الهجرة الدولية 2013 " . الذى قام به الجهاز المركزى للتعبئة العامة والإحصاء بتنفيذ مسح الهجرة الدولية عام 2013 (EGYPT – HIMS) كجزء من برنامج مسح الهجرة الدولية فى دول البحر الأبيض المتوسط (MED- HIMS) الذى يمثل مبادرة من الاتحاد الأوروبي، البنك الدولى، صندوق الأمم المتحدة للسكان، المفوضيه السامية للأمم المتحدة لشئون الأجنئين، منظمة العمل الدولية، منظمة الهجرة الدولية، وجامعة الدول العربية بالتعاون مع أجهزة الإحصاء القومية فى الدول العربية فى منطقة جنوب وشرق المتوسط وذلك مطلع عام 2009. والذى يهدف إلى توضيح العلاقات التداخلية للهجرة والتنمية بين مصر والدولة المستقبلية للمهاجرين ، إلى جانب توفير البيانات حول توقيتات أماكن، وكيفية حدوث الهجرة ودراسة أبعادها. وقد بلغت عينتها 90 ألف أسرة والتي تضمنت عدد 5 استمارات استبيان وحيث ان اهداف البحث تهتم بدوافع الهجرة والعودة من والى مصر. وقد استخدمت البيانات الخام لهذا المسح استمارة المهاجر العائد منذ 1 / 1 / 2000 والتي بلغت حجم عينتها 5085 مهاجر عائد كانت اعمارهم 15 سنة فأكثر وقت العودة من الخارج وذلك بنسبة استجابة بلغت 99.1%، واستمارة المهاجر الحالي وبلغت حجم العينة 5847 مهاجر حالي كانت اعمارهم 15 سنة فأكثر وقت اجراء المسح بنسبة استجابة بلغت 99.9%.

نتائج الدراسة

لتحقيق هدف الدراسة فى التعرف على الاختلافات المعنوية بين دوافع المهاجرين (دوافع ظروف العمل فى مصر – دوافع اقتصادية – دوافع اجتماعية – دوافع اخرى) المهاجرين الحاليين والعائدين للهجرة تبعاً لخصائص هؤلاء المهاجرين استخلصت الدراسة النتائج التى سنعرضها على النحو التالى :-

1- دوافع المهاجرين تبعاً للنوع

- اوضحت النتائج ان حوالى 96% من المهاجرين العائدين من الذكور ، فى حين بلغت تلك النسبة للمهاجرين الحاليين حوالى 98% مما يفسر انتقائية الهجرة تبعاً للنوع حيث تبين وجود فرق معنوى بين هاتين النسبتين حيث بلغت قيمة $Z= 5,172$ ، $P\text{-Value}= 0,00$. مما يتضح ان هناك اتجاه تزايد لانتقائية هجرة الذكور فى الوقت الحالى عنهم كان قبل عام 2000 وقد يرجع هذا الى اتجاه دول المهجر واحتياجاتهم للمهاجرين الذكور اكثر من احتياجاتهم ورغبتهم فى استقدام الاناث .
- ابرزت نتائج تحليل دوافع هجرة المصريين سواء العائدين او الحاليين تبعاً للنوع ان نسبة الذكور المهاجرين حالياً بسبب دوافع اجتماعية بلغت حوالى 91% فى حين كانت نسبتهم بين المهاجرين العائدين حوالى 39% وتحليل اختبارات الفروض بين تلك النسبتين تبين وجود فرق معنوى حيث بلغت قيمة $Z= 13,78$ ، $P\text{-Value}= 0,00$.
- فى حين تبين عدم وجود فروق معنوية بين باقى دوافع الهجرة ، كذلك بقياس الاختلاف بين نوع المهاجر بدوافع الهجرة لكل من المهاجرين الحاليين والعائدين
- بينت النتائج عدم وجود اختلاف بين النوع ودوافع هجرة المهاجرين الحاليين والعائدين حيث بلغت قيمتى كا2 (x^2) المحسوبة= $(301.1 ، 1394.8)$ على الترتيب ، وهذا ما يتضح من الجدول رقم (1)

2- دوافع المهاجرين تبعاً لسابقة العمل

- كما ابرزت نتائج دوافع هجرة المصريين سواء العائدين او الحاليين بدافع اقتصادى تبعاً لسابقة عملهم قبل الهجرة ان نسبة المهاجرين العائدين بهذا الدافع حوالى 97% من سبق لهم العمل فى حين بلغت تلك بالنسبة للمهاجرين الحاليين حوالى 92% وتحليل اختبارات الفروض بين تلك النسبتين تبين وجود فرق معنوى بينهما حيث بلغت قيمة $Z = 6,58$ ، $P\text{-Value} = 0,00$ وهذا قد يرجع لانخفاض الاوضاع الاقتصادية بصفة عامة ولمن سبق لهم العمل قبل الهجرة بصفه خاصة .
- كذلك الامر بين دافع الهجرة لمن سبق لهم العمل قبل الهجرة بين المهاجرين العائدين والحاليين بسبب ظروف العمل فى مصر حيث بلغت نسبتهم حوالى 79% ، حوالى 59% وبلغت قيمة $Z = 7,66$ ، $P\text{-Value} = 0,00$.
- فى حين تبين انخفاض نسبة من سبق لهم العمل قبل الهجرة بسبب الدوافع الاجتماعية بين المهاجرين العائدين والحاليين (حوالى 44% ، حوالى 69%) على الترتيب حيث بلغت قيمة $Z = 5,55$ ، $P\text{-Value} = 0,00$ ، وهذا ما يوضحه الجدول رقم (1) .

3- دوافع المهاجرين تبعاً لفئات العمر عند الهجرة

- اوضحت النتائج ان حوالى 45% من المهاجرين العائدين هاجروا وهم فى الفئة العمرية (25-34) ، فى حين بلغت تلك النسبة للمهاجرين الحاليين حوالى 42% مما يفسر انتقائية الهجرة تبعاً للعمر حيث تبين عدم وجود فرق معنوى بين هاتين النسبتين، حيث بلغت قيمة $Z = 2.1$ ، $P\text{-Value} = 0.018$. وقد يرجع هذا الى اتجاه دول المهجر واحتياجاتهم للمهاجرين بالفئة العمرية (25-34) اكثر من الفئات الاخرى .
- ابرزت نتائج تحليل دوافع هجرة المصريين سواء العائدين او الحاليين تبعاً للعمر عند الهجرة ان الدافع الاقتصادي للمهاجرين هو من أكثر الدوافع للهجرة ، حيث بلغت نسبتهم تبعاً للفئة العمرية للمهاجر عند الهجرة (25-34) سنة حوالى 46% ، فى حين يأخذ الدافع الاقتصادي لنفس الفئة

بين المهاجرين الحاليين المرتبة الثانية من دوافعهم للهجرة ، حيث بلغت نسبتهم 43% وبإجراء اختبار الفروض بين تلك النسبتين تبين عدم وجود فرق معنوي بينهما ، حيث بلغت قيمة $z=1.9$, p-value 0.029

- كذلك تبين عدم وجود فروق معنوية بين نسب باقى دوافع الهجرة سواء بين المهاجرين الحاليين أو العائدين من حيث فئات أعمارهم عند الهجرة .

- بينت النتائج عدم وجود اختلاف بين فئات العمر عند الهجرة ودوافع هجرة المهاجرين الحاليين العائدين حيث بلغت قيمتى χ^2 (x^2) المحسوبة = (247.7 ، 214.6) على الترتيب ، كما هو موضح بالجدول رقم (1)

4- دوافع المهاجرين تبعاً للحالة التعليمية عند الهجرة

- اوضحت النتائج ان حوالى 51% من المهاجرين العائدين كانت حالتهم التعليمية وقت الهجرة (ثانوي وأقل من جامعي) ، فى حين بلغت تلك النسبة للمهاجرين الحاليين حوالى 56% مما يفسر انتقائية الهجرة تبعاً للمرحلة التعليمية حيث تبين وجود فرق معنوى بين هاتين النسبتين حيث بلغت قيمة $P\text{-Value}= 0,00$ ، $Z= 4.11$. مما يتضح ان هناك اتجاه متزايد لانتقائية الهجرة فى التعليم ، وقد يرجع هذا الى اتجاه دول المهجر واحتياجاتهم للمهاجرين ذوي المستوى التعليمي الأعلى .

- ابرزت نتائج تحليل دوافع هجرة المصريين سواء العائدين او الحاليين تبعاً للحالة التعليمية عند الهجرة ان المرحلة التعليمية ثانوى واقل من جامعى من المهاجرين حالياً بسبب دوافع اقتصادية بلغت حوالى 58% فى حين كانت نسبتهم بين المهاجرين العائدين حوالى 55% وتحليل اختبارات الفروض بين تلك النسبتين تبين وجود فرق معنوى حيث بلغت قيمة $Z= 2.5$ ، p-value = 0.006 .

- فى حين تبين عدم وجود فروق معنوية بين باقى دوافع الهجرة ، كذلك بقياس الاختلاف عمر المهاجر بدوافع الهجرة لكل من المهاجرين الحاليين والعائدين

- بينت النتائج عدم وجود اختلاف بين العمر عند الهجرة ودوافع هجرة المهاجرين الحاليين العائدين حيث بلغت قيمتى كاي² (χ^2) المحسوبة= (149.9 ، 174.8) على الترتيب على الترتيب ، كما هو موضح بالجدول رقم (1)

5- دوافع المهاجرين تبعاً لمدة الهجرة

- اوضحت النتائج ان حوالى 56% من المهاجرين العائدين كانت مدة هجرتهم بالخارج ما بين عام الى خمسة أعوام ، فى حين بلغت تلك النسبة للمهاجرين الحاليين حوالى 41% ، حيث تبين وجود فرق معنوى بين هاتيين النسبتين حيث بلغت قيمة $Z = 9.88$ ، $P\text{-Value} = 0,00$.
- بينت النتائج عدم وجود اختلاف بين العمر عند الهجرة ودوافع هجرة المهاجرين الحاليين العائدين حيث بلغت قيمتى كاي² (χ^2) المحسوبة= (44.0 ، 57.3) على الترتيب على الترتيب ، كما هو موضح بالجدول رقم (2)

6- دوافع المهاجرين تبعاً لصاحب قرار الهجرة

- اوضحت النتائج ان حوالى 91% من المهاجرين العائدين كانوا اصحاب القرار الاول فى اهجرتهم هو(المهاجر نفسه) ، فى حين بلغت تلك النسبة للمهاجرين الحاليين حوالى 83% ، حيث تبين وجود فرق معنوى بين هاتيين النسبتين حيث بلغت قيمة $P\text{-Value} = 0,00$ ، $Z = 9.88$.
- ابرزت نتائج تحليل دوافع هجرة المصريين سواء العائدين او الحاليين تبعاً لصاحب قرار الهجرة (المهاجر نفسه) من المهاجرين حالياً بسبب دوافع اقتصادية بلغت حوالى 96% فى حين كانت نسبتهم بين المهاجرين العائدين حوالى 98% وتحليل اختبارات الفروض بين تلك النسبتين تبين وجود فرق معنوى حيث بلغت قيمة $P\text{-Value} = 0,00$ ، $Z = 3.73$.

- فى حين تبين وجود فروق معنوية بين باقى دوافع الهجرة ، كذلك بقياس الاختلاف صاحب قرار الهجرة (المهاجر نفسه) بدوافع الهجرة لكل من المهاجرين الحاليين والعائدين
- بينت النتائج عدم وجود اختلاف بين العمر عند الهجرة ودوافع هجرة المهاجرين الحاليين العائدين حيث بلغت قيمتى χ^2 (المحسوبة= 302.9 ، 2286.2) على الترتيب على الترتيب ، كما هو موضح بالجدول رقم (2)

7- دوافع المهاجرين تبعاً للتمويل المالى للهجرة

- اوضحت النتائج ان حوالى 81% من المهاجرين العائدين اعتمد على نفسه فى تمويل تجربة هجرته ، فى حين بلغت تلك النسبة للمهاجرين الحاليين حوالى 61% ، حيث تبين وجود فرق معنوى بين هاتيين النسبتين حيث بلغت قيمة $Z = 17.9$ P-Value= 0,00 .
- ابرزت نتائج تحليل دوافع هجرة المصريين سواء العائدين او الحاليين للتمويل المالى للهجرة (المهاجر نفسه) من المهاجرين حالياً بسبب الدوافع الاقتصادية بلغت حوالى 60% فى حين كانت نسبتهم بين المهاجرين العائدين حوالى 84% وتحليل اختبارات الفروض بين تلك النسبتين تبين وجود فرق معنوى حيث بلغت قيمة $Z = 18.3$ ، P-Value= 0,00 كما هو موضح بالجدول رقم (2)

التوصيات

- على الدولة تشجيع الباحثين على القيام باعداد الدراسات والابحاث المتخصصة للهجرة الدولية حيث تعد ظاهرة مجتمعية مركبة ومتعددة الابعاد، وهو من شأنه الاهتمام بالدراسات الاحصائية للحصول على بيانات دقيقة وممثلة للمجتمع المصرى فيما يخص اعداد المصريين المهاجرين بالخارج وذلك برصد بيانات المهاجرين بشكل دورى.

- توفير فرص استثمارية للمهاجرين العائدين سواء انشاء شهادات تامين بالدولار تدفع من الخارج باقساط مما يشجعهم على استثمار اموالهم وهم بالخارج وبالتالي استفادة الاقتصاد من هذه التحويلات في الاقتصاد الرسمي.
- بالرغم من غلبة التعليم المتوسط وفوق المتوسط على المستوى التعليمي للمهاجرين الحاليين والعائدين فإنه لا يمكن تجاهل تمتع نسبة منهم بمهارات عالية تميزت باحتكاك خارجي ساعد على نموها وهو ما يتطلب بناء روابط مع اصحاب المهارات منهم للاستفادة بها في المشروعات الاقتصادية سواء التي تتحمل مستوى تكنولوجيا عالي او تحتاج لمستوى إداري واحترافي.

المراجع

- السعدني، سمية؛ محمود. (1993). دوال للهجرة الدولية للعمل في حالة مصر. دراسة دكتوراه، معد الدراسات والبحوث الاحصائية، جامعة القاهرة. مصر.
- إمام، مجدة. (2009). هجرة الشباب الدولية والتنمية «الفرص والتحديات». قراءات شبابية: التنمية اجملمتعية واحكم الصالح، مركز اعلام حقوق الانسان والديمقراطية – شمس. 178-208.
- زهري ايمن؛ اخرون (2009). إتجاهات الشباب المصري حول الهجرة لأوروبا. قطاع الهجرة بوزارة القوى العاملة.
- محمود عبد الكريم، عبد الله عبد الرازق و مديحة عبد الحليم (2009). ورقة عمل عن واقع إحصاءات الهجرة الدولية فى جمهورية مصر العربية. مقدمة إلى ورشة العمل التدريبية حول إحصاءات الهجرة الدولية التى تنظمها اللجنة الاقتصادية والاجتماعية لغرب آسيا (الاسكوا) خلال الفترة من 6/30-2009/7/2 بالقاهرة.
- البرعي، نادية. مرسي؛ مظلوم، نهلة. كمال؛ عياد، مها. السيد؛ الضو، نوال. عبدالمنعم؛ جلال الدين، محمد؛ عبدالوهاب، فكرات؛ أحمدى، مرفت والعشري، فاطمة. محمد. (2013). الهجرة الدولية للشباب. مجلة السكان بحوث ودراسات، الجهاز المركزي للتعبئة العامة والاحصاء. 1-41. مصر.

- البرنامج الإنمائي للأمم المتحدة. (2013). تقرير التنمية البشرية 2013 "نهضة الجنوب: تقدم بشري في عالم متنوع". نيويورك.
- الجهاز المركزي للتعبئة العامة والإحصاء. (2014). مسح الهجرة الدولية 2013. مصر.
- متولي سها أحمد. (2014). تأثير الهجرة الداخلية على الفقر في مصر. دراسة دكتوراه، معهد الدراسات والبحوث الإحصائية، جامعة القاهرة. مصر.
- حسن محمد مصطفى، زهري أيمن، قنديل سعد محمد ، عبدالمنعم محمد. (2016). الهجرة الدولية وعلاقات الفقر والمكان في مصر: أدلة تطبيقية من مسح الدخل والانفاق والاستهلاك. المؤتمر السنوي 51 للإحصاء، علوم الحاسب وبحوث العمليات من 5 إلى 8 ديسمبر 2016.
- عبدالعزيز خالد سلطان. (2016). الدوافع. موسوعة المقاتل. (تاريخ الدخول نوفمبر 2016:

<http://www.moqatel.com/openshare/Behoth/Mnfsia15/Dawafeia/sec01.do>
(c_cvt.htm)

- Masja van Meeteren, Godfried Engbersen, Erik Snel and Marije Faber (2014). Understanding Different Post-Return Experiences The Role of Preparedness, Return Motives and Family Expectations for Returned Migrants in Morocco. CMS, 2(3): 335-360.
- Velizarova, M (2009). Motivation for Migration, Return and Sustainability of the Return a Case Study of Bulgarian Student Mobility. Regime Change and Transitions across the Danubian Region, 1989-2009, 153-164.
- Houte, M. van and T. Davids. (2014). Moving Back or Moving Forward? Return Migration, Development and Peace-building. New Diversities, 16(2) (Special issue: Migration and Development: Rethinking

Recruitment, Remittances, Diaspora Support and Return, edited by Ninna Nyberg Sørensen): 71-87.

جدول (1)

التوزيع النسبي للمهاجرين العائدين والحاليين حسب خصائصهم طبقا لدوافع هجرتهم

المهاجرين الحاليين							المهاجرين العائدين							البيان
ك2	d.f	النسبة الاجمالية %	دوافع أخرى	دوافع اجتماعية	دوافع اقتصادية	ظروف العمل في مصر	ك2	d.f	النسبة الاجمالية %	دوافع أخرى	دوافع اجتماعية	دوافع اقتصادية	ظروف العمل في مصر	الدافع للهجرة الخصائص
301.1	3	98 1.7	84.5 15.5	90.8 9.2	99.6 0.4	99.2 0.8	1394.8	3	96.2 3.8	93.2 6.8	38.8 61.2	99.7 0.3	99.2 0.8	النوع ذكر أنثى
710.3	3	81.8 18.2	60.6 39.4	69.1 30.9	92.4 7.6	58.6 41.4	541.1	3	90.6 9.4	89.8 10.2	44.2 55.8	96.8 3.2	78.8 21.2	سابقة العمل قبل الهجرة سبق له العمل لم يسبق له العمل
247.4	15	42.3 42.0 11.8 2.7 0.3 0.9	47.9 33.8 15.5 1.4 - 1.4	65.1 28.2 2.1 0.6 0.2 3.8	38.0 43.4 14.6 3.2 0.3 0.5	43.2 44.9 8.3 2.6 0.4 0.7	214.6	15	36.2 44.0 14.4 3.1 0.9 1.3	32.4 37.0 16.7 2.8 4.6 6.5	42.0 36.5 8.9 2.9 2.7 7.0	34.7 46.1 15.5 2.8 0.5 0.4	39.2 41.8 13.3 4.8 0.7 0.2	فئات العمر عند الهجرة 24-15 34-25 44-35 54-45 64-55 65 فأكثر
149.9	9	18.5 6.3 57.3 17.9	11.3 4.8 50.0 33.9	17.7 8.7 55.9 17.7	21.1 7.0 58.3 13.7	12.3 3.5 56.0 28.2	174.8	9	23.1 6.4 53.8 16.7	13.8 13.8 41.4 31.0	10.9 5.1 51.3 32.7	27.2 6.5 54.8 11.5	17.6 5.6 53.7 23.1	الحالة التعليمية عند الهجرة أقل من أجمالي أجمالي و أقل من ثانوي ثانوي و أقل من جامعي جامعي فأعلى العدد
		4837	71	524	3154	1088			2712	59	147	2035	471	

المصدر: محسوب بمعرفة الباحث من بيانات مسح الهجرة الدولية 2013، الجهاز المركزي للتعبئة العامة والإحصاء، مصر.

جدول (2)

التوزيع النسبي للمهاجرين العائدين والحاليين حسب خصائصهم طبقاً لدوافع هجرتهم

المهاجرين الحاليين							المهاجرين العائدين						البيان	
كأ	d.f	النسبة الاجمالية %	دوافع أخرى	دوافع اجتماعية	دوافع اقتصادية	ظروف العمل في مصر	كأ	d.f	النسبة الاجمالية %	دوافع أخرى	دوافع اجتماعية	دوافع اقتصادية	ظروف العمل في مصر	الدافع للهجرة الخصائص
														<u>مدة الهجرة</u>
														أقل من سنة
														1 – 5
														5 – 10
														10 – 15
														15 - فاكثر
														<u>صاحب قرار الهجرة</u>
														المهاجر نفسه
														الزوج/ة أو الخطيب/ة
														الأبناء والأقارب
														صاحب العمل
														<u>التمويل المالي للهجرة</u>
														المهاجر نفسه
														أخرى
														العدد
		4837	71	524	3154	1088			2712	59	147	2035	471	

المصدر: محسوب بمعرفة الباحث من بيانات مسح الهجرة الدولية 2013، الجهاز المركزي للتعبئة العامة والإحصاء، مصر.