CONTENTS

- *TABLE OF CONTENTS.*

- *LIST OF TABLES.*

- *LIST OF GRAPHS.*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF GRAPHS

DEDICATION

✳ *To my dear devoted father*

Who has always been encouraging me and paving the avenues and win ding roads at the cost of his priceless blood and life. To his all streamlin ed sacrifices, endless benevolence and everlasting giving.

✳ *To my lovely mother*

Who has been teaching me the right things. To her patience, toleran ce and intimate motherhood welfare.

✳ *To my dear husband*

Who has been continuously encouraging me and fully supporting m e.

✳ *To my sincere brothers & sisters*

Who always encourage me.

✳ *To my dearest and lovely children.*

✳ *To my faithful friends.*

The Researcher..

CHAPTER ONE

INTRODUCTION

1 - 0   INTRODUCTION.

1 - 1   RESEARCH PROBLEM.

1 - 2   RESEARCH IMPORTANCE.

1 - 3   RESEARCH OBJECTIVES.

1 - 4   RESEARCH HYPOTHESES.

1 - 5   RESEARCH DATA.

1 - 6   METHODOLOGY.

1 - 7   PREVIOUS STUDIES.

1 - 8   RESEARCH PLAN.

1 - 9   RESEARCHER PUBLISHED PAPERS.

CHAPTER ONE

INTRODUCTION

1. <u>INTRODUCTION:-</u>

The endeavors to operate many types of aeroplanes in Sudan, offering a complete range of modern services, was reached by the Sudanese Authorities, a considerable time ago; but this dream could not be materialized.

Sudan, the heart of Africa and the largest country in Africa, uniting widely separated regions and populations of diverse characters for whom, through history, has been the sole focal point of mutual contact and trade.

Sudan is located in Northeastern Africa. It is bordered by Egypt to the north, the Red Sea to the northeast, Eritrea and Ethiopia to the east, South Sudan to the south, the Central African Republic to the southwest, Chad to the west and Libya to the northwest. It is the third largest country in Africa. (It had been the largest until South Sudan became an independent state on 9 July 2011) .Geographic coordinates are 15°00′N 30°00′E. After separation from the South three years ago, Sudan's total remaining area amounted to some 1,861,484 square km (Sultan, 2012).

On account of its geographical and economic characteristics, Sudan has a pronounced vocation to become the main air transport centre of the whole Continent, as well as regional and international air traffic.

The government orientations in the national strategic planning, investment encouragement, production & export of Sudanese petroleum resources, tourism development, peace and political stability, all combine to

generate very high demands for air traffic in the near future. Most of this will be focused, having good reliable airlines, to offer good facilities of the highest international standards.

1.1. RESEARCH PROBLEM:-

1.1.1. Statistical Data:

Difficulty in getting statistical data from the airlines for so many unjustifiable reasons! Most companies were reluctant in revealing their actual revenues. They considered this issue as top confidential! Such a negative attitude is ethically questionable, as we do not appreciate their hindering the progressing development of such scientific economical national studies, which are vitally important for Sudan development in the aviation field.

1.1.2. Preface:

Research Problem Limits:

1. Covers all States of Sudan.
2. Domestic civil airlines movements in Sudan carrying passengers and or cargo during 2004- 2013.

1.1.3. Reasons for Research Problems:

1. The vital strategic importance of civil aviation services and role in Sudan.
2. Few, poor, unsafe, unpaved, narrow roads all over the country, complete destruction and stoppage of railways and river way transportation services due to crucial, devastating political reasons since May Military Coup Revolution in 1969.

3. There is deterioration in Sudan airlines domestic services in terms of specific factors, e.g. quality and quantity.

4. There is a gradual rise in the total operating costs of the main airlines firms in Sudan.

5. There are specific factors contributing to the reduced efficiency of Sudan airlines.

1.2. <u>RESEARCH IMPORTANCE:-</u>

Sudan is a republic in north-eastern Africa, the largest country of the African continent. It is bordered on the north by Egypt, on the east by the Red Sea, Eritrea, and Ethiopia, on the south by South Sudan south, and the Central African Republic to the southwest and on the west by Chad, and Libya. Sudan has a total area 1,861,484 .sq. km. Khartoum is the capital and largest city. South Sudan became an independent state on 9 July 2011.

Such a large strategic country, in the heart of Africa, definitely requires many reliable, capable modern domestic and international airline carriers to meet its development in various fields of investment and discoveries, as transportation of passengers & cargo/freight, oil exploration, mining, industry, agriculture, etc.

Because of that, we are interested in finding out how the total cost behaves in relation to the output, in revenue passenger and cargo, fuel price and load factor. The result shall lead us to estimate an airline function.

A properly estimated cost model allows airlines to more accurately

forecast cost:

- As a function of changes in average fares.
- Given recent or planned changes to frequency of service.
- To account for changes in market or economic conditions.

1.3. RESEARCH OBJECTIVES:-

1. Evaluation of Sudanese Civil Airlines domestic Services in Sudan in regards to Passenger and Cargo Movements.
2. Statistical estimation of a Sudanese airline cost function to identify the extent of aviation development in Sudan.
3. Determination of the main causes of the inefficiency of Sudanese civil airlines services compared to the large area of the country.
4. Endeavours to upgrade & promote civil aviation activities in Sudan to meet the future requirements for Sudan developing economy.

1.4. RESEARCH HYPOTHESES:-

Cost models are mathematical representation of the relationship between the total cost and explanatory variables (the output, in revenue passenger and cargo, fuel price and load factor), model specification reflects expectations of cost behavior, by using:

- The Classical Normal Linear Regression Model (CNLRM), to evaluate the model for forecasting, by satisfying the main features of a good regression model.
- Panel Regression Models: Pooled Ordinary Least Square (OLS) model, Fixed Effects Least Squares Dummy Variable (LSDV) model, Fixed Effects within-group (WG) model, and the Random Effects model (REM). The results were evaluated to determine the best

17

suitable model to estimate an airline cost function.

1.5. <u>RESEARCH DATA</u>:-

The Data was obtained from:

1. Sudan Civil Aviation Authority.
2. Sudanese Airlines Companies (Sudan Airways, Marsland Aviation, Sun Air, Nova Air, Mid Airlines and Badr Airlines).
3. Aircraft Operating Calculators for specifications of each aircraft type.
4. Nile Bakri Aviation Co. Ltd. for aircraft fuel prices ,
5. Bank of Sudan for currency rates.

Aircraft Fuel Cost Calculation Per Annum is based on:

1. Aircraft Type.
2. Aircraft Annual usage.
3. Aircraft load Factor.
4. Aircraft Cruise Speed.
5. Average Flight Distance.
6. Fuel Price per Liter.

Aircraft Type Total Cost Calculation Per Annum is based on:

1. Fuel Cost.
2. Finance Cost.
3. Fixed Costs.
4. Tyres Cost.
5. Maintenance Cost.
6. Service Cost.

1.5.1. Average Number of Flights:

The average number of flights per annum calculation is based on the average of forty weeks per annum multiplied by seven days per week (7×40=280). The remaining twelve weeks of the year are always catered for normal grounding time, scheduled aircraft maintenance, unscheduled aircraft maintenance, defect certifications, operations, traffic, loading, etc...

1.5.2. Load Factor:

In aeronautics, the load factor is defined as the ratio of the lift of an aircraft to its weight and represents a global measure of the stress (load) to which the structure of the aircraft is subjected:

$$LF = \frac{L}{W}$$

Where:

LF= Load factor   L=Lift   W=Weight

Since the load factor is the ratio of two forces, it is dimensionless. However, its units are traditionally referred to as **G**, because of the relation between load factor and apparent acceleration of gravity felt on board the aircraft. A load factor of one (1G) represents conditions in straight and level flight, where the lift is equal to the weight. Load factors greater or less than one (or even negative) are the result of maneuvers or wind gusts.

In this research the lift (L) represents the number of passengers per annum in a passenger aircraft, while it represents the freight per annum in a freight aircraft. While the Weight (W) is calculated as passenger capacity per day multiplied by the average of flights per annum.

1.6. <u>METHODOLOGY</u>:-

1.6.1. Theoretical Frame:

Analysis of the data obtained from the Planning Directorate of Sudan Civil Aviation Authority, Air Transport Directorate and Sudan Airways Directorate of Central Planning and some other currently active Sudanese airlines, namely Sudan Airways, Marsland Airlines, Badr Airline, Nova Airlines, Sun Air Airlines and Mid Air Airlines shall be conducted. These data consist of the total number of passengers and freight/cargo carried domestically in Sudan, and also the data consist of the total number and types of aircraft in each Sudanese airlines through the years from 2004 to 2013. For each airline these data shall be tabulated for each year separately, in addition to the data of Fuel price obtained from Nile Bakri Aviation Co.Ltd., and price of currency from Bank of Sudan; which were used in calculation of airlines total cost.

1.6.2. Specifications and Estimation of the Models:

- The data analyze the annual cost of the six Sudanese airlines companies for the period from 2004 to 2013.
- The researcher is interested in finding out how the total cost (TC) behaves in relation to the domestic output, in revenue passenger (PAX) and cargo/ freight (FRT), fuel cost (FC) and load factor (LF). The result shall lead us to estimate an airline cost function, by using Classical Normal Linear Regression Model (CNLRM), and evaluate the model for forecasting.

- The data statistically analyze the annual cost function of the five Sudanese airlines companies for the period from 2004 to 2013, for a total of 50 Balanced Long Panel Data observations.
- Panel Data Regression Models shall be applied in this research to study the four possibilities models: Pooled Ordinary Least Square (OLS) model, Fixed Effects Least Squares Dummy Variable (LSDV) model, Fixed Effects within-group model, and the Random Effects model (REM). These results shall be evaluated to determine the best suitable model for estimating an airline cost function.
- Panel Data Regression models shall be applied by using the Eview Statistical Package to estimate an airline cost function.

1.7. PREVIOUS STUDIES:-

1. In (2013), Department of Management, Okan University, Orhan Sivrikaya presented a study titled as" Demand Forecasting for Domestic Air Transportation in Turkey". In this study a semi-logarithmic regression model is generated in order to estimate the domestic air travel demand by means of a number of passengers carried per city pair. Airline passenger data of 2011 out of 42 served cities in Turkey, are used to establish the model. Then 2010 data are used to test prediction performance of the model. Accuracy level is found to be significantly successful by estimating passenger demand for any domestic city pair. Due to its city pair basis and acceptable level of accuracy, the estimation model can be utilized in many areas of aviation industry.

2. In November (2013), Department of Economics, North West University, Mr. Olebogeng Ambrocius Baikgaki, presented a study titled by: "The Determination of Domestic Air Passenger Demand in the Republic of South Africa". The methodology used involved collection of data for passenger movement in South Africa for the period 1971-2012 to determine the pattern of air travels. Data on macro and micro – economic variables were considered to affect demand for air passenger. The models were selected to identify and measure the relation between domestic air passenger demand and the economic and demographic factors in South Africa. The economic model analyzes the demand for air passenger movement through establishing a relationship between the candidate independent variables and the dependent variable (domestic air passenger demand). It was then established that only four variables (population, airfares, oil prices and the level of household consumption) are most significant and main determinants of domestic air passenger demand. Also, this model is very good in terms of goodness of fit measures and doesn't suffer from multicollinearity.

3. In August (2012), School of Economics, Erasmus University Rotterdam, Coglar Demirsoy, presented a study titled by: "Analysis of Stimulated Domestic Air Transport Demand In Turkey, What Are The Main Drivers?" Which investigates phenomenal growth of the Turkish domestic air transport demand in the last three decades, using panel data approach with fixed effect specification model. As a result, the study suggests that income and population changes are determined as the main drivers of Turkish domestic air transport demand, and they are important

estimators for forecasting prospective traffic and passenger increases in domestic air transport sector.

4. In (2011), Journal of Transport Geography, Dobruszkes, presented a study titled by:" An Analysis of the Determinants of Air Traffic Volume for European Metropolitan Areas". In this study he used total air service as a dependent variable and step-wise approach was applied to decide which variables were significant to determine air transportation demand in urban regions in Europe. From population, gross domestic product, national administrative function, international administrative function, economic decision-power, knowledge and scientific research, tourism and distance to the nearest main air market, only GDP, economic decision-power, tourism and distance to the main air market turned out to be significant. Consequently, the model with four independent variables explains 70% of variation in the changing air transport service demand, where, GDP is the strongest determinant. This study found relatively small GDP coefficient with respect to other researches. Regression shows that GDP has a 0.39 coefficient while standardized coefficient of economic decision power is 0.29. Moreover, the distance to the nearest main air market and tourism have positive influence on the total air services, respectively 0.27 and 0.23.

5. In (2010), University of Ilorin, Nigeria, Adekunle J.Aderamo, presented a study titled by:" Demand for Air Transport in Nigeria". The study seeks to assess the demand for domestic air transport in Nigeria and the factors responsible for it. In particular, it looks at passenger, aircraft and freight traffic and the relative demand for them in the country. The

Methodology used involved collection of data on passenger, aircraft and cargo movements in Nigeria for the period 1975-2006 to determine the pattern of air travels. Data on macro and micro-economic variables considered to affect demand for air travel were also collected. Multiple regression method was then used to develop models of demand in respect of the three types of movement. The results showed that, of the selected variables, Index of Agricultural Production, Index of Manufacturing Production, Gross Domestic Product, Inflationary Rate and Consumer Price Index are important in the explanation of the demand for air transport in Nigeria. The study suggested that, in order to encourage the demand for air travel in Nigeria, the government need arises to improve the transportation system in the country.

6. In (2010), Journal of Transportation Research, Marazzo, presented a study titled by:" Air Transport Demand and Economic Growth in Brazil: A Time Series Analysis". In this study he investigated only the inter-related relationship between GDP and passenger numbers. Research results showed that passenger number is reacting strongly to changes in the GDP, but GDP reacts to changes in the passenger number slower.

7. In March (2008), University of British Columbia, Tae Hoon Oum presented a study titled by: "Ownership Forms Matter for Airport Efficiency: A Stochastic Frontier Investigation of Worldwide Airports". This studies the effects of ownership forms on airports' cost efficiency by applying stochastic frontier analysis to a panel data of the world's major airports. He estimated a stochastic frontier cost model in translog form via a Bayesian approach in order to measure the effects of

ownership and institutional forms on the efficiency of airports. These findings imply the following:

- Countries concerned to privatize airports should transfer 100% or a majority ownership to private sector; and should avoid the mixed ownership with government majority in favor of even 100% government owned public firm;

- US should reconsider ownership and management of airport by port authorities;

- Although average efficiency of the airports owned and operated by cities/states are lower than those operated by independent airport authorities, the difference is not statistically significant. As such, this issue needs careful further examinations.

8. In October (2008), University of Mauritius, Reduit. Jameel Khadaroo presented a study titled by: "The Role of Transport Infrastructure in International Tourism Development: A Gravity Model Approach". This study employs a gravity framework to evaluate the importance of transport infrastructure in determining the tourism attractiveness of destination. The analysis is based on a panel data set of bilateral tourism flows among 28 countries over the decade 1990-2000. It found that, on top of tourism infrastructure and other classical determinants, transport infrastructure is a significant determinant of tourism inflows into a destination. Disaggregated continent-wise analysis reveals that the sensitiveness of tourism flows to transport infrastructure does vary, depending on origins and destination. Also, found evidence of repeated tourism around the world, the more so from high-income origins and to high-income destinations.

9. In (2008), Journal of Air Transport Management, Bahadra, presented a study titled by: "A Basic Empirical Analysis of Domestic Passenger Demand". In this study the author used personal income and distances between two O&D destinations. He segmented the markets according to market type. These markets are super thin, and semi-thick and tick markets. However, population turns out to have a correlation with other sub-samples and has to be removed from the model. The model that this study built, explains 60% of variation in the thick markets demand. In conclusion, fares, real personal incomes, and distances are strong determinants of air transport demand according to the outcome of this research.

10. In (2007), University of Mannheim, Tobias Grosche, presented a study titled by:" Gravity Model for Airline Passenger Volume Estimation". This study presented two gravity models for the estimation of air passenger volume between city-pairs. The models included variables describing the general economic activity and geographical characteristic of city-pairs instead of variable describing air service characteristic. Thus, both models can be applied to city-pairs where currently no air service is established, historical data is not available, or for which factors describing the current service level of air transportation are not accessible or accurately predictable. One model was limited to city-pairs with airports not subject to competition from airport in the vicinity, while the other models included all city-pairs. Booking data of flight between Germany and 28 European countries was used for calibration. Both

models showed a good fit to the observed data and were statistically tested and validated.

11. In (2005), Public Works Management and Policy, Bhadra, presented a study titled by: "Air Travel by State: Its Determinants and Contributions in the United States'. In this study, the authors used O&D passenger number as a dependent variable and average fare, gross state production, population and number of hubs as an independent variable. Result showed that there was a close relation (unit elastic) between GSP and passenger numbers with a coefficient of GSP 0.95. In other words, increase of 1% in the state product results in 0.95% increase in passenger number. Furthermore, passenger demand and location of airports are strongly affected by population and Economic activities.

12. In April (2005), Carlos III University, Madrid, Spain, Vicent Inglada presented a study titled by: "Liberalisation and Efficiency in International Air Transport". This study was set out to compare the economic and technical efficiency of international air transport companies, within the new liberalisation framework that characterises the period of 1996-2000. For this purpose, two stochastic frontiers were estimated, one for cost function, the other for production function. From these estimations we obtain indexes for, respectively, economic and technical efficiency. Our evidence suggests that the benefits of increasing competition in terms of efficiency, is being large for the Asian companies.

13. In (2004), Cranfield University, John, F. O☐ connell, presented a study titled by: "Passengers☐ Perception of Low Cost Airlines and Full Service Carriers – A case Study involving Ryanair, Air Lingus, Air Asia and Malaysia Airlines". This study attempted to provide answers about direct competition between full service airlines and no-frills carriers is intensifying across the world. American and European full service airlines have lost significant proportion of their passengers to low cost carriers, data collected during 1997-2004. As a result of this study, it is apparent that passengers travelling on incumbents, place strong emphasis on reliability, quality, flight schedules, connections, frequent flyer programs and comfort, while travelers taking low cost carriers focus almost exclusively on fare. It would seem therefore, that passengers would like to see the two airline models become ever closer.

14. In (2002), Deel University of Trieste, Lorenzo Castelli, presented a study titled by: "An Airline-Based Multilevel Analysis of Airfare Elasticity for Passenger Demand". In this study price elasticity of passenger demand for a specific airline was estimated. The main drivers affecting passenger demand for air transportation were identified. First, an Ordinary Least Squares regression analysis was performed. Then, a multilevel analysis-based methodology to investigate the pattern of variation of price elasticity of demand among the various routes of the airline under study was proposed. The experienced daily passenger demands on each fare-class were grouped for each considered route. 9 routes were studies for the months of February and May in years from 1999 to 2002, and two fare-classes were defined (business and economy). The analysis had revealed that the airfare elasticity of

passenger demand significantly varies among the different routes of the airline.

15. In May (2001), King Abdul-Aziz University, Seraj Y. Abed presented a study titled by: "An Econometric Analysis of International Air Travel Demand in Saudi Arabia". This study was for analyzing and forecasting international air travel market in Saudi Arabia using econometric model. In this study an attempt was made to develop several models for the air travel demand with different combination of explanatory variables, utilizing stepwise regression technique. The model with the two variables (i.e., total expenditures and population size) is the most appropriate model to represent the demand for international air travel in Saudi Arabia.

16. In May (2001), Edinburgh University, Scotland, Dargay, presented a study titled by: "The Determinants of the Demand for International Air Travel to and from the UK". This study used income, airfares, foreign trade, exchange rates and domestic price levels to find out the factors that were affecting air transport demand to and from UK. In the empirical part, this research used pooled time-series cross-section approach (panel data) with Fixed Effects Model specification, which allows them to have country specific evaluation. Panel method was used for leisure trips/business trips to 20 countries and non-UK residents' leisure/business trips to and from the UK. The authors preferred to use pooled time-series cross-section approach because of the limited number of observations for time-series model. Eventually, the results clarified that fares had a negative effect on passenger demand while income had a

positive effect on it in the UK air transport market. Moreover, income elasticity of UK international leisure air travelers was determined to be 0.43 in the short-run and 1.05 in the long-run, which means that one unit change in the income drives number of air travel to increase by 0.43 and 1.05 units respectively.

17. In (2000), King Abdulaziz University, Abdullah O. Ba-Fail presented a study titled by: "The Determinants of Domestic Air Travel Demands in Kingdom of Saudi Arabia". In this study, the author used passenger numbers as a dependent variable and non-oil gross domestic product, consumer price index, imports of goods and services, population size, total expenditures and the total consumption expenditure as explanatory variables. He used four different model specifications in order to see forecasting performance of each model. As a result, he found out that the model with population size and total expenditure was the best model to explain passenger demand for both domestic and international air transportation. It means that increasing population and expenditures drive the increase of international and domestic air transport demand. However, the model they built up had deficiencies. Despite the fact that they had considered different models, they did not try to create a model with more explanatory variables.

18. In (2000), Journal of Transport Management, Graham, presented a study titled by: "Demand Air for Leisure Air Travel and Limits to Growth". This study was on income elasticity of UK leisure air travel market. According to this study, air travel segments in UK are facing lower income elasticities than before. For instance, according to the outcome

of this study, income elasticities of international holidays decreased from 0.74 during period 1970-1988 to 0.55 during period 1984-1998. Also, R-square (explanatory power of model) of the elasticity models decreases, which means that the relationship between income and passenger numbers is diminishing. Accordingly, this lead the author to draw conclusion that the UK leisure air travel market is facing maturity.

19. In May (1996), University of British Columbia, Tae Hoon Oum presented a study titled by: "The Effect of Airlines Codesharing Agreements on Firm Conduct and International Air Fares". In this study a model based on the profit-maximising behavior of n firms in the market had been developed and applied to a panel data of 57 transpacific air routes for the 1982-1992 period. The model had been used to measure the effect of a codesharing agreement between non-leaders on the market leader's equilibrium price and passenger volume. It had been found that "complementary" codesharing between non-leaders rotates the market leader's supply relation curve clockwise (that is, it makes the leader behave more competitively), and moves its residual demand curve upward. As a result, a codesharing agreement increases the annual equilibrium quantity of the market leader by 10,052 passengers, while reducing the leader's equilibrium price by about $83 per passenger.

20. In (1989), Journal of Transportation Research, Fridström, presented a study titled by: "An Econometric Air Travel Demand Model for the Entire Conventional Domestic Network: The Case of Norway". This study uses pooled cross section time series data, and fares, travel time, income and population taken as independent variables in the model.

Fares and travel times variables were used for air travel and for the fastest surface transportation mode. The results were as expected. Population and income had positive demand elasticities while time fare and travel time had negative demand elasticities. Magnitudes for population and income were 1.46 for each variable. On the other hand, fare and travel time were respectively -1.23 and -0.94 for the long run.

Most of the previous studies were interested in determining the relation between maximum three variables, while in this research the relation between five variables were studied. Additionally, only five out of the twenty previous studies summarized in this research, used only the Fixed Effect Model of the Panel Regression Models, while in this research all four Panel Regression Models had been used for statistical analysis.

1.8. <u>RESEARCH PLAN:-</u>

The research contains five chapters:

The first chapter (Introduction), as introduced, includes: an introduction and specifications; that illustrates the research problem, importance, objectives and hypothesis, to determine the research methodology and overview the previous studies.

The second chapter (Literature Review), includes a brief background about the airlines in Sudan and the aircraft cost and implication for analysis.

The third chapter (Basic Methods of Regression Models & Panel Data Models), includes: approaches and definition of Regression analysis and panel data models that were applied in the research to determine the causes of insufficient Sudanese airlines domestic services and to estimate an

airline cost function.

The fourth chapter (Results & analysis), contains analysis of Sudanese domestic air transport data, by using Classical Normal Linear Regression Model (CNLRM) & Panel Regression Models. Also, to discuss the outcomes of the analysis and to study the possibility of using study techniques to estimate an equation of the airline data, so as to evaluate the data to obtain the results.

The fifth chapter (Conclusions & Recommendations), provides conclusions and recommendations regarding future of Sudan domestic air transport.

1.9. <u>RESEARCHER PUBLISHED PAPERS:-</u>

The researcher had published two papers from this study, titled as the following:

First Paper Title:

"Past, Present & Future of Airlines Domestic Services in Sudan". In (May – Jun. 2016), IOSR Journal of Mathematics (IOSR-JM). e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 3 Ver. IV, PP 74-85. (Sultan **et al.**, 2016).

Second Paper Title:

"Estimation of Sudanese Airlines Domestic Services Cost Function". In (May – Jun. 2016), IOSR Journal of Mathematics (IOSR-JM). e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 3 Ver. VI, PP 42-50. (Sultan **et al.**, 2016).

The next chapter includes a brief background about the airlines in Sudan, aircraft cost and implication for analysis.

CHAPTER TWO

LITERATURE REVIEW

2 - 0 INTRODUCTION.

2 - 1 SPECIFICATION.

2 - 2 THE ROLE OF AVIATION IN SUDAN
    DEVELOPMENT.

2 - 3 AIR TRANSPORTATION IN SUDAN.

2 - 4 ACHIEVING MORE EFFICIENT OPERATIONS.

2 - 5 FREEDOMES OF THE AIR.

2 - 6 TYPES OF AIRCRAFT TO SELECT.

2 - 7 TYPES OF AIRCRAFT OPERATIONS.

2 - 8   INTERNAL MARKET (DOMESTIC).

2 - 9 AIRLINES IN SUDAN.

2 - 10 AIRCRAFT COST.

2 - 11 IMPLICATION FOR ANALYSIS.

CHAPTER TWO

LITERATURE REVIEW

2. <u>INTRODUCTION:-</u>

Being one of the largest African countries in area, Sudan has always been in need for air transport, both for domestic and international links. Aviation industry in Sudan has received close attention and encouragement by successive governments, and Sudan Civil Aviation Authority has always been a technical, legislative and administratively, a priority institution. This institutional prominence has born fruition in establishing and managing a range of more than 40 airports and airstrips in towns and cities across the country , and in areas which otherwise are very difficult to access over land. The country has from the outset managed to plan and develop its aviation industry by associating and acquiring the membership and signing agreements for all matters, technical or legal, for the promotion, control, and coordination of aviation locally, regionally and internationally. At present , the nation locates seven international airports in the north, south, east and west of the country, with Khartoum Airport as the main country hub. It is possible to read within such a vocational configuration the strategic nature of an aviation system for Sudan, the region and beyond. Technically, aviation in Sudan has shown constant development and eagerness to adopt and employ the latest technologies for communication, control, and safety operations in the skies and on the land (D.G., C.A.A., 2002), (Sultan & Siddig, 2006). At present, the aviation scene is passing through an important transitional period, the main feature of

which are upgrading, renovation and new constructions. Khartoum New International Airport is an important project for future development (D.G., C.A.A., 2002).

Being the largest country in Africa, it becomes vitally important to link the different areas of the country with each other. Currently, there are several international airports in the country, in addition to a considerable number of landing strips.

2.1. SPECIFICATION:-

2.1.1 Aviation Growth:

Air traffic growth has averaged about 5% per year during the period 1980 to 2013, and it is continuing to experience the fastest growth among all modes of transport. If the strong growth in air travel continues, world air traffic volume may increase up to five-fold to twenty-fold by 2050 compared to the 1990 level and account for roughly two-thirds of global passenger-miles traveled. (S.A.C., 2012)

Given the strong growth in air travel and increasing concerns associated with effects of aviation emissions on the global atmosphere, the aviation industry is likely to face a significant environmental challenge in the near future. Current estimates show that global air traffic volume is growing so fast that the total aviation fuel consumption and subsequent aviation emissions' impacts on climate change will continue to grow despite future improvement in engine and airframe technologies and aircraft operations.

As a result, continuing growth in world population and (GDP) gross domestic product are expected to lead to a high growth in air travel

demand in the future.

Aviation has now become a major mode of transportation and integral part of the infrastructure of modern society. Currently, aircraft account for more than 10% of world's passenger miles travelled (Sultan, 2012).

Aviation directly impacts the global economy in the form of commercial passenger travel, freighter transport, and business travelers, involving the suppliers and operators of aircraft, component manufacturers, fuel suppliers, airports, and air navigation services providers. In 1994, the aviation sector accounted for 24 million jobs globally and financially provided $1,140 billion in annual gross output (S.A.C., 2012), (Sultan, 2012).

2.1.2. Past Situation of Sudan Aviation Field:

In mid-1991, scheduled domestic air services were provided by Sudan Airways, a government-owned enterprise operated by Sudan Airways Company. The company began its operations in 1947 as a government department. It has been operating commercially since the late 1960s, holding in effect a monopoly on domestic services. In 1991 Sudan Airways had scheduled flights from Khartoum to twenty other domestic airports, although it did not always adhere to its schedules. It also provided international services to several European countries, including Britain, Germany, Greece, and Italy. Regional flights were scheduled to North Africa and the Middle East as well as to Chad, Ethiopia, Kenya, Nigeria, and Uganda. Sudan Airways fleet in 1991 consisted of thirteen aircraft, including five Boeing 707s used in international flights, two Boeing 737s and two Boeing 727s employed in domestic and regional services, and four Fokker F-27s used for domestic flights (Director S.A.C. Planning, 2012).

Sixteen international airlines provided regular flights to Khartoum. The number of domestic and international passengers increased from about 478,000 in 1982 to about 485,000 in 1984. Air freight increased from 6 million tons per kilometer in 1982 to 7.7 million tons per kilometer in 1984. As compared with the previous years, in 1989 passenger traffic of Sudan Airways fell by 32% to 363,181 people, reducing the load factor to 34.9%. By contrast, freight volume increased by 63.7% to 12,317 tons. At the end of 1979, Sudan Airways had entered into a pooling agreement with Britain's Trade Wind Airways to furnish charter cargo services between that country and Khartoum under a subsidiary company, Sudan Air Cargo. A new cargo terminal was built at Khartoum Airport (D.G., C.A.A. 2002), (Director S.A.C. Planning, 2012), (S.C.A.A., 2012).

Sudan Airways' operations have generally shown losses, and in the early 1980s the corporation was reportedly receiving an annual government subsidy of about 500,000 Sudanese pounds. In 1987 the government proposed to privatize Sudan Airways, precipitating a heated controversy that ultimately led to a joint venture between the government and private interests. Like the railroads and river transport operators, however, Sudan Airways suffered from a shortage of skilled personnel, overstaffing, and lacked hard currency and credit for spare parts and proper maintenance (Sultan, 2012).

In the early 1980s, the country's civilian airports, with the exception of Khartoum International Airport and the airport at Juba, sometimes closed during rainy periods because of runway conditions. After the 1986 drought, which caused major problems at regional airports, the government launched a programme to improve runways, to be funded locally. Aeronautical communications and navigational aids were minimal and at

some airports relatively primitive. Only Khartoum International Airport was equipped with modern operational facilities, but by the early 1990s, Khartoum and seven other airports had paved runways. In the mid-1970s, International Development Aid (IDA) and the Saudi Development Fund agreed to make funds available for construction of new airports at Port Sudan and Wau, reconstruction and improvement of the airport at Malakal, and substantial upgrading of Juba Airport; these four airports accounted for almost half of the domestic traffic. Because the civil war resumed, improvements were made only at Port Sudan. Juba airport runways were rebuilt by a loan from the European Development Fund, but the control tower and navigational equipment remained incomplete (Sultan, 2012).

2.1.3. Present Situation of Sudan Aviation Field:

At present, the aviation scene is passing through an important transitional period, the main feature of which are upgrading, renovation and new constructions, Khartoum New International Airport is an important project for future development (D.G., C.A.A. 2002).

Being the largest country in Africa, it becomes vitally important to link the different areas of the country with each other. Currently , there are several international airports in the country , in addition to a considerable number of landing strips. Also, the following main projects have been processed: (D.G., C.A.A. 2002), (Director C.A.A. Planning, 2012)

1. A rehabilitation process to Khartoum International Airport and some of the other regional airports (El Obayed, Nyala, El Fashir, Dongola, Port Sudan, etc.) had been accomplished .

2. Technical and economic feasibility studies for the establishment of regional airports and landing strips , compatible with the international standards had been prepared. Moreover, the updating of the technical and economic feasibility studies for Khartoum New International Airport is now ready, and the execution of the project is now in progress.

3. Systems of communication, navigation and metrology have been modernized to secure air safety travels.

4. Several bilateral and multilateral agreements were conducted with some countries, such as Nigeria and South Africa, with respect to crossing the Sudanese airspace, and with some airlines with respect to having joint venture with Sudan Airways company.

5. Modern maintenance units have been established to meet the requirements of the new aircraft types.

6. Modern ground equipment for air handling modern aircraft types are introduced.

7. Training centers for Civil Aviation and Sudan Airways are also established, and a competent management of Technical Information System is also introduced.

8. Many chances have been granted to private sector companies to operate domestic and international flight services of transport, cargo and air handling.

Internationally, major changes in the Civil Aviation Regulations have been introduced to cope with globalization. Air transport is now in the process of liberalization , and fare pricings liberalization has been adopted in some regions in the world . Some airlines joints with others improve cost effectiveness and financial efficiency , while other public

and national carriers preferred privatization.

More than sixty five Sudanese airline companies were registered and approved by Sudan Civil Aviation Authority since Sudan independence in 1956. Unfortunately, now only 10% of them are active at a very low profile. Most of them used to operate more than three aircraft, but now, including Sudan Airways which is the first National Carrier, each one is operating one aircraft only!

2.2. <u>THE ROLE OF AVIATION IN SUDAN DEVELOPMENT:-</u>

The fallowing considerations are of paramount importance in a developing country such as Sudan: (Sultan, 2012)

- Contribution to Sudan international trade.
- Generation of direct and indirect employment in aviation related industries through a multiplier effect.
- Development of skills and creation of a pool of expertise in the fields of advanced and intermediate technology.
- Increase of government and private income flows.
- Additions to the gross block of Aviation institutions from internally generated funds.
- Impact made by airports on states and regional economics.
- Activation of passengers and freight.

2.3. <u>AIR TRANSPORTATION IN SUDAN:-</u>

The establishment of Air Transportation Company in accordance with the laws and regulations – Sudan Civil Aviation Authority (S.C.A.A.) Act 1999 as implemented in the International Civil Aviation Organization (ICAO) annexes and documents of Chicago Convention in 1944 are for the fallowing Objectives: (C.A.A., 2012)

- Carriage of passengers, cargo and mail by transportation.
- Self-reliability and protection of state security from foreign Interference.
- Training of national cadres to replace and take-over from foreigners i.e. Pilots, Engineers, Operation and Traffic Officers, Navigators and all relative aviation staff.
- To create clean and fair economical competitions between companies engaged in air transport.
- Development of air transport industry within Sudan as well as internationally.
- Adoption of the International Civil Aviation Organization (ICAO) and the International Airlines Transport Association (IATA) Laws, Regulations and Resolutions.
- To maintain safety in an orderly manner.
- To create opportunities for more employments.
- To create profitable markets domestically and internationally.
- To create new feasible flying routes and rehabilitate the old ones.
- To utilize the maximum capacity of aerodrome and airstrips.
- To participate in the development of all trade industries that are related to aviation.
- To participate in the acceleration of the trend of Sudan Economy.
- To develop tourism in Sudan and the neighboring countries.
- To import new funds to Sudan through import aircraft and handling equipment.

2.4. ACHIEVING MORE EFFICIENT OPERATIONS:-

      Once reaching the stage of a capable national carrier in operation and airworthiness (as designed by the "C.A.A. Sudan") and fulfilling the requirements of ICAO documents and annexes, the company shall enjoy all privileges of the permanent Air Operator Certificate (AOC):

- Flying routes and destinations according to Jeppesen and Sudan AIP on charter basis.
- Through the CAA-Sudan- Air Transportation may be granted the right to enjoy the privileges of the FREEDOMS of the air.
- This right is subject to the full satisfaction and acceptance of the Sudan Civil Aviation Authority.

2.5. FREEDOMS OF THE AIR:- (ICAO FAQ, 2011)

- FREEDOM 1: The right to fly over the territory of a state without landing.

- FREEDOM 2: The right to land in another state for non traffic purposes (repairs, refueling).

- FREEDOM 3: The right to fly into the territory of another state and discharge passengers, mail and cargo which is originated in the flag state of the carrier.

- FREEDOM 4: The right to take on passengers, mail and cargo in the territory of another state and transport them to the flag state of the carrier.

- FREEDOM 5: The right to fly into the territory of another state for the purpose of taking on or discharging passengers, mail and cargo destined for a third state.

- FREEDOM 6: The right to take on and or discharge passengers, mail and cargo destined for, or coming from, the territory of another state via the flag state of the carrier.

- FREEDOM 7: The right to establish a base in the territory of another state for the purpose of taking on and discharging passengers, mail and cargo destined for, or coming from, other states.

- FREEDOM 8: The right to fly into the territory of a state for the purpose of taking on and discharging passengers, mail and cargo destined for another location within the territory of that state. It's also known as (limited cabotage).

- FREEDOM 9: The right for a foreign carrier to take on and discharge passengers, mail and cargo destined for another location within the territory of a state without being a part of an international flight. It is also known as (full cabotage).

2.6. TYPES OF AIRCRAFT TO SELECT:-

For the selection of Aircraft types to be operated in Sudan, the following steps are stated here below: (S.C.A.A., 2012)

- The aircraft type, design shall be related to the conditions and serviceability of aerodrome (s) used.

- The capacity, payload of aircraft shall be considered in relation to the aerodrome serviceability.
- The availability of maintenance facility within Sudan or abroad shall be considered.
- The availability of crew experienced with the appropriate license endorsed with the type rating selected.
- The availability of experienced engineers appropriately licensed with the type rating selected.
- The availability of an experienced handling agent equipped with the appropriate handling equipment and appliances.
- The selected type to comply with Sudan C.A.A requirements in respect of aircraft airworthiness (Certificate of Airworthiness i.e. C. of A).
- Fuel and lubricants quantities and quality are available.

2.7. TYPES OF AIRCRAFT OPERATIONS:-

There are two distinct types of aircraft operations: Specialized operation with all cargo aircraft and Carriage of freight / passengers and freight aircraft "COMBI" are provided. The Certificate of Airworthiness (C of A). shall be category "COMBI" or Approved by CAA- Operations Directorate.

In Sudan scheduled international freighter air services are operated by Sudan Airways under bilateral agreements or special agreements with other states. Specialized Freight Charters are operated by several tramp charter operators that are approved by CAA- Air Transport directorate. Regular, Frequent Charters are operated by Saudi Airlines, Sudan Airways and some other approved CAA Airlines. Occasional Freighter Charters are

operated on some domestic routes (C.A.A., 2012).

2.8. <u>INTERNAL MARKET (DOMESTIC):-</u>

Due to the lack of paved safe roads and to insufficient modern transport vehicles to avail a continuous flow of transport, the need is now entirely dependent on air transport for everything. Since most of this volume of cargo consists of consumables, then this demand shall continue to persist as long as the safety situation remains unchanged. The air cargo movement include: Livestock, Meat, Fruit and vegetables, Relief goods and generals goods (C.A.A., 2012).

2.9. AIRLINES IN SUDAN:-

 The airlines currently operating in Sudan are listed in the table below:

Table (2.1): Airlines in Sudan:

AIRLINE	IATA	ICAO	CALLSIGN
Azza Transport	------	AZZ	AZZA TRANSPORT
Badr Airlines	J4	BDR	BADRAIR
Blue Bird Aviation	------	BLB	BLUEBIRD SUDAN
Dove Air Services	------	DOV	DOVAIR
Marsland Aviation	M7	MSL	MARSLANDAIR
Mid Airlines	7Y	NYL	NILE
Nova Airline	O9	NOV	NOVANILE
El Magal Aviation	------	MGL	ELMAGALAIR
Sudan Airways	SD	SUD	SUDANAIR
Tarco Airlines	------	TRC	TARCO AIR
Sun Air	1Y	SUN	SUN GROUB
Alfa Airlines	----	AAJ	ALFA AIR

Source: S.C.A.A.(2012). Sudan Civil Aviation Integrated Statistics.

The researcher collected some data from two airlines, namely Sudan Airways and Marsland Air. Others were reluctant to provide any actual genuine data! It took him a long time asking each airline management , but … at last the rest of the data were hardly collected from Sudan Civil Aviation Authority , Department of Air transport, which is supposed to be the custodian data bank of aviation activities and resources . Although more than twenty Sudanese airlines were declared active and enlisted by Sudan Civil Aviation Authorities, the researcher was astonishingly informed that there were no data recorded or any useful records kept anywhere!

It is worthy to remarkably note that some airlines lost their fleet either due to bankruptcy or accidents, and thereby, were renamed and resumed operations as new firms!

In this research ten years annual data of total domestic passengers and cargo, from 2004 to 2013 , are statistically analyzed to determine a cost function for six airlines, namely Sudan Airways, Marsland Aviation, Badr Airlines, Mid Airlines and Sun Air as follows:-

2.9.1. Sudan Airways: (S.A.C., 2012)

(IATA Code: SD | ICAO Code: SUD | Call sign: SUDANAIR):

Sudan Airways is a national airline in Sudan, headquartered in Khartoum since 1946 .The carrier is a member of the International Air Transport Association (IATA), of the Arab Air Carriers Organization since 1965, and of the African Airlines Association since 1968, becoming a founding member along with ten companies.

In 2007, the Sudanese government privatized the airline & sold it to a Kuwaiti private group for four years. The Airline has been included in the list of airlines banned in the European Union since March, 2010. As of December 2011, the company is reowned by the Government of Sudan, and has 1,700 employees.

As of July 2011, the airline serves twenty four destinations in Africa , ten of them within Sudan and the Middle East.

Sudan Airways fleet consisted of the following aircraft, with an average age of seventeen years:

Table (2.2): Sudan Airways Fleet (Seventeen Years Ago):

Aircraft	In Fleet
Airbus A300B4 /-600R	2
Airbus A320-200	2
Boeing 737-500	1
Fokker 50	3
Yakovlev Yak-42D	1
Total	**9**

Source: S.A.C. Planning Directorate (2012).

Sorrowfully, Sudan Airways currently operates only one aircraft due to its privatization and selling 70% of its shares in 2007 to a non-aviation- minded foreign Arab businessman, who had paralyzed this famous first aviation pioneer international Airline in Africa and all Arab countries and the Middle East. Due to his absolute failure to run the Airline, he had to sell its stake back to Sudan in 2011.

As of June, 2014, Sudan Airways active fleet consists of the following aircraft:

Table (2.3): Sudan Airways Active Fleet:

Aircraft	In Fleet
Airbus A300B4-600R	2
Airbus A320-200	2
Fokker 50	3
Total	**7**

Source: S.A.C. Planning Directorate.

The company has flown more than thirty equipment of various

types through its long history.

During the last forty years the airline experienced 21 accidents/incidents, 7 of which lead to fatalities …The worst accident was in July, 2003 near Port Sudan, when 117 were killed.

In this research ten years annual data of total domestic passengers and cargo, from 2004 to 2013, are statistically analyzed to determine a cost function.

2.9.2. Marsland Aviation:-

(IATA Code: M7 | ICAO Code: MSL | Call sign: MARSLAND AIR):

Marsland Aviation Company was incorporated in 2001, and commenced operations the same year with flights to Western Sudan (Elgenaina), initially with one Antonov 24, while currently it operates flights to 10 destinations (8 domestic) (G.M. Marsland Airlines Statistic Section, 2012).

Marsland Aviation fleet consist of the following aircraft: (as of December 2012):

Table (2.4): Marsland Aviation Active Fleet:

Aircraft	In Fleet
Antonov AN-24	1
Boeing 737-200	1
Boeing 737-500	2
Yakovlev Yak-42D	1
Total	**5**

Source: G.M. Marsland Airlines Statistic Section (2012).

But now, Marsland Air stopped activities by the end of 2013.

In this research ten years annual data of total domestic passengers and cargo, from 2004 to 2013, are statistically analyzed to determine a cost function.

2.9.3. Badr Airlines (Sarit Airlines):

(IATA Code: J4 | ICAO Code: BDR | Callsign: BADR AIR):

The company was founded as Sarit Air Lines in 1997 and was renamed in 2004, due to numerous fatal air crashes, when the Antonov 12 was carrying food and money on a routing flight from Khartoum, and after obtaining permission to land, it exploded in the air about 6km from Wau Air port (2003), in which proceeding investigations showed that the airline disregarded many CAA [clarification needed] safety regulations, including but not limited, to overloading of the aircraft and bribing of its air crew to overlook overloading of the aircraft !. (S.C.A.A., 2012).

Badr Air lines fleet consist of the following aircraft:

Table (2.5): Badr Air Active Fleet:

Aircraft	In Fleet
Boeing 737-300	2
Boeing 737-500	2
Ilyushin IL-76 TD	2
Antonov AN-26B-100	1
Antonov AN-74 T 200	1
Antonov AN-74 TK-100	1
Antonov 2	1
Antonov 28	1
Total	**11**

Source: S.C.A.A. Sudan Civil Aviation Integrated Statistics, (2012).

In this research ten years annual data of total domestic passengers and cargo, from 2004 to 2013, are statistically analyzed to determine a cost function.

2.9.4. Mid Airlines

(IATA Code: 7Y | ICAO Code: NYL | Callsign: NILE):

Mid Airlines is an airline based in Khartoum, Sudan. It operates domestic passenger services. Its main base is Khartoum. The airline was established in 2002 and started operations in May 2003.Mid Airlines operates scheduled domestic destinations to Khartoum, Rumbek, and Port Sudan (as of January 2005) (S.C.A.A, 2012).

The Mid Airlines fleet includes the following aircraft (as of August 2013):

Table (2.6): Mid Airlines Active Fleet:

Aircraft	In Fleet
FOKKER F-50	2
Total	**2**

Source: S.C.A.A. Sudan Civil Aviation Integrated Statistics, (2012).

In this research eight years annual data of total domestic passengers and cargo, from 2004 to 2011, are statistically analyzed to determine a cost function up to the date of that operation in 2011.

2.9.5. Sun Air (Air West):

2.9.5.1. Air West: (S.C.A.A, 2012)

(IATA Code: --- | ICAO Code: AWZ | Callsign: ----):

Air west is an airline based in Khartoum, Sudan. It operates domestic passenger services and international cargo charters. Its main base is Khartoum International Airport, with a hub at Sharjah International Airport. The airline is on the List of air carriers banned in the European Union.

The airline was established in April 1992 and started operations in October 1992. Air west was the first Sudanese airline to start to fly between Khartoum and Rumbek, where the Sudan People's Liberation Movement (SPLM) has its headquarters.

On February 3, 2005, an Air West Ilyushin IL-76 aircraft, owned by East-West Airlines based in Sharjah, left Sharjah with 46 tons of humanitarian cargo for Nyala, Darfur province in Sudan, with a planned enroute stop at Khartoum. The aircraft crashed about 15 km from Khartoum killing all seven crew; six of the crew members were Ukrainian and the pilot of the flight was Sudanese.

Air West fleet consists of the following aircraft (as of 4 April 2008):

Table (2.7): Air West Fleet:

Aircraft	In Fleet
Antonov AN-12	2
Antonov AN-24RV	2
Antonov AN-26	1
Antonov AN-28	2
Antonov AN-32	1
Yakovlev Yak-42D	1
Total	**9**

Source: S.C.A.A. Sudan Civil Aviation Integrated Statistics, (2012).

2.9.5.2. Sun Air: (S.C.A.A, 2012)

(IATA Code: 1Y | ICAO Code: SNR | Callsign: SUN GROUP)

The company was founded as Air West Airlines and was renamed in 2008, due to fatal air crashes.

Sun Air is a private airline based in Khartoum, Sudan. The airline is owned by a Sudanese Owner and his partners of Sun Air group. Sun Air was established as a company in accordance with the Companies Act of 1925 and started its operations on June 2008 with a fleet of one Airbus A310-300 aircraft, and three Boeing 737-200. The airline currently operates on domestic routes and international routes. The destinations covered by Sun Air are Sharjah, Juba, Nyala, El fashir, Elgenana, Port Sudan and shortly Jeddah, Cairo, Damascus, and Asmara.

Sun Air currently operates with a brand new fleet of 5 Boeing 737 and one Airbus A310 aircraft. It was the first private airline in Sudan to operate with all new aircraft. Sun Air is also the first private airline to order the Airbus A310 and Boeing 737-400. It placed orders for 2 Boeing 767

aircraft. In a deal valued at over $15 million in June 15, 2011. Delivery of the Boeing 767 is due to start in late 2010, followed by the A300s in 2013. Sun Air fleet consists of the following aircraft (as of 5 July 2011):

Table (2.8): Sun Air Fleet:

Aircraft	In Fleet
Boeing 737-200	2
Boeing 737-300	1
Total	**3**

Source: S.C.A.A. Sudan Civil Aviation Integrated Statistics.

In this research ten years annual data of total domestic passengers and cargo, from 2004 to 2013, are statistically analyzed to determine a cost function.

2.9.6. Nova Air:

(IATA Code: O9 | ICAO Code: NOV | Callsign: NOVANILE):

Nova Airways is an airline based in Khartoum, Sudan and established in the year 2000. Nova Airways fleet includes the following aircraft (as of 4 July 2009):

Table (2.9): Nova Air Fleet:

Aircraft	In Fleet
Bombardier CRJ200ER	4
FOKKER F-50	1
Boeing 737-500	1
Total	**6**

Source: S.C.A.A. Sudan Civil Aviation Integrated Statistics.

In this research ten years annual data of total domestic passengers and cargo, from 2004 to 2013, are statistically analyzed to determine a cost function.

It is noted that all airlines were established by certain aircraft fleet as presented. But, in fact, the actual working fleet may be different, because most of airlines operated some aircraft from others as needed, (Routing policy).

2.10. AIRCRAFT COST:-

Direct Operating Cost (DOC) and price are the two major elements of aircraft cost. While price is a one-time cost for aircraft acquisition, DOC is a recurring cost over the lifetime of an airplane.

However, in practice, both elements appear together as part of aircraft operating cost, DOC and investment, as the value of an airplane is depreciated over a large fraction of its lifetime). (Rajiv & Johnston, 1993), (Ptetronas Marketing, 2012).

2.10.1. Cost Categories:

The input cost categories listed below are subsets of traditional broad economic categories and are based on the operating expense categories of the Uniform System of Accounts used by the carriers to file financial data with Civil Aeronautics Board (CAB) and Department of Transportation (DOT). The dependent variables are physical measures of the inputs when these are available and deflated cost measures otherwise: (Rajiv & Johnston, 1993), (Ptetronas Marketing, 2012).

1. Fuel, gallons of jet fuel and oils.

2. Flying operations labor, hours of labor of flight crews, including pilots, copilots, navigators, and flight engineers.

3. Passenger service labor, hour of labor of flight attendants.

4. Aircraft traffic servicing labor, hours of labor of ground personnel servicing aircraft and handling passengers at gates, baggage, and cargo.

5. Promotions and sales labor, hours of labor of reservations and sales agents primarily, but also of personnel involved in advertising and publicity.

6. Maintenance labor, hours of labor involved in maintenance of flight equipment and ground property and equipment.

7. Maintenance materials and overhead, total cost of maintenance of property and equipment, deflated by Producer Price Index for fabricated metals.

8. General overheads, total expenses corresponding to supplies, general and administrative personnel, utilities, insurance, and communications.

9. Ground property and equipment, flows of service from ground property and equipment, calculated with the method developed by Christensen and Jorgenson (1969) and including landing fees deflated by the Air Transport Association cost index for landing fees and rental expenses for ground property and equipment deflated by Producer Price Index for fixed nonresidential structures.

10. Flight equipment, flows of service from flight equipment (airframes, aircraft engines, avionics, etc.), calculated by imputing fair market rental values deflated by Producer Price Index for fixed-wing aircraft to owned and leased aircraft by aircraft categories.

2.11. <u>IMPLICATIONS FOR ANALYSIS:-</u>

- Dichotomy of airline demand and supply complicates many facet of airline economic analysis.

- Difficulty, in theory, to answer seemingly "simple" economic questions, for example:

  a) Because we cannot quantify "supply" to an individual market, we cannot determine if the market is in "equilibrium".

  b) Cannot determine if the airline's service to that market is "Profitable", or whether fares are "too high" or "too low".

  c) Serious difficulties in proving predatory pricing against low-fare new entrants, given joint supply of seats to multiple markets and inability to isolate costs of serving each market.

- In practice, assumptions about cost and revenue allocation are required:

Estimates of flight and/or route profitability are open to question. The airline industry worldwide has been going through some significant changes. These were mainly brought about by the fact that historically airlines were being run by governments, but due to a lack of funding they have been privatized (AFRAA, 1999).

Modernization of fleet has been forced on the airlines by the stricter noise and safety regulations, and the need to improve Africa's air transport services and industry. Because of the above financial and management problems, privatization of airlines has been adopted but with only a few successes. There are few investors interested in airlines, which continually operate at a loss.

The next chapter includes approaches and definition of Regression

analysis and panel data models that are applied in the research for determining the causes of insufficient Sudanese airlines domestic services and estimation of an airline cost function.

CHAPTER THREE

BASIC METHODS OF REGRESSION MODELS & PANEL DATA MODELS

3 - 0   INTRODUCTION.

3 - 1   DATA TYPES.

3 - 2   THE NATURE OF REGRESSION ANALYSIS.

3 - 3   LINEAR REGRESSION MODELS.

3 - 4   A MEASURE OF GOODNESS OF FIT.

3 - 5   MULTICOLLINEARITY.

3 - 6   FEATURES OF A GOOD REGRESSION MODEL.

3 - 7   RESIDUAL TESTS.

3 - 8   NORMALITY TESTS.

3 - 9   HETEROSCEDASTICITY.

3 - 10 AUTOCORRELATION.

3- 11 MEASURES OF FORECAST ACCURACY.

3- 12 DUMMY VARIABLE REGRESSION MODELS.

3 - 13 PANEL DATA.

3 - 14 PANEL DATA MODELS.

3 - 15 SIMPLE EXPONENTIAL SMOOTHING.

CHAPTER THREE

BASIC METHODS OF REGRESSION MODELS & PANEL DATA MODELS

3. INTRODUCTION:-

Regression analysis is concerned with the study of the dependence of one variable, the dependent variable, on one or more other variables, the explanatory variables.

The objective of such analysis is to estimate and predict the mean or average value of the dependent variable on the basis of the known or fixed values of the explanatory variables.

In practice, the success of regression analysis depends on the availability of the appropriate data.

3.1. DATA TYPES:-

The data can be classified to the following types: (Gujarati & Porte, 2009), (Michael, 1978).

- Time Series Data:

A time series data is a set of observations value that a variable takes at different times, whether may be collected at regular time intervals such as daily, weekly, monthly, quarterly, annually, quinquennially (every five years), decennially (every ten years).

- Cross-Section Data:

     Are data of one or more variables collected at the same point in time.

- Pooled Data:

     In pooled; or combined, data are elements of both time series and cross-section data.

- Panel Data:

     This is a special type of pooled data in which the same cross-sectional unit is surveyed over time.

3.2. THE NATURE OF REGRESSION ANALYSIS:- (Rawlings et al., 1932), (Gujarati & Porte, 2009)

Regression analysis is concerned with the study of dependence of one variable, the dependent variable, on one or more other variables, the explanatory variables, with a view to estimating or predicting the (population) mean or average value of the former in terms of known or fixed values of the latter.

The key concept underlying regression analysis is the concept of the conditional expectation function (CEF), or population regression function (PRF). The PRF is an idealized concept, since, in practice, one rarely has access to the entire population of interest. Usually, one has a sample of observations from the population. Therefore, one uses the stochastic sample regression function (SRF) to estimate the

PRF.

There are simpler possible regression analyses for PRF, namely as follow:

- Bivariate regression, two-variable, regression in which the dependent (the regressand) is related to a single explanatory variable (the regressor).

$$Y_t = \beta_1 + \beta_2 X_t + u_t \qquad (3.2.1)$$

- Multiple regression analysis, in which the regressand is related to two or more regressors, is in many ways a logical extension of the two-variable case.

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t \qquad (3.2.2)$$

Where: $\boldsymbol{\beta_1}$, $\boldsymbol{\beta_2}$ and $\boldsymbol{\beta_3}$ are unknown, but fixed parameters known as the regression coefficients,

$\boldsymbol{\beta_1}$: Intercept: It gives the mean or average effect on Y of all variables excluded from the model, although mechanical interpretation is the average value of Y when X's are set equal to zero.

$\boldsymbol{\beta_2}, \boldsymbol{\beta_3}$: Slope coefficients: Are also called the partial regression coefficients.

Y_t : Regressand.

X_t: Regressor.

u_t: Random Disturbance Term.

The stochastic disturbance term plays a critical role in estimating the PRF.

The sample regression line which represents the SRF written as follows:

$$\widehat{Y}_t = \widehat{\beta}_1 + \widehat{\beta}_2 X_{2t} + \widehat{\beta}_3 X_{3t} + \widehat{u}_t \qquad (3.2.3)$$

Where:

$\widehat{Y}_t$ = Estimator of $E(Y|X_{2t}, X_{3t})$.

$\widehat{\beta}_1$ = Estimator of β_1.

$\widehat{\beta}_2$ = Estimator of β_2.

$\widehat{\beta}_3$ = Estimator of β_3.

$\widehat{u}_t$ = Estimator of u_t. It denotes the sample residual term.

3.2.1. OLS Estimation:

To obtain the OLS estimate of β, let us first, write the k-variable sample regression (SRF):

$$Y_t = \widehat{\beta}_1 + \widehat{\beta}_2 X_{2t} + \widehat{\beta}_3 X_{3t} + \cdots + \widehat{\beta}_K X_{Kt} + \widehat{u}_t \quad (3.2.4)$$

which can be written more compactly in matrix notation as:

$$y = \widehat{X}\beta + \widehat{u} \qquad (3.2.5)$$

and in matrix form as:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \vdots \\ \widehat{\beta}_n \end{bmatrix} + \begin{bmatrix} \widehat{u}_1 \\ \widehat{u}_2 \\ \vdots \\ \widehat{u}_n \end{bmatrix} \qquad (3.2.6)$$

As in the two- and three-variable models, in the k-variable case, the OLS estimators are obtained by minimizing

$$\sum \widehat{u}_t^2 = \sum (Y_t - \widehat{\beta}_1 - \widehat{\beta}_2 X_{2t} - \widehat{\beta}_3 X_{3t} - \cdots - \widehat{\beta}_K X_{Kt})^2 \qquad (3.2.7)$$

Where $\sum \widehat{u}_t^2$ is the residual sum of squares (RSS). In matrix notation, this amounts to minimizing $\acute{\widehat{u}}\widehat{u}$ since

$$\acute{\widehat{u}}\widehat{u} = [\widehat{u}_1 \quad \widehat{u}_2 \quad \cdots \quad \widehat{u}_n] \begin{bmatrix} \widehat{u}_1 \\ \widehat{u}_2 \\ \vdots \\ \widehat{u}_n \end{bmatrix} = \widehat{u}_1^2 + \widehat{u}_2^2 + \cdots + \widehat{u}_n^2 = \sum \widehat{u}_t^2 \quad (3.2.8)$$

$$\widehat{u} = y - X\widehat{\beta} \qquad\qquad (3.2.9)$$

$$\acute{\widehat{u}}\widehat{u} = (y - X\widehat{\beta})(y - X\widehat{\beta}) \qquad (3.2.10)$$

$$\acute{\widehat{u}}\widehat{u} = \acute{y}y - 2\widehat{\beta}\acute{X}y + \widehat{\beta}\acute{X}X\widehat{\beta} \qquad (3.2.11)$$

This process yields k simultaneous equations in k unknowns, the normal equations of the least-squares theory. These equations are follows:

$$n\widehat{\beta}_1 + \widehat{\beta}_2 \sum X_{2t} + \widehat{\beta}_3 \sum X_{3t} + \cdots + \widehat{\beta}_K \sum X_{Kt} = \sum Y_t \qquad (3.2.12)$$

$$\widehat{\beta}_1 \sum X_{2t} + \widehat{\beta}_2 \sum X_{2t}^2 + \widehat{\beta}_3 \sum X_{2t}X_{3t} + \cdots + \widehat{\beta}_K \sum X_{2t}X_{Kt} = \sum X_{2t}Y_t$$

$$\widehat{\beta}_1 \sum X_{3t} + \widehat{\beta}_2 \sum X_{3t}X_{2t} + \widehat{\beta}_3 \sum X_{3t}^2 + \cdots + \widehat{\beta}_K \sum X_{3t}X_{Kt} = \sum X_{3t}Y_t$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\widehat{\beta}_1 \sum X_{kt} + \widehat{\beta}_2 \sum X_{kt}X_{2t} + \widehat{\beta}_3 \sum X_{kt}X_{3t} + \cdots + \widehat{\beta}_K \sum X_{kt}^2 = \sum X_{kt}Y_t$$

In matrix form, Eq (3.2.12) can be represented as:

$$
\begin{bmatrix}
n & \sum X_{2t} & \sum X_{3t} & \cdots & \sum X_{Kt} \\
\sum X_{2t} & \sum X_{2t}^2 & \sum X_{2t}X_{3t} & & \sum X_{2t}X_{Kt} \\
\sum X_{3t} & \sum X_{3t}X_{2t} & \sum X_{3t}^2 & & \sum X_{3t}X_{Kt} \\
\vdots & \vdots & \vdots & & \vdots \\
\sum X_{kt} & \sum X_{kt}X_{2t} & \sum X_{kt}X_{3t} & & \sum X_{kt}^2
\end{bmatrix}
\begin{bmatrix}
\widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \widehat{\beta}_3 \\ \vdots \\ \widehat{\beta}_K
\end{bmatrix}
=
\begin{bmatrix}
1 & 1 & 1 & 1 \\
X_{21} & X_{22} & \cdots & X_{2n} \\
X_{31} & X_{32} & \cdots & X_{3n} \\
\vdots & \vdots & \vdots & \vdots \\
X_{k1} & X_{k2} & \cdots & X_{kn}
\end{bmatrix}
\begin{bmatrix}
Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n
\end{bmatrix}
\quad (3.2.13)
$$

$$(\acute{X}X)\widehat{\beta} = \acute{X}y \qquad (3.2.14)$$

$$\widehat{\beta} = (\acute{X}X)^{-1}\acute{X}y \qquad (3.2.15)$$

3.3. LINEAR REGRESSION MODELS:-

3.3.1. Classical Linear Regression Model (CLRM): (Greene, 1951), (Gujarati & Porte, 2009)

The Gaussian, standard, or Classical Linear Regression Model, which is the cornerstone of most econometric theory, makes seven assumptions, that are applied on the regression function of four variables, as the researcher concentrates in this study.

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + u_t \qquad (3.3.1)$$

Assumptions:

- Linear regression model, or linear in the parameters.

- Fixed X value independent of the error term. Here, this means we require zero covariance between u_{it} and each X variable. (Davidson, 2000)

$$Cov(u_t, X_{2t}) = Cov(u_t, X_{3t}) = Cov(u_t, X_{4t}) = Cov(u_t, X_{5t}) = 0$$

- Zero mean value of disturbance u_t. (Malinvaud, 1966).

$$E(u_t | X_{2t}, X_{3t}, X_{4t}, X_{5t}) = 0 \quad for\ each\ t$$

- Homoscedasticity or constant variance of u_t.

$$var(u_t) = \sigma^2$$

- No autocorrelation, or serial correlation, between the disturbances.

$$Cov(u_t, u_s) = 0 \quad t \neq s$$

- The number of observations n must be greater than the number of parameters to be estimated.

- There must be variation in the values of X variables.

- No exact collinearity (no linear relationship) between X variables.

- There is no specification bias. The model is correctly specified.

The classical theory of statistical inference consists of two branches, namely, estimation and hypothesis testing. Using the method of Ordinary Least Square (OLS), we were able to estimate the parameters and σ^2. Under the assumptions of CLRM, that enables us to show that the estimators satisfy several desirable statistical properties (BLUE property).

But this is half of the battle; and hypothesis testing is the other battle. So, since the method of OLS does not make any assumption about the probabilistic nature of u_t, it is of little help for the purpose of drawing inferences, the Gauss-Markov theorem notwithstanding. This void can be filled; if the regression context is that it is usually assumed that the disturbances term is normal distribution.

3.3.2. Classical Normal Linear Regression Model (CNLRM):

Adding the normality assumption for u_t to the assumptions of CLRM, we obtain what is known as Classical Normal Linear Regression Model (CNLRM). (Greene, 1951),

(Gujarati & Porte, 2009).

The CNLRM, assume that each u_{it} is distributed normally with:

Mean: $E(u_t) = 0$

Variance: $E[u_t - E(u_t)]^2 = E(u_t^2) = \sigma^2$

Covariance: $E(u_t u_s) = 0 \qquad t \neq s$

Zero covariance means that u_t and u_s are not only uncorrelated, but are also, independently distributed (NID).

$$u_t \sim NID(0, \sigma^2)$$

3.3.2.1. Normality Assumption:

Employ the normality assumption for the following reasons:

- The central limit theorem (CLT) of statistics, shows that if there are a large number of independent and identically distributed random variables, then with a few exceptions, the distribution of their sum tends to a normal distribution, as the number of such variables increases indefinitely (Ross, 2000). It is the CLT that provides a theoretical justification for assumption of normality of u_t.

- A variant of the CLT states that, even if the number of variables is not very large, or if these variables are not

strictly independent, their sum may still be normally distributed (Cramer, 1946).

- With the normality assumption, the probability distributions of OLS estimators can be easily derived, because, one property of the normal distribution is that any linear function of normally distributed variables is itself normally distributed. Therefore, if u_t are normally distributed, the estimated coefficients $\hat{\beta}_t$ make our task of hypothesis testing very straightforward.

- The normal distribution is comparatively a simple distribution involving only two parameters (mean and variance); its theoretical properties have been extensively studied in mathematical statistics.

- If the sample size is small or finite, the normality assumption assumes a critical role. Beside it helps us to derive the exact probability distributions of OLS estimators it enables us to use the t, F and χ^2 statistical tests of regression models.

- In large samples, t and F statistics have approximately the t and F probability distributions; so that tests are based on the assumption that the error term is normally distributed & can still be applied validly (Heij **et al.**, 2004). Therefore, the normality assumption may not be very crucial in large data sets.

The researcher preferred to use the framework of Classical Normal Linear Regression model (CNLRM) that extended to multiple regression models with normality assumption of disturbance term.

3.3.3. Properties of Least-Squares Estimators: The Gauss-Markov Theorem: (Rawlings et al., 1932)

Given the assumptions of the CLRM, the least-squares estimators, in the class of unbiased linear estimators, have minimum variance, that is, they are the best linear unbiased estimator (BLUE). As the following properties estimators:

- They are linear, that is linear function of random variable, such as the dependent variable Y in the regression model.
- They are unbiased.
- They have minimum variance. Combined with 1, this means that they are minimum variance unbiased, or efficient estimators.
- They have consistency; that is, as the sample size increases indefinitely, the estimators converge to their true population values.
- The coefficients estimators are normally distributed (under the normality assumption).

$$\widehat{\boldsymbol{\beta}}_t \sim \boldsymbol{N}(\boldsymbol{\beta}_t, \sigma^2_{\widehat{\beta}_t})$$

3.4. A MEASURE OF GOODNESS OF FIT:- (Greene, 1951)

The coefficient of determination R^2 (multiple regressions) is a summary measure that tells how the sample regression line fits the data. Verbally, R^2 measures the proportion or percentage of the total variation in Y explained by the regression model.

Two properties of R^2 may be noted: (Kennedy, 1981)

1. It is nonnegative quantity.
2. Its limits are $0 \leq R^2 \leq 1$.

An R^2 of 1 means a perfect fit, that is, $\hat{Y}_t = Y_t$ for each i. on the other hand, an R^2 of zero means that there is no relationship between regressand and the regressor whatsoever (the estimated slope coefficients are equal zero). In this case, the best prediction of any Y value is simply its mean value ($\hat{Y}_t = \hat{\beta}_1 = \bar{Y}$). In this situation, therefore, the regression line will be horizontal to the X axis.

3.5. MULTICOLLINEARITY:-

The term multicollinearity is due to Frisch (1934). Originally it meant the existence of a "perfect," or exact, linear relationship among some or all explanatory variables of a

regression model. For the k-variable regression involving explanatory variable X_1, X_2, . . . , X_k (where $X_1 = 1$ for all observations to allow for the intercept term), an exact linear relationship is said to exist if the following condition is satisfied:

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_K X_K = 0 \qquad (3.5.1)$$

Where $\lambda_1, \lambda_2, \ldots, \lambda_K$ are constants, such that not all of them are zero, simultaneously.

The fallowing equation shows that the case where the X variables are intercorrelated, but not perfectly so, as follows:

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_2 X_K + v_t = 0 \qquad (3.5.2)$$

Where v_t is a stochastic error term.

If multicollinearity is perfect in the sense of Eq (3.5.1), the regression coefficients of the X variables are indeterminate and their standard errors are infinite. If multicollinearity is less than perfect, as in Eq (3.5.2), the regression coefficients, although determinate, possess large standard errors (in relation to the coefficients themselves), which means the coefficients cannot be estimated with great precision or accuracy.

3.5.1. Detection of Multicollinearity:

There are some rules of thumb, some informal and some formal, but the rules of thumb are all the same. We now consider some of these rules:

1. High R^2 but few significant t ratios:

   The clearest sign of multicollinearity is when R^2 is very high, but none or very few of the regression coefficients is statistically significant on the basis of the conventional t test. This case is, extreme (Gujarati & Porter, 2009).

   The researcher uses this rule to detect multicollinearity.

2. High pair-wise correlations among regressors:

   In models involving just two explanatory variables, a fairly good idea of collinearity can be obtained by examining the zero-order, or simple, correlation coefficient between the two variables. If this correlation is high, multicollinearity is generally the culprit (Gujarati & Porter, 2009).

3. Examination of partial correlations:

   However, the zero-order correlation coefficients can be misleading in models involving more

than two X variables, since it is possible to have low zero-order correlations and yet find high multicollinearity. In situations like these, one may need to examine the partial correlation coefficients. If R^2 is high, but the partial correlations are low, multicollinearity is a possibility. Here one or more variables may be superfluous (Farrar & Glauber, 1967). But if R^2 is high and the partial correlations are also high, multicollinearity may not be readily detectable (O'Hagan & McCabe, 1975), (Kumar, 1975).

4. Auxiliary regressions:

Therefore, one may regress each of the X_i variables on the remaining X variables in the model and find out the corresponding coefficients of determination R_i^2. A high R_i^2 would suggest that X_i is highly correlated with the rest of the X's. Thus, one may drop that X_i from the model, provided it does not lead to serious specification bias (Hill **et al.**, 1982).

5. Tolerance and variance inflation factor (TOL & VIF).

As R_j^2, the coefficient of determination in the regression of regressor X_j on the remaining regressors in the model, increases toward unity, that is, as the collinearity of X_j with the other regressors

increases,VIF also increases and in the limit it can be infinite.

Some authors, therefore, use the VIF as an indicator of multicollinearity. The larger the value of VIF_j, the more "troublesome" or collinear the variable X_j. As a rule of thumb, if the VIF of a variable exceeds 10, which will happen if R_j^2 exceeds 0.90, that variable is said be highly collinear (Kleinbaum **et al.**, 1988).

One could use TOL_j as a measure of multicollinearity in view of its intimate connection with VIF_j . The closer TOL_j is to zero, the greater the degree of collinearity of that variable with the other regressors. On the other hand, the closer TOL_j is to 1, the greater the evidence that X_j is not collinear with the other regressors.

3.5.2. Remedial Measures:

3.5.2.1. Do Nothing:

The "do nothing" school of thought is expressed by Blanchard as follows (Blanchard, 1967):

When students run their first ordinary least squares (OLS) regression, the first problem that they usually encounter is that of multicollinearity. Many of them conclude that there is

something wrong with OLS; some resort to new and often creative techniques to get around the problem. But, we tell them, this is wrong. Multicollinearity is God's will, not a problem with OLS or statistical technique in general.

3.5.2.2. Rule of Thumb Procedures: (Gujarati & Porter, 2009)

1. A priori information.
2. Combining cross-sectional and time series data (pooling the data).
3. Dropping a variable(s) and specification bias.
4. Transformation of variables.
5. Additional or new data.
6. Other methods of remedying multicollinearity.

The researcher uses the procedures of dropping a variables and specification bias in this study.

3.6. FEATURES OF A GOOD REGRESSION MODEL:-

The features of a good regression are represented as follow: (Rawlings **et al.**, 1932), (Greene, 1951), (Maddala, 1992), (Gujarati & Porter, 2009).

Feature (1):

Regression line must be fitted to data strongly. Value of R-square should be more than 60%, because the higher R-

square value; better is the model or model fitted.

Feature (2):

Most of explanatory variables (at least 50%) should individually be significant to explain dependent variables. Here, t-test should be performed.

Feature (3):

Explanatory variables should be jointly significant to explain dependent variables. Here, F-test should be performed.

Feature (4):

Residuals of the model have no serial correlation, no heteroscedasticity, and are normally distributed.

3.7. <u>RESIDUAL TESTS:-</u> (Greene, 1951), (Maddala, 1992)

- Normality:

That means the normality of the disturbance term u_i , note that the t and F tests used before require that the error term follows the normal distribution.

- Heteroscedasticity:

This is a very general test for non constancy of the variance of the residuals. If the residuals have non constant variance then ordinary least squares is not the best estimation technique – weighted least squares should be used instead.

Heteroscedasticity test is Breusch-Pagan-Godfrey, which tests the null hypothesis; the residuals are not heteroscedastic, which is homoscedastic; while the alternative hypothesis; the residuals are heteroscedastic.

- ▪ Serial Correlation:

Serial correlation (auto- correlation) is statistically a term used to describe the situation, when the residual is correlated with lagged value of itself. Serial correlation can be formed in the model by incorrect model specifications, or omitted variables, or incorrect functional form or by incorrectly transformed data. An approach of detecting serial correlation is Breusch-Godfrey Serial Correlation LM test, which tests the null hypothesis; the residuals are not serial correlated, which is not auto correlated; while the alternative hypothesis; the residuals are serial correlated.

3.8. <u>NORMALITY TESTS</u>:- (Gujarati & Porter, 2009)

There are several tests of normality, the researcher considered just three: histogram of residuals, normal probability plot; a graphical device; and the Jarque-Bera test.

3.8.1. Histogram of Residuals:

A histogram of residuals is a simple graphic device that is used to learn something about the shape of the Probability Density Function (PDF) of a random variable. On the horizontal axis, we divide the values of the variable of interest (e.g., OLS residuals) into suitable intervals, and in each class interval we erect rectangles equal in height to the number of observations (i.e., frequency) in that class interval. If you mentally superimpose the bell shaped normal distribution curve on the histogram, you will get some idea as to whether normal (PDF) approximation may be appropriate.

This diagram shows that the residuals are not perfectly normally distributed; for a normally distributed variable the skewness should be zero and kurtosis should be 3.

3.8.2. **Normal Probability Plot (NPP):**

A comparatively simple graphical device to study the shape of the probability density function (PDF) of a random variable is the normal probability plot (NPP) which makes use of normal probability paper, a specially designed graph paper. On the horizontal, or *x*, axis, we plot values of the variable of interest and on the vertical, or *y*, axis, we show the expected value of this variable if it were normally distributed. Therefore, if the variable is in fact from the normal population, the NPP will be approximately a straight line.

3.8.3. **Jarque–Bera (JB) Test of Normality:** (Jaque & Bera, 1987)

It is based on the OLS residuals. This test first computes the skewness and kurtosis measures of the OLS residuals and uses the following

test statistic:

$$JB = n\left[\frac{s^2}{6} + \frac{(k-3)^2}{24}\right] \sim \chi_2^2 \qquad (3.8.1)$$

where n = sample size, S = skewness coefficient, and K = kurtosis coefficient.

For a normally distributed variable, $S = 0$ and $K = 3$. Therefore, the JB test of normality is a test of the joint hypothesis that S and K are 0 and 3, respectively. In that case the value of the JB statistic is expected to be 0.

Under the null hypothesis that the residuals are normally distributed, the JB statistic given in (3.8.1) follows the chi-square distribution with 2 df. If the computed p value of the JB statistic in an application is sufficiently low, which will happen if the value of the statistic is very different from 0, one can reject the hypothesis that the residuals are normally distributed. But if the p value is reasonably high, which will happen if the value of the statistic is close to zero, we do not reject the normality assumption.

3.9. HETEROSCEDASTICITY:– (Maddala,1992), (Gujarati & Porte, 2009).

Heteroscedasticity is an important assumption of the classical linear regression model so that the variance of each disturbance term u_t, conditional on the chosen values of the explanatory variables, is some constant number equal to σ^2. This is the assumption of homoscedasticity, or equal (homo)

spread (scedasticity), that is, *equal variance*. Symbolically,

$$E(u_t^2) = \sigma^2 \qquad i = 1, 2, ..., n$$

A critical assumption of the classical linear regression model is that the disturbances u_t have all the same variance, σ^2. If this assumption is not satisfied, there is heteroscedasticity.

Heteroscedasticity does not destroy the unbiasedness and consistency properties of OLS estimators. But these estimators are no longer minimum variance or efficient. That is, they are not BLUE, by reverting two variable models:

$$Y_t = \beta_1 + \beta_2 X_t + u_t \qquad (3.9.1)$$

But applying the usual formula, the OLS estimator of β_2 is

$$\widehat{\beta}_2 = \frac{n\sum X_t Y_t - \sum X_t \sum Y_t}{n\sum X_t^2 - (\sum X_t)^2}$$

$$\widehat{\beta}_2 = \frac{\sum x_t y_t}{\sum x_t^2}$$

$$= \sum k_t Y_t \quad , \qquad (3.9.2)$$

where $$k_t = \frac{x_t}{\sum x_t^2} \qquad (3.9.3)$$

which shows that $\hat{\beta}_2$ is a linear estimator because it is a linear function of Y; actually it is a weighted average of Y_t

with k_t serving as the weights. It can similarly be shown that $\hat{\beta}_1$ too is a linear estimator.

Incidentally, note these properties of the weights k_t :

1. Since the X_t are assumed to be nonstochastic, the k_t are nonstochastic too.

2. $\sum k_t = 0$.

3. $\sum k_t^2 = \dfrac{1}{\sum x_t^2}$

4. $\sum k_t x_t = \sum k_t X_t = 1$.

Substitute the PRF $Y_t = \beta_1 + \beta_2 X_t + u_t$ into equation (3.9.2) to obtain:

$$\begin{aligned}
\hat{\beta}_2 &= \sum k_t(\beta_1 + \beta_2 X_t + u_t) \\
&= \beta_1 \sum k_t + \beta_2 \sum k_t X_t + \sum k_t u_t \\
&= \beta_2 + \sum k_t u_t
\end{aligned}$$

(3.9.4)

Where use is made of properties of k_t noted above.

Now taking expectation of (3.9.4) on both sides and noting that k_t, being nonstochastic, can be treated as constants, we obtain:

$$E(\hat{\beta}_2) = \beta_2 + \sum k_t E(u_t) = \beta_2$$

since $E(u_t) = 0$ by assumption; therefore, $\hat{\beta}_2$ is an unbiased

estimator of β_2. Likewise, it can be proved that $\hat{\beta}_1$ is also an unbiased estimator of β_1 .

If we introduce heterocedasticity by letting $E(u_t^2) = \sigma_t^2$, then,

$$Var(\widehat{\beta}_2) = \sum k^2{}_t \sigma^2{}_t = \sum \left[\left(\frac{x_t}{\sum x^2{}_t} \right)^2 \sigma^2{}_t \right]$$

since

$$K_t = \frac{x_t}{\sum x^2{}_t} \qquad = \frac{\sum x^2{}_t \sigma^2{}_t}{\left(\sum x^2{}_t \right)^2}$$

Which is obviously different from the usual variance formula obtained under the assumption of homoscedasticity (Kmenta, 1986), namely,

$$Var(\widehat{\beta}_2) = \frac{\sigma^2}{\sum x^2{}_t}$$

In the presence of heteroscedasticity, the variances of OLS estimators are not provided by the usual OLS formulas. But if we persist in using the usual OLS formulas, the t and F tests based on them can be highly misleading, resulting in erroneous conclusions.

3.9.1. Detection of Heteroscedasticity:

Examine some of the informal and formal methods of detecting heteroscedasticity as the following:

3.9.1.1. Informal Methods:

- Nature of the Problem.
- Graphical Method.

3.9.1.2. Formal Methods:

- Park Test. (Park, 1966)
- Glejser Test. (Glesjser, 1969)
- Spearman's Rank Correlation Test. (Yule **et al.,** 1953)
- Goldfeld-Quandt Test. (Goldfeld & Quandt, 1972)
- Breusch-Pagan-Godfrey Test. (Breusch & Pagan, 1979), (Godfrey, 1978)
- White's General Heteroscedasticity Test. (White, 1980)
- Koenker-Bassett (KB) Test. (Harrison & McCabe, 1979)

The researcher preferred to use Breusch-Pagan-Godfrey Test, because it is suitable for this study.

3.9.2. Breusch-Pagan-Godfrey Test: (Breush & Pagan, 1979), (Godfrey, 1978)

To illustrate (BPG) test, consider the $K-$ variable linear regression model

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t \qquad (3.9.5)$$

Assume that the error variance σ^2_t is described as:

$$\sigma^2_t = f(\alpha_1 + \alpha_2 Z_{2t} + \cdots + \alpha_m Z_{mt})$$

That is σ^2_t is some function of the nonstochastic Z variables; some or all of the $X's$ can serve as $Z's$. Specifically, assume that:

$$\sigma^2_t = \alpha_1 + \alpha_2 Z_{2t} + \cdots + \alpha_m Z_{mt}$$

that is, σ^2_t is a linear function of the $Z's$.

if $\qquad \boldsymbol{\alpha_2 = \alpha_3 = \cdots = \alpha_m = 0}$, $\sigma^2_t = \boldsymbol{\alpha_1}$, which is a constant.

Therefore, to test whether σ^2_t is homoscedastic, one can test the hypothesis that $\boldsymbol{\alpha_2 = \alpha_3 = \cdots = \alpha_m = 0}$. This is the basic idea behind the Breusch – Pagan – Godfrey test.

The actual test procedure is as follows:

Step 1: Estimate (3.9.5) by OLS and obtain the residuals $\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n$.

Step 2: Obtain $\tilde{\sigma}^2 = \sum \hat{u}_t^2 / n$. That this is maximum likelihood (ML) estimator of σ^2.

Step 3: Construct variables p_t defined as: $\quad p_t = \hat{u}_t^2 / \tilde{\sigma}^2$ which is

simply each residual squared divided by $\tilde{\sigma}^2$.

Step 4: Regress p_t thus constructed on the Z's as:

$$p_t = \alpha_1 + \alpha_2 Z_{2t} + \cdots + \alpha_m Z_{mt} + v_t \quad (3.9.6)$$

where v_t is the residual term of this regression.

Step 5: Obtain the ESS (explained sum of squares) from (3.9.6) and define: (Adrian, 1994)

$$\ominus = \frac{1}{2}(ESS) \quad (3.9.7)$$

Assuming u_t are normally distributed, one can show that if there is homoscedasticity and if the sample size n increases indefinitely, then

$$\ominus \sim_{asy} \chi^2_{m-1} \quad (3.9.8)$$

that is, $\ominus$ follows the chi-square distribution with $(m-1)$ degrees of freedom. (*asy* means asymptotically).

Therefore, if in an application the computed $\ominus$ (χ^2) exceeds the critical χ^2 value at the chosen level of significance, one can reject the hypothesis of homoscedasticity; otherwise one does not reject it.

3.10. AUTOCORRELATION:- (Gujarati & Porter, 2009).

The term autocorrelation may be defined as "correlation between members of series of observations ordered in time [as in time series data] or space [as in cross-sectional data]."(Kendall & Buckland, 1971) in the regression context, the classical linear regression model assumes that such autocorrelation does not exist in the disturbances u_t. Symbolically,

$$cov(u_t, u_s | x_t, x_s) = E(u_t u_s) = 0 \quad t \neq s \qquad (3.10.1)$$

Now a common practice to treat the terms autocorrelation and serial correlation synonymously, some authors prefer to distinguish the two terms. For example, Tintner (1951) defines autocorrelation as "lag correlation of a given series with itself, lagged by a number of time units, whereas he reserves the term serial correlation to "lag correlation between two different series." Thus, correlation between two time series such as $u_1, u_2, \ldots, u_{10}$ and $u_2, u_3, \ldots, u_{11}$, where the former is the latter series lagged by one time period, is *autocorrelation,* whereas correlation between time series such as $u_1, u_2, \ldots, u_{10}$ and $v_2, v_3, \ldots, v_{11}$, where u and v are two different time series, is called serial correlation.

If the assumption of the classical linear regression model that the errors or disturbances u_t entering into the population regression function (PRF) are random or uncorrelated is violated, the problem of serial or autocorrelation arises.

Autocorrelation can arise for several reasons, such as inertia or sluggishness of economic time series, specification bias resulting from excluding important variables from the model or using incorrect functional form, the cobweb phenomenon, data messaging, and data transformation.

Although in the presence of autocorrelation the OLS estimators remain unbiased, consistent, and asymptotically normally distributed, they are no longer efficient. As a consequence, the usual t, F, and χ^2 tests cannot be legitimately applied. Hence, remedial results may be called for.

The mechanism that is commonly assumed is the Markov first-order autoregressive scheme, which assumes that the disturbance in the current time period is linearly related to the disturbance term in the previous time period, the coefficient of autocorrelation ρ providing the extent of the interdependence. This mechanism is known as the AR (1) scheme.

Revert to the two variable regression model:

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

Assume that the disturbance terms are generated by the following mechanism:

$$u_t = \rho u_{t-1} + \varepsilon_t \quad , \quad -1 < \rho < 1 \qquad (3.10.2)$$

where ρ is known as the coefficient of autocovariance and ε_t is the stochastic disturbance term such that it satisfies the standard OLS assumptions:

$$E(\varepsilon_t) = 0$$
$$Var(\varepsilon_t) = \sigma_\varepsilon^2$$
$$Cov(\varepsilon_t, \varepsilon_{t+s}) = 0 \qquad s \neq 0$$

From first order autoregressive scheme; in equation (3.10.2) (AR(1)), it can be shown that,

$$E(u_t) = \rho E(u_{t-1}) + E(\varepsilon_t) = 0$$
$$Var(u_t) = \rho^2 Var(u_{t-1}) + Var(\varepsilon_t)$$

Because the u's and $\varepsilon's$ are uncorrelated.

Since

$$Var(u_t) = Var(u_{t-1}) = \sigma^2 \quad and \quad Var(\varepsilon_t) = \sigma_\varepsilon^2$$

we get

$$Var(u_t) = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$$Cov(u_t, u_{t-1}) = E(u_t u_{t-1}) = E\left[\rho u_{t-1}^2 + u_{t-1}\varepsilon_t\right] = \rho E(u_{t-1}^2)$$

$$Cov(u_t, u_{t-1}) = \rho \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$$Cov(u_t, u_{t-s}) = \rho^s \frac{\sigma_\varepsilon^2}{1-\rho^2} \qquad (3.10.3)$$

$$Cor(u_t, u_{t-s}) = \rho^s$$

Since ρ is a constant between -1 and $+1$, equation (3.10.3) shows that under the AR(1) scheme, the variance of u_t is still homoscedastic, but u_t is correlated not only with its immediate past value but its values several periods in the past. It is critical to note that $|\rho| < 1$, that is, the absolute value of rho is less than one. If $|\rho| < 1$, we say that the AR(1) process is stationary; that is, the mean, variance, and covariance of u_t do not change over time. If $|\rho|$ is less than one, then it is clear from (3.10.3) that the value of the covariance will decline as we go into the distant past.

3.10.1. Detecting Autocorrelation:

- A General Test of Autocorrelation: The Breush-Godfrey (BG) Test. (Breush,1978), (Godfrey, 1978)
- Durbin-Waston d test. (Durbin & Waston, 1951)
- The Runs Test. (Geary, 1970)
- Graphical Method. (Draper & Smith, 1998)

The researcher preferred to use Breush-Godfrey (BG)

Test, because it is more powerful in statistical sense.

3.10.1.1. The Breusch-Godfrey (BG) Test: (Breusch, 1978), (Godfrey, 1978)

Statisticians Breusch and Godfrey have developed a test of autocorrelation that is general in the sense that it allows for (1) no stochastic regressors, such as the lagged values of the regress and; (2) higher-order autoregressive schemes, such as AR(1), AR(2), etc.; and (3) simple or higher-order moving averages of white noise error terms.

The BG test, which is also known as the Lagrange Multiple Principle (LM) test, proceeds as follows; we use the two-variable regression model to illustrate the test, although many regressors can be added to the model. Also, lagged values of the regress and can be added to the model. Let:

$$Y_t = \beta_1 + \beta_2 x_t + u_t \qquad\qquad (3.10.4)$$

Assume that the error term u_t follows the pth-order autoregressive, AR(p), scheme as follows:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots \rho_p u_{t-p} + \varepsilon_t \qquad (3.10.5)$$

Where ε_t is a white noise error term. The null hypothesis H_0 to be tested is that,

$$H_0: \rho_1 = \rho_2 = \cdots \rho_p = 0$$

That is, there is no serial correlation of any order. The BG test involves the

Following steps:

1. Estimate (3.10.4) by OLS and obtain the residuals, $\hat{u}_t$.

2. Regress $\hat{u}_t$ on the original x_t (if there is more than one X variable in the original model, include them also) and $\hat{u}_{t-1}, \hat{u}_{t-2}, \ldots, \hat{u}_{t-p}$, where the latter are the lagged values of the estimated residuals in step 1. Thus, if $p = 4$, we will introduce four lagged values of the residuals as additional regressor in the model. Note that to run this regression:

$$\hat{u}_t = \alpha_1 + \alpha_2 X_t + \hat{\rho}_1 \hat{u}_{t-1} + \hat{\rho}_2 \hat{u}_{t-2} + \cdots \hat{\rho}_p \hat{u}_{t-p} + \varepsilon_t$$
$$(3.10.6)$$

and obtain R^2 from this (auxiliary) regression.

3. If the sample size is large (technically, infinite), Breusch and Godfrey have shown that

$$(n - p)R^2 \sim \chi_p^2 \qquad (3.10.7)$$

That is, asymptotically, $n - p$ times the R^2 value obtained from the auxiliary regression (3.10.6) follows

the chi-square distribution with p df. If in an application, $(n-p)R^2$ exceeds the critical chi-square value at the chosen level of significance, we reject the null hypothesis, in which case at least one ρ in (3.10.5) is statistically significantly different from zero.

3.10.1.2. Durbin-Watson d Test: (Durbin & Watson, 1951)

The most celebrated test for detecting serial correlation is that developed by statisticians Durbin and Watson. It is popularly known as the Durbin–Watson d statistic, which is defined as:

$$d = \frac{\sum_{t=2}^{t=n}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{t=n}\hat{u}_t^2} \qquad (3.10.8)$$

which is simply the ratio of the sum of squared differences in successive residuals to the RSS. Note that in the numerator of the d statistic the number of observations is n -1 because one observation is lost in taking successive differences.

A great advantage of the d statistic is that it is based on the estimated residuals, which are routinely computed in regression analysis. Because of this advantage, it is now a common practice to report the Durbin–Watson d along with summary measures, such as R^2, adjusted R^2, t, and F.

$$d = \frac{\sum \hat{u}_t^2 + \sum \hat{u}_{t-1}^2 - 2\sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2} \qquad (3.10.9)$$

Since $\sum \hat{u}_t^2$ and $\sum \hat{u}_{t-1}^2$ differ in only one observation, they are approximately equal. Therefore, setting $\sum \hat{u}_{t-1}^2 \approx \sum \hat{u}_t^2$; so eq. $(3.10.9)$ is written as follows:

$$d \approx 2\left(1 - \frac{\sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2}\right) \quad (3.10.10)$$

Now let define; $\qquad \hat{\rho} = \dfrac{\sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2}$

As the sample first-order coefficient of autocorrelation, an estimator of ρ. We can express eq. $(3.10.10)$ as follow:

$$d \approx 2(1 - \hat{\rho}) \quad (3.10.11)$$

But since $-1 \leq \rho \leq 1$, eq. $(3.10.11)$ implies that:

$$0 \leq d \leq 4$$

These are the bounds of d; any estimated d value must lie within these limits.

It is apparent from Eq. $(3.10.10)$ that if $\hat{\rho} = 0$, $d = 2$; that is, if there is no serial correlation (of the first-order), d is expected to be about 2. Therefore, as a rule of thumb, if d is found to be 2 in an application, one may assume that there is no first-order autocorrelation, either positive or negative. If $\hat{\rho} = +1$, indicating perfect positive correlation in the residuals, $d \approx 0$. Therefore, the closer d is to 0, the greater the evidence of positive serial correlation.

If $\hat{\rho} = -1$, that is, there is perfect negative correlation among successive residuals, $d \approx 4$. Hence, the closer d is to 4, the greater the evidence of negative serial correlation.

If a regression model contains lagged value(s) of the regressand, the d value in such cases is often around 2, which would suggest that there is no (first-order) autocorrelation in such models.

3.10.1.3. The Runs Test: (Geary, 1970)

Sometimes also known as the Geary test. To e explain the runs test, let us simply note down the signs (+ or −) of the residuals. By examining how runs behave in a strictly random sequence of observations, one can derive a test of randomness of runs. If there are too many runs, it would mean that the residuals change sign frequently, thus indicating negative serial correlation. Similarly, if there are too few runs, they may suggest positive autocorrelation.

N = total number of observations = $N_1 + N_2$.

N_1 = number of (+) symbols (+ residuals).

N_2 = number of (−) symbols (− residuals).

R = number of runs.

Mean: $$E(R) = \frac{2N_1N_2}{N} + 1$$

(3.10.12)

Variance: $$\sigma_R^2 = \frac{2N_1N_2(2N_1N_2 - N)}{N^2(N-1)}$$

If the null hypothesis of randomness is sustainable, following the properties of the normal distribution, we should expect that:

Prob $[E(R) - 1.96\, \sigma_R \leq R \leq E(R) + 1.96\, \sigma_R] = 0.95$

That is, the probability is 95 percent that the preceding interval will include R. Therefore we have this rule:

Do not reject the null hypothesis of randomness with 95% confidence if R, the number of runs, lies in the preceding confidence interval; reject the null hypothesis if the estimated R lies outside these limits.

3.10.1.4. **Graphical Method:** (Draper & Smith, 1998)

The importance of producing and analyzing plots of [residuals] as a standard part of statistical analysis cannot be overemphasized. Besides occasionally providing an easy to understand summary of a complex problem, they allow the simultaneous examination of the data as an aggregate while clearly displaying the behavior of individual cases.

There are various ways of examining the residuals. We can simply plot them against time, the time sequence plot, standardized residuals.

3.10.2. **The Phenomenon of Spurious Regression:** (Yule, 1926), (Granger & Newbold, 1974)

Yule (1926) showed that (spurious) correlation could persist in nonstationary time series even if the sample is very large. That there is something wrong in the preceding regression is suggested by the extremely

low Durbin–Watson d value, which suggests very strong first-order autocorrelation. According to Granger and Newbold (1974), *an $R^2 > d$* is a good rule of thumb to suspect that the estimated regression is spurious.

3.11. <u>MEASURES OF FORECAST ACCURACY:-</u> (Balcila, 2007), (Cipan, 2004)

- Good forecasts have three desirable characteristics:
  1. Unbiased: We want $E(\hat{y}_t) = y_t$, which implies $E(\hat{u}_t) = 0$. Therefore, models with small average errors are preferred.
  2. Efficient: We want small forecast error variance
  $$E(y_t - \hat{y}_t)^2 = E(u_t^2)$$
  Therefore, models with small sums of squared errors are preferred.
  3. No pattern in $\hat{u}_t$ series.
- Measures of absolute forecast accuracy:
  1. Mean Error (ME), Mean Absolute Error (MAE).
  2. Regression based measures: Standard error of regression (SEE) and coefficient of determination ($0 \leq R^2 \leq 1$); the closer to 1 is better the fit.
- Visually checking goodness of fit:
  1. Fitted line, plot y_t and $\hat{y}_t$ with a good fit, $\hat{y}$ captures the basic pattern in y.
  2. Residual plot, plot $\hat{u}_t$ approximate 95% confidence interval.
- More sophisticated measure of accuracy:
  1. Useful breakdown of MSE:

  Bias%+ Variance%+ Covariance%=100

  Bias and variance proportions show the extent to which the forecast

$\hat{y}$ missed the mean and variance of y. For a good forecast, both of these will be small.

2. Theil's Inequality Coefficient ($0 \leq I \leq 1$), I=0 implies a perfect fit.

3.12. <u>DUMMY VARIABLE REGRESSION MODELS:-</u> (Maddala, 1992), (Gujarati & Porter, 2009).

Dummy variables are a data-classifying device in that they divide a sample into various subgroups based on qualities or attributes and implicitly allow one to run individual regressions for each subgroup. If there are differences in the response of the regressand to the variation in the qualitative variables in the various subgroups, they will be reflected in the differences in the intercepts or slope coefficients, or both, of the various subgroup regressions.

Dummy variables, taking values of 1 and 0 (or their linear transforms), are a means of introducing qualitative regressors in regression models.

Being a versatile tool, the dummy variable technique needs to be handled carefully. At first, if the regression contains a constant term, the number of dummy variables must be one less than the number of classifications of each qualitative variable. Secondly, the coefficient attached to the dummy variables must always be interpreted in relation to the base, or reference, the group that receives the value of zero.

The base chosen will depend on the purpose of research at hand.

3.13. PANEL DATA:-

This is a special type of pooled data in which the same cross-sectional unit is surveyed over time. (In short, panel data have space as well as time dimension.) (Gujarati & Porter, 2009). There are other names for panel data, such as:

- Combination of time series and cross section data.
- Micropanel data.
- Longitudinal data: a study over time of a variable or group of subjects.
- Event history analysis: studying movement over time of subjects through successive states or conditions.
- Cohort analysis.

Although there are subtle variations, all these names essentially connote movement over time of cross-sectional units.

3.13.1. Advantages of Panel Data:

The following advantages of panel data: (Baltagi, 1995)

1. Since panel data relate to individuals, firms, states, countries, etc; over time, there is bound to heterogeneity in these units. The techniques of panel data estimation can take such heterogeneity explicitly into account by allowing for subject-specific variables.

2. By combining time series of cross-section observations, panel data gives, more informative data, more variability, less collinearity among variables, more degrees of freedom and more efficiency'.

3. By studying the repeated cross-section observations, panel data are better suited to study the dynamics change.

4. Panel data can better detect and measure effects that simply cannot be observed in pure cross-section or pure time series data.

5. Panel data enables us to study more complicated behavioral models.

6. By making data available for several thousand units; panel data can minimize the bias that might result.

3.13.2. Balance/ Unbalanced Data: (Gujarati & Porter, 2009)

A panel is said to be balanced if each subject has the same number of observations. If each entity has a different

number of observations, then we have an unbalanced panel.

3.13.3. Long/ Short Panel Data: (Gujarati & Porter, 2009)

In a short panel the number of cross-sectional subjects, N, is greater than the number of time periods, T. In long panel, it is T that is greater than N. The estimation techniques can depend on whether short or long panel.

3.14. PANEL DATA MODELS:-

The variables are defined as:

I: Airline id.

T: Year id.

PAX: Output1: Passenger.

FRT: Output2: Cargo.

TC: Total cost in US $.

FC: Fuel price in US $.

LF: Load factor, the average capacity utilization of the fleet.

The researcher was interested in finding out how the total cost (TC) behaves in relation to passenger (PAX), cargo (FRT) fuel price (FC), and load factor (LF).There are four possibilities models:

3.14.1. Pooled OLS Regression Model:

The researcher pools all 50 observations and estimates a grand regression, neglecting the cross-section and time series nature of our data.

Consider the following model:

$$TC_{it} = \beta_1 + \beta_2 PAX_{it} + \beta_3 FRT_{it} + \beta_4 LF_{it} + \beta_5 FC_{it} + u_{it} \quad (3.14.1)$$

$$i = 1, 2, \dots, 5 \quad , \quad t = 1, 2, \dots, 10$$

Where;

i is ith subject and t is the time period for the variables defined.

Assumptions: (Cameron & Trivedi, 2005), (Gujarati & Porter, 2009)

1. The regression coefficients are the same for all the airlines. That is, there is no distinction between the airlines.
2. The explanatory variables are nonstochastic.
3. The error term is $u_{it} \sim iid(0, \sigma_u^2)$, that is; it is independently and identically distributed with zero mean and constant variance.

3.14.2. The Fixed Effect Least-Squares Dummy Variable (LSDV) Model:

The least-square dummy variable (LSDV) model allows for heterogeneity among subjects by allowing each entity to have its own intercept value. (Cameron & Trivedi, 2005), (Gujarati & Porter, 2009)

Consider the following model:

$$TC_{it} = \beta_{1i} + \beta_2 PAX_{it} + \beta_3 FRT_{it} + \beta_4 LF_{it} + \beta_5 FC_{it} + u_{it} \qquad (3.14.2)$$

$$i = 1, 2, \dots, 5 \qquad t = 1, 2, \dots 10$$

Where;

i is ith subject and t is the time period for the variables defined.

Note:

The subscript i on the intercept term to suggest that the intercepts of the five airlines may be different; due to special features of each airline. This model is known as **Fixed Effects (regression) Model (FEM).** The term "fixed effect" is due to the fact that, although the intercept may differ across the five airlines, each entity's intercept does not vary over time (time invariant). (Maddala, 1992), (Gujarati & Porter, 2009)

Assumption:

Assumes that the slope coefficients of regressors do not vary across individuals or over the time.

The researcher will use the **Differential Intercept Dummy Technique**, to determine how actually the intercept (fixed effect) vary among airlines. This is defining as follow:

$$TC_{it} = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \alpha_5 D_{5i} + \beta_2 PAX_{it} + \beta_3 FRT_{it} +$$
$$\beta_4 LF_{it} + \beta_5 FC_{it} + u_{it}$$
(3.14.3)

Where:

$D_{2i} = 1\ for\ airline\ 2, 0\ otherwise$;

$D_{3i} = 1\ for\ airline\ 3, 0\ otherwise$; and so on.

The researcher has introduced only four dummy variables to avoid falling into the **dummy variable trap**.

Here the researcher treating airline 1 (Sudan Airways) as the base, or reference. As a result, the intercept α_1 is the intercept value of airline 1 and the other α coefficients represent by how much the intercept values of the other airlines differ from the intercept value of the first airline. Thus, α_2 tells by how much the intercept value of the second airline differs from α_1. The sum $(\alpha_1 + \alpha_2)$ gives the actual value of the intercept for airline 2. The intercept values of the other airlines will be computed similarly. (Maddala, 1992), (Gujarati & Porter, 2009)

3.14.2.1. The F-Test Approach: Restricted Least Squares:

To compare the unrestricted and restricted least squares regression, we can apply the F test as follows: (Theil, 1971)

Let:

$\sum \hat{u}_{UR}^2$ = RSS of the unrestricted regression.

$\sum \hat{u}_R^2$ = RSS of the restricted regression.

m = number of linear restrictions.

k = number of parameter in the unrestricted regression.

n = number of observations.

Then,

$$F = \frac{(RSS_R - RSS_{UR})/m}{RSS_{UR}/(n-k)}$$

$$F = \frac{(\sum \hat{u}_R^2 - \sum \hat{u}_{UR}^2)/m}{\sum \hat{u}_{UR}^2/(n-k)}$$

Follows the F distribution with m, (n-k) df.

The F test above can also be expressed in terms of R^2 as follows:

$$F = \frac{(R_{UR}^2 - R_R^2)/m}{(1 - R_{UR}^2)/(n-k)}$$

According to F-test approach, the researcher applied Wald test to test the following hypothesis: (Baltagi, 1995)

Null hypothesis: all dummy variables are equal to zero, that mean pooled regression model is appropriate.

Alternative hypothesis: fixed effect least square dummy variable (LSDV) is appropriate.

3.14.3. The Fixed Effect Within-Group (WG) Estimator:

One way to estimate a pooled regression is to eliminate the fixed effect, β_{1i} , by expressing the values of the dependent and explanatory variables for each airline as deviations from their respective mean values. Thus, for airline1 we will obtain the sample mean values of TC, PAX, FRT, LF and FC, ($\overline{TC}, \overline{PAX}, \overline{FRT}, \overline{LF}, \overline{FC}$, respectively) and subtract them from individual values of these variables (Cameron & Trivedi, 2005), (Gujarati & Porter, 2009). The resulting values are called "de-meaned" or mean corrected values. So the researcher does this for each airline and then pools all 50 mean-corrected values and run an OLS regression. Letting tc_{it} , pax_{it} , frt_{it} , lf_{it} and fc_{it} represent the mean-corrected values and run the following regression:

$$tc_{it} = \beta_2 pax_{it} + \beta_3 frt_{it} + \beta_4 lf_{it} + \beta_5 fc_{it} + u_{it} \qquad (3.14.4)$$

Where,

$$i = 1, 2, \dots, 5 \qquad t = 1, 2, \dots 10$$

Note that equation (3.14.4) does not have an intercept term, because of differencing.

We obtain the estimates of the intercept of the ith airline using the WG method, by subtracting from the mean value of the dependent variable the mean values of the explanatory variables for the airline times the estimated slope coefficients from the WG estimators. Note that the estimated slope coefficients remain the same for all airlines and the estimated intercept of each airline represents the subject-specific characteristics of each airline, but not able to identify these characteristics individually. (Cameron & Trivedi, 2005), (Gujarati & Porter, 2009).

$$\hat{\alpha}_i = \overline{TC_i} - \hat{\beta}_2 \overline{PAX_i} - \hat{\beta}_3 \overline{FRT_i} - \hat{\beta}_4 \overline{LF_i} - \hat{\beta}_5 \overline{FC_i} \qquad (3.14.5)$$

Where bars over the variables denote the sample mean values of the variables for the ith airline.

3.14.4. The Random Effects Model (REM):

It is known also as Error Components Model (ECM), is so named because the composite error term consists of two or more error components. (Kmenta, 1986), (Gujarati & Porter, 2009).The basic idea is starting in the following model:

$$TC_{it} = \beta_{1i} + \beta_2 PAX_{it} + \beta_3 FRT_{it} + \beta_4 LF_{it} + \beta_5 FC_{it} + u_{it} \quad (3.14.6)$$

Instead of treating β_{1i} as fixed, we assume that it is a random variable with a mean value of β_1. The intercept value for an individual company can be expressed as follows:

$$\beta_{1i} = \beta_1 + \varepsilon_i$$

Where ε_i is a random error term with mean value of zero and variance of σ_ε^2. The individual differences in the intercept values of each company are reflected in the error term ε_i. The equation obtains as follow:

$$TC_{it} = \beta_1 + \beta_2 PAX_{it} + \beta_3 FRT_{it} + \beta_4 LF_{it} + \beta_5 FC_{it} + w_{it} \quad (3.14.7)$$
Where,

$$w_{it} = \varepsilon_i + u_{it}$$

The composite error term w_{it} consists of two components:
ε_i : Which is the cross-section or individual-specific error component.

u_{it}: Which is the combined time series and cross-section error component and is sometimes called idiosyncratic term, because it varies over cross-section as well as time.

The assumptions made by the ECM are that: (Kmenta, 1986), (Gujarati & Porter, 2009)

$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$$
$$u_{it} \sim N(0, \sigma_u^2)$$
$$E(\varepsilon_i u_{it}) = 0 \ , \ E(\varepsilon_i \varepsilon_j) = 0 \quad (i \neq j)$$
$$E(u_{it} u_{is}) = E(u_{jt} u_{js}) = E(u_{it} u_{js}) = 0 \quad (i \neq j \ ; t \neq s)$$

That is, the individual error components are not correlated with each other and are not autocorrelated across both cross-section and time series units. Note that w_{it} is not correlated with any of the explanatory variables included in the model. Since ε_i is a component of w_{it} , it is possible that the latter is correlated with the explanatory variables. If that is indeed the case, the ECM will result in inconsistent estimation of the regression coefficients. So that we use Hausman test, which tells us in a given application if w_{it} is correlated with the explanatory variables, that is whether ECM is the appropriate model.

There are substantial difference between FEM and ECM. In FEM each cross-sectional unit has its own (fixed)

intercept value, in all N such values for N cross-sectional units. In ECM, on the other hand, the intercept represents the mean value of all (cross-sectional) intercepts, and the error component ε_i represents the (random) deviation of individual intercept from this mean value (Kmenta, 1986), (Gujarati & Porter, 2009).

3.14.4.1. Hausman's Specification Test for the Random Effects Model: (Hausman, 1978)

The specification test devised by Hausman is used to test for orthogonality of random effects and the regressors. The test is based on the idea that under the hypothesis of no correlation, both OLS in the LSDV model and GLS are consistent, but OLS is inefficient, whereas under the alternative, OLS is consistent, but GLS is not. Therefore, under the null hypothesis, the two estimates should not differ systematically, and a test can be based on the difference. Hausman's essential result is that the covariance of an efficient estimator with its difference from an inefficient estimator is zero.

The null of the REM/ ECM and the alternative of FEM correspond to the Hausman situation:

- In the REM, the GLS-type RE estimator is efficient by construction for Gaussian errors; the FE estimator and even the OLS estimator are consistent.

111

- In the FEM, the RE estimator is inconsistent, because of omitted-variable effect, while FE estimator is consistent by construction.

The Hausman test statistic is defined as:

$$m = \acute{q}(var(\widehat{\beta}_{FE}) - var(\widehat{\beta}_{RE}))^{-1}q$$

$$\text{with} \quad q = \widehat{\beta}_{FE} - \widehat{\beta}_{RE}$$

Under RE, the matrix difference in brackets is positive, as the RE estimator is efficient and any other estimator has a large variance.

The statistic m is distributed χ^2 under the null of RE, with degrees of freedom determined by the dimension of $, k$.

So that the researcher needs to determine which model is appropriate ECM or FEM. Here, the researcher applied Hausman Test to test the following hypothesis:

Null hypothesis: random effects model (ECM)/(REM) is appropriate. Alternative hypothesis: fixed effects within group (FEM) is appropriate.

3.15. <u>SIMPLE EXPONENTIAL SMOOTHING:-</u> (Hyndman, 2014)

The simplest of the exponentially smoothing methods is naturally called "simple exponential smoothing" (SES). This method is suitable for

forecasting data with no trend or seasonal pattern. Using the naïve method, all forecasts for the future are equal to the last observed value of the series,

$$\hat{y}_{T+h|T} = y_T \quad , \quad h = 1, 2, \dots$$

Hence, the naïve method assumes that the most current observation is the only important one and all previous observations provide no information for the future. This can be thought of as a weighted average where all the weight is given to the last observation.

Using the average method, all future forecasts are equal to a simple average of the observed data,

$$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=1}^{T} y_t \quad , \quad h = 1, 2, \dots$$

Hence, the average method assumes that all observations are of equal importance and they are given equal weight when generating forecasts.

Forecasts are calculated using weighted averages where the weights decrease exponentially as observations come from further in the past; the smallest weights are associated with the oldest observations:

$$\hat{y}_{T+h|T} = \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \cdots$$

Where $0 \leq \alpha \leq 1$ is the smoothing parameter. The one-step-ahead forecast for time T+1 is a weighted average of all the observations in the series $y_1, \dots, y_T$. The rate at which the weights decrease is controlled by the parameter α.

3.15.1. Weighted Average Form:

The forecast at time t+1 is equal to a weighted average between the most recent observation y_t and the most recent forecast $\hat{y}_{t|t-1}$,

$$\hat{y}_{t+1|t} = \alpha y_t + \alpha(1-\alpha)\hat{y}_{t|t-1}$$

For $t = 1, ..., T$, where $0 \leq \alpha \leq 1$ is the smoothing parameter.

The process has to start somewhere, so we let the first forecast of y_1 be denoted by ℓ_0. Then,

$$\hat{y}_{2|1} = \alpha y_1 + (1-\alpha)\ell_0$$
$$\hat{y}_{3|2} = \alpha y_2 + (1-\alpha)\hat{y}_{2|1}$$
$$\vdots$$
$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha)\hat{y}_{T|T-1}$$

3.15.2. Component Form:

The component form of simple exponential smoothing is given by:

Forecast equation $\hat{y}_{t+1|t} = \ell_t$

Smoothing equation $\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$

where ℓ_t is the level (or the smoothed value) of the series at time t. The forecast equation shows that the forecasted value at time t+1 is the estimated level at time t. The smoothing equation for the level (usually referred to as the level equation) gives the estimated level of the series at each period t. Applying the forecast equation for time T gives, $\hat{y}_{T+1|T} = \ell_T$, the most recent estimated level. If we replace ℓ_t by $\hat{y}_{t+1|t}$ and ℓ_{t-1} by $\hat{y}_{t|t-1}$ in the smoothing equation, we will recover the weighted average form of simple exponential smoothing.

The next chapter contains analysis of Sudanese domestic air transport data, by using Classical Normal Linear Regression Model (CNLRM) & Panel Regression Models; and to discuss outcomes of the analysis and the study possibility of using study techniques to estimate equation of airline data and evaluate the data to obtain the results.

CHAPTER FOUR

RESULTS & ANALYSIS

4 - 0   INTRODUCTION.

4 - 1   DESCRIPTION OF DATA.

4 - 2   ANALYSIS OF DATA BY USING
CLASSICAL LINEAR REGRESSION
MODEL.

4 - 3   ANALYSIS OF DATA BY USING PANEL
DATA REGRESSION MODEL.

CHAPTER FOUR

RESULTS & ANALYSIS

4. INTRODUCTION:-

This chapter describes the methods used to analyze the data obtained in this research for determining a cost function for each aircraft type operated by each airline during the period (2004 – 2013) in this research.

4.1. DESCRIPTION OF DATA:-

As in appendix (A), Table (A.1): Sudan Airways, Table (A.2): Marsland Aviation, Table (A.3): Sun Air, Table (A.4): Nova Air, Table (A.5): Mid Airlines and Table (A.6): Badr Airlines; describe Freight per annum in kilograms, Number of Passenger per annum, Load Factor (LF) in revenue passenger per annum, Fuel Cost per annum in US dollars and the Total Cost per annum in US dollars.

4.2. ANALYSIS OF DATA BY USING CLASSICAL LINEAR REGRESSION MODEL:-

In this part the researcher interest to study how the total cost (TC) behaves in relation to the number of passenger

(PAX), freight (FRT), load factor (LF) and fuel cost (FC), for each six airlines separately, and to evaluate the estimated model of total cost; by using Classical Normal Linear Regression Model (CNLRM); during ten years from 2004 to 2013, and then evaluate the model for forecasting. The model is represented as follows:

$$TC_t = \beta_0 + \beta_1 PAX_t + \beta_2 FRT_t + \beta_3 LF_t + \beta_4 FC_t + u_t \quad (4.2.1)$$

$$t = 1, \dots, 10$$

4.2.1. Sudan Airways:

4.2.1.1. Evaluation of the Model:

According to the data in appendix (A); table (A.1), the estimated value of the total cost by using (CNLRM) is represented as follows:

Table (4.1): Significant Value of the Total Cost (TC) Regression Model (4.2.2) of Sudan Airways:

Variable	Coefficient	Prob.
C	4848827	0.0648
PAX	-0.608596	0.9465
FRT	0.008605	0.9642

LF	-5131981	0.0788
FC	1.098333	0.0001

$$TC_t = 4848827 - 0.608596PAX_t + 0.008605FRT_t - 5131981LF_t + 1.098333FC_t$$

(4.2.2)

As shown in the table (4.1); there is 25% of the explanatory variables represented only on fuel cost (FC); are statistically significant at level 5%; to influence the dependent variable; total cost (TC), **(Appendix C.1.1)**. According to that, the researcher suspect there is a problem of multicollinearity in the model and expects there is high correlation between any two explanatory variables, which will be determined by using the following table:

Table (4.2): Correlation Matrix of Explanatory Variables:

Explanatory Variables	PAX	FRT	FC	LF
PAX	1	0.237142	−0.593871	0.787198
FRT	0.237142	1	−0.476148	0.251174
FC	−0.593871	−0.476148	1	−0.780397
LF	0.787198	0.251174	−0.780397	1

Source: Prepared by the researcher.

According to this problem, the Chi square P-value of Breusch-Godfrey Serial Correlation LM test equal **0.0358** is statistically significant at 5% level, so we can reject the null hypothesis; that residual is not serial correlation. **(Appendix C.1.3)**.

As shown in the table (4.2); there is a higher correlation between the load factor (LF) and passenger (PAX) rather than correlation between the load factor (LF) and freight (FRT), are results as not statistical significant in the model (4.2.2). So, the researcher had to drop one of the two variables (LF and PAX), that registered high correlation (79%).

By dropping PAX; which has higher P-value **(0.9465)**; the estimated value of total cost regression model is transferred as follow:

Table (4.3): Significant Value of the Total Cost (TC) Regression Model (4.2.3) of Sudan Airways:

Variable	Coefficient	Prob.
C	4861494	0.0407
FRT	0.007127	0.967
LF	-5238669	0.0178
FC	1.09743	00000

Source: Prepared by the researcher.

$$TC_t = 4861494 + 0.007127FRT_t - 5238669LF_t + 1.09743FC_t$$

(4.2.3)

As shown in the table (4.3), there is 67% of the explanatory variables: load factor (LF) and fuel cost (FC); are statistically significant at level 5%, to influence the dependent variable; total cost (TC), **(Appendix C.1.4)**.

As expected, the partial coefficient 0.007127 of freight says that; if the freight increases by hundred kgs, the total cost increases by 0.71 US dollars. The partial coefficient -5238669 of load factor says that; if the load factor increases by one

revenue average, the total cost decreases by **5238669** US dollars. The partial coefficient **1.09743** of fuel cost says that; if the fuel cost increases by one US dollar, the total cost increases by **1.097** US dollars.

The intercept value of about **4861494**, mechanically interpreted means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean total cost would be about **4861494** US dollars.

Table (4.4): Tests Results of Goodness of Fit of the Estimated Regression Model (4.2.3) of Sudan Airways:

R-squared	0.990847
Prob (F-statistic)	0.000002
Durbin-Watson Stat.	2.332225
Jarque-Bera-Normality Prob.	0.613088
Breusch-Godfrey Serial Correlation Chi-square Prob.	0.2347
Breusch-Pagan-Godfrey Heteroscedasticity Chi-square Prob.	0.6117

Source: Prepared by the researcher.

The R-squared value of about **0.990847** is statistically a significant value (more than 60%), means that about 99% of variation in the total cost is explained by freight, load factor

and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (2.332225) is found to be 2, so there is no first-order autocorrelation either positive or negative. Also, the R-squared (0.990847) is less than Durbin-Watson statistic, which means that this model is not spurious. (Appendix C.1.4).

The probability of F-statistic equal to 0.000002 is statistically significant at level 5%, means that the explanatory variables: freight, load factor and fuel cost are jointly significant to influence the total cost.

The P-value of Jarque-Bera normality test equal to 0.613088 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are normally distributed. (Appendix C.1.7).

The Chi square P-value of Breusch-Godfrey serial correlation LM test equal to 0.2347 is not statistically significant at 5% level; so, we cannot reject the null hypothesis; that residuals are not serial correlation. (Appendix C.1.8).

The Chi square P-value of Breusch-Pagan-Godfrey heteroscedasticity test equal to 0.6117 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are not heteroscedasticity. (Appendix C.1.9).

According to the above results, the researcher concludes that the residuals are normally distributed, not autocorrelated and homoscedastic; so, this result means that the estimated regression model (4.3.3) makes sense, and is acceptable to predictive purposes and forecasting.

4.2.1.2. Forecasting:

Table (4.5): Result of Forecasting Sample 2004-2013 of Sudan Airways:

Root Mean Squared Error	**615993.2**
Theil Inequality Coefficient	**0.034086**
Bias Proportion	0.00000

Source: Prepared by the researcher.

As shown in table (4.5), the root mean squared error is equal to 615993.2, while Theil Ineqauality coefficient equal 0.034086, which is close to zero that means; the predictive power of this model is very strong. Bias proportion equals zero that means, there is no gap between the actual total cost and the predictive total cost, and they are moving closely. (Appendix C.1.10).

Graph (4.1): Forecasting Sample 2004-2013 of Sudan Airways:
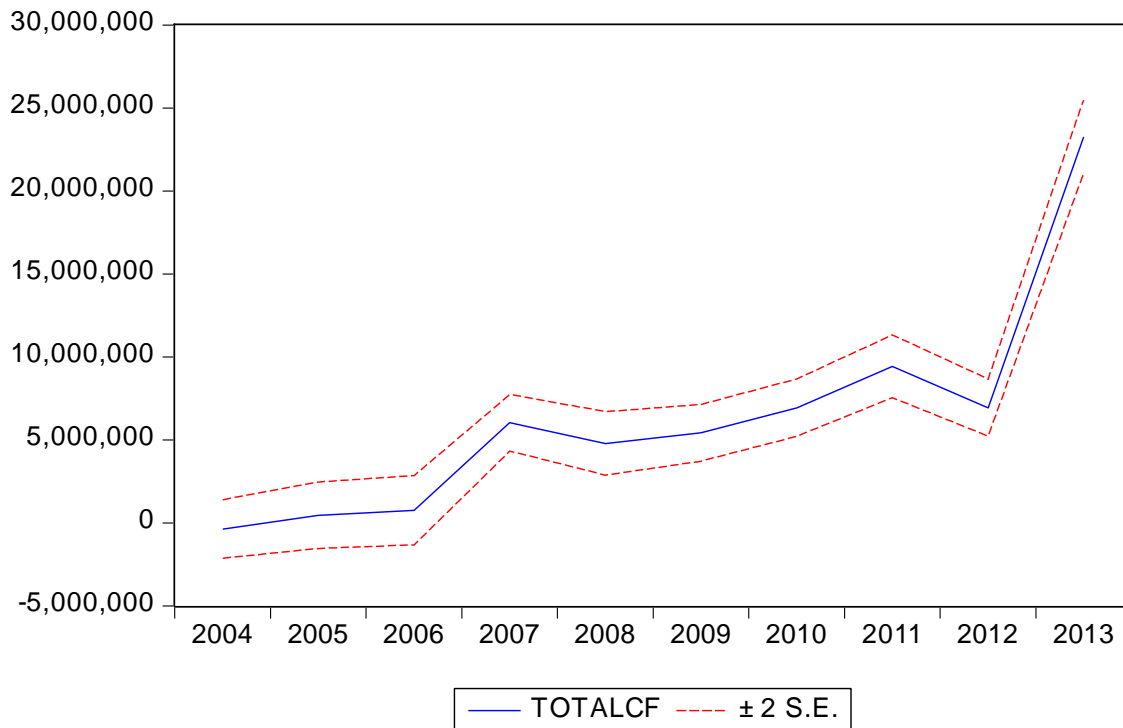
Source: Prepared by the researcher.

As shown in graph (4.1), the total cost value has been forecast and is passing throw 50% confidence interval; so, the forecasting of the total cost is significant and the ability of forecasting model is satisfactory.

4.2.2. Marsland Aviation:

4.2.2.1. Evaluation of the Model:

According to the data in appendix (A); table (A.2), the estimated value of the total cost (CNLRM) is represented as follows:

Table (4.6): Significant Value of the Total Cost (TC) Regression Model (4.2.4) of Marsland Aviation:

Variable	Coefficient	Prob.
C	3141858	0.2104
PAX	-3.678765	0.4985
FRT	-3.989002	0.0059
LF	-209550.1	0.9398
FC	1.048604	0.0000

Source: Prepared by the researcher.

$$TC_t = 3141858 - 3.678765PAX_t - 3.989002FRT_t - 209550.1LF_t + 1.048604FC_t$$

(4.2.4)

As shown in the table (4.6), there is 50% of the explanatory variables: freight (FRT) and load factor (LF); are statistically significant at level 5%, to influence the dependent

variable; total cost (TC), (Appendix C.2.1).

As expected, -3.678765 is the partial regression coefficient of passenger says that; if the number of passenger increases by one passenger per annum, the total cost decreases by about 3.68 US dollars. The partial coefficient -3.989002 of freight says that; if the freight increases by hundred kgs, the total cost decreases by 398.9 US dollars. The partial coefficient -209550.1 of load factor says that; if the load factor increases by one revenue average, the total cost decreases by 209550 US dollars. The partial coefficient 1.048604 of fuel cost says that; if the fuel cost increases by one US dollar, the total cost increases by 1.05 US dollars.

The intercept value of about 3141858, mechanically interpreted means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean total cost would be about 3141858 US dollars.

Table (4.7): Tests Results of Goodness of Fit of the Estimated Regression Model (4.2.4) of Marsland Aviation:

R-squared	0.990896
Prob(F-statistic)	0.000027
Durbin-Watson stat.	1.625457

Jarque-Bera-Normality Prob.	0.659185
Breusch-Godfrey Serial Correlation Chi-square Prob.	0.1463
Breusch-Pagan-Godfrey Heteroscedasticity Chi-square Prob.	0.864

Source: Prepared by the researcher.

The R-squared value of about 0.990896 is statistically a significant value (more than 60%), means that about 99% of variation in the total cost is explained by passenger, freight, load factor and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (1.625457) is found to be 2, so there is no first-order autocorrelation, either positive or negative. Also, the R-squared (0.990896) is less than Durbin-Watson statistic, which means this model is not spurious. **(Appendix C.2.1)**.

The probability of F-statistic equal to 0.000027 is statistically significant at level 5%, means that the independent variables: passenger, freight, load factor and fuel cost are jointly significant to influence the total cost.

The P-value of Jarque-Bera normality test equal to 0.659185 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are normally distributed. **(Appendix C.2.4)**.

The Chi square P-value of Breusch-Godfrey serial correlation LM test equal to 0.1463 is not statistically

significant at 5% level; so, we cannot reject the null hypothesis; that residuals are not serial correlation. **(Appendix C.2.5).**

The Chi square P-value of Breusch-Pagan-Godfrey heteroscedasticity test equal to 0.864 is not statistically significant value at 5% level, so we cannot reject the null hypothesis; that residuals are not heteroscedasticity. **(Appendix C.2.6).**

According to the above results, the researcher concludes that the residuals are normally distributed, not autocorrelated and homoscedastic; so, this result means that the estimated regression makes sense and is acceptable to predictive purposes and forecasting.

4.2.2.2. Forecasting:

Table (4.8): Result of Forecasting Sample 2004-2013 of Marsland Aviation:

Root Mean Squared Error	385280.7
Theil Inequality Coefficient	0.023852
Bias Proportion	0.00000

Source: Prepared by the researcher.

As shown in table (4.8), the root mean squared error is equal to 385280.7, while Theil Ineqauality coefficient equal to

0.023852, which is close to zero, that means, the predictive power of this model is very strong. Bias proportion is equal zero that means, there is no gap between the actual total cost and the predictive total cost, and they are moving closely. (Appendix C.2.7).

Graph (4.2): Forecasting Sample 2004-2013 of Marsland Aviation:
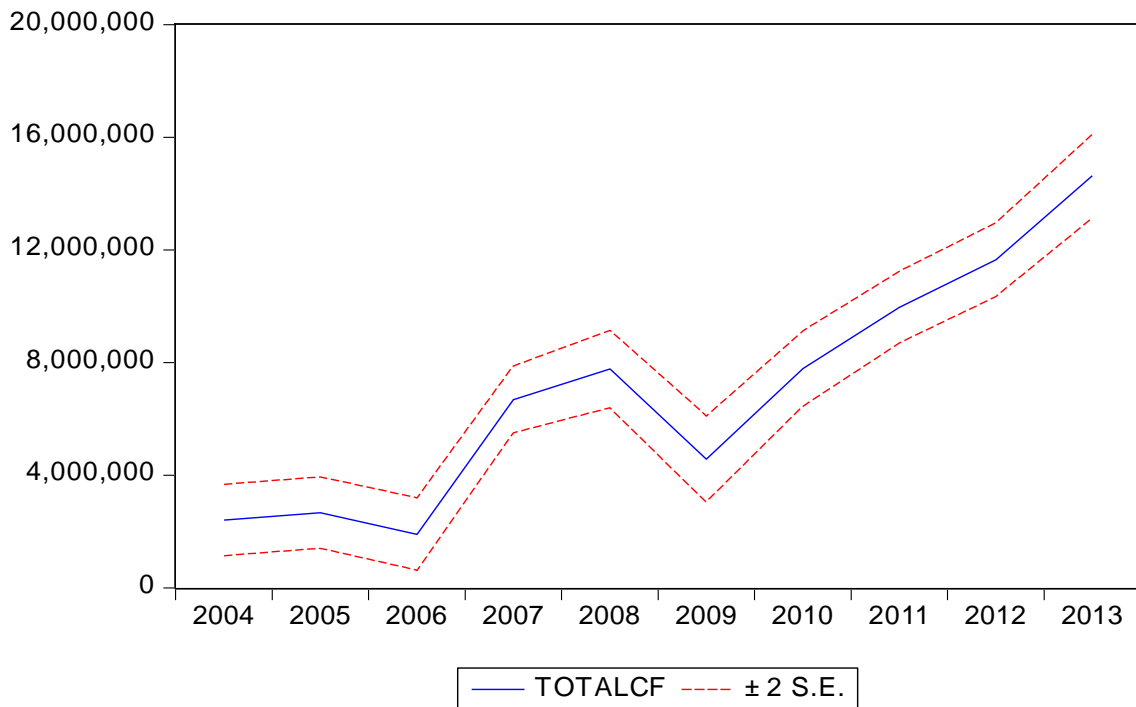
Source: Prepared by the researcher.

As shown in graph (4.2), the total cost value has been forecast and is passing throw 50% confidence interval; so, the forecasting of the total cost is significant and the ability of forecasting model is satisfactory.

4.2.3. Sun Air:

4.2.3.1. Evaluation of the Model:

According to the data in appendix (A); table (A.3), the estimated value of the total cost (CNLRM) is represented as follows:

Table (4.9): Significant Value of the Total Cost (TC) Regression Model (4.2.5) of Sun Air:

Variable	Coefficient	Prob.
C	93500.7	0.025
PAX	3.375324	0.0014
FRT	-0.446059	0.1774
LF	-107630.7	0.1007
FC	0.99644	0.0000

Source: Prepared by the researcher.

$$TC_t = 93500.7 + 3.375324PAX_t - 0.446059FRT_t - 107630.7LF_t + 0.99644FC_t$$

(4.2.5)

As shown in the table (4.9), there is 50% of the explanatory variables: passenger (PAX) and fuel cost (CF); are statistically significant at level 5%, to influence the dependent

variable; total cost (TC), **(Appendix C.3.1)**.

As expected, 3.375324 is the partial regression coefficient of passenger says that; if the number of passenger increases by one passenger per annum, the total cost increases by about 3.375 US dollars. The partial coefficient -0.446059 of freight says that; if the freight increases by hundred kgs, the total cost decreases by 44.6 US dollars. The partial coefficient -107630.7 of load factor says that; if the load factor increases by one revenue average, the total cost decreases by 107630.7 US dollars. The partial coefficient 0.99644 of fuel cost says that; if the fuel cost increases by one US dollar, the total cost increases by 0.99644 US dollars.

The intercept value of about 93500.7, mechanically interpreted means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean total cost would be about 93500.7 US dollars.

Table (4.10): Tests Results of Goodness of Fit of the Estimated Regression Model (4.2.5) of Sun Air:

R-squared	0.999844
Prob(F-statistic)	0.000000
Durbin-Watson stat.	1.88048

Jarque-Bera-Normality Prob.	0.611736
Breusch-Godfrey Serial Correlation Chi-square Prob.	0.8706
Breusch-Pagan-Godfrey Heteroscedasticity Chi-square Prob.	0.9278

Source: Prepared by the researcher.

The R-squared value of about 0.999844 is statistically significant value (more than 60%), means that about 99.98% of variation in the total cost is explained by passenger, freight, load factor and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (1.88048) is found to be 2, so there is no first-order autocorrelation, either positive or negative. Also, the R-squared (0.999844) is less than Durbin-Watson statistic, which means this model is not spurious. (Appendix C.3.1).

The probability of F-statistic equal to 0.000000 is statistically significant at level 5%, means that the independent variables: passenger, freight, load factor and fuel cost are jointly significant to influence the total cost.

The P-value of Jarque-Bera normality test equal to 0.611736 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are normally distributed. (Appendix C.3.4).

The Chi square P-value of Breusch-Godfrey serial correlation LM test equal to 0.8706 is not statistically

significant at 5% level; so, we cannot reject the null hypothesis; that residuals are not serial correlation. (Appendix C.3.5).

The Chi square P-value of Breusch-Pagan-Godfrey heteroscedasticity test equal to 0.9278 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are not heteroscedasticity. (Appendix C.3.6).

According to the above results, the researcher concludes that the residuals are normally distributed, not autocorrelated and homoscedastic; so, this result means that the estimated regression makes sense and is acceptable to predictive purposes and forecasting.

4.2.3.2. Forecasting:

Table (4.11): Result of Forecasting Sample 2004-2013 of Sun Air:

Root Mean Squared Error	22469.18
Theil Inequality Coefficient	0.003792
Bias Proportion	0.00000

Source: Prepared by the researcher.

As shown in table (4.11), the root mean squared error is equal to 22469.18, while Theil Ineqauality coefficient equal to

0.003792, which is close to zero, that means, the predictive power of this model is very strong. Bias proportion is equal to zero, that means, there is no gap between the actual total cost and the predictive total cost, and they are moving closely. (Appendix C.3.1).

Graph (4.3): Forecasting Sample 2004-2013 of Sun Air:

Source: Prepared by the researcher.

As shown in graph (4.3), the total cost value has been forecast and is passing throw 50% confidence interval, which is very small intervals; so, the forecasting of the total cost is significant, and the ability of forecasting model is very strong.

135

4.2.4. Nova Air:

4.2.4.1. Evaluation of the Model:

According to the data in appendix (A); table (A.4), the estimated value of the total cost (CNLRM) is represented as follows:

Table (4.12): Significant Value of the Total Cost (TC) Regression Model (4.2.6) of Nova Air:

Variable	Coefficient	Prob.
C	352633.7	0.0005
PAX	0.988587	0.0000
FRT	-0.073409	0.7052
LF	-409289.9	0.0111
FC	6.793882	0.0109

Source: Prepared by the researcher.

$$TC_t = 352633.7 + 0.988587PAX_t - 0.073409FRT_t - 409289.9LF_t + 6.793882FC_t$$

(4.2.6)

As shown in the table (4.12), there is 75% of the explanatory variables: passenger (PAX), load factor (LF) and fuel cost (CF); are statistically significant at level 5%, to influence the dependent variable; total cost (TC). (Appendix

136

C.4.1).

As expected, 0.988587 is the partial regression coefficient of passenger says that; if the number of passenger increases by one passenger per annum, the total cost increases by about 0.988587 US dollars. The partial coefficient -0.073409 of freight says that; if the freight increases by hundred kgs, the total cost decreases by 7.34 US dollars. The partial coefficient -409289.9 of load factor says that; if the load factor increases by one revenue average; the total cost decreases by 409289.9 US dollars. The partial coefficient 6.793882 of fuel cost says that; if the fuel cost increases by one US dollar, the total cost increases by 6.793882 US dollars.

The intercept value of about 352633.7, mechanically interpreted means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean total cost would be about 352633.7 US dollars.

R-squared	0.99996
Prob(F-statistic)	0.000000
Durbin-Watson stat.	1.53558
Jarque-Bera-Normality Prob.	0.87321

Breusch-Godfrey Serial Correlation Chi-square Prob.	0.4533
Breusch-Pagan-Godfrey Heteroscedasticity Chi-square Prob.	0.3192

Table (4.13): Tests Results of Goodness of Fit of the Estimated Regression Model (4.2.6) of Nova Air:

Source: Prepared by the researcher.

The R-squared value of about 0.999962 is statistically significant value (more than 60%), means that about 99.996% of variation in the total cost is explained by passenger, freight, load factor and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (1.53558) is found to be 2, so there is no first-order autocorrelation, either positive or negative. Also, the R-squared (0.99996) is less than Durbin-Watson statistic, which means this model is not spurious, (Appendix C.4.1).

The probability of F-statistic equal to 0.000000 is statistically significant at level 5%, means that the independent variables: passenger, freight, load factor and fuel cost are jointly significant to influence the total cost.

The P-value of Jarque-Bera normality test equal to 0.87321 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are normally distributed. (Appendix C.4.4).

The Chi square P-value of Breusch-Godfrey serial

correlation LM test equal to 0.4533 is not statistically significant at 5% level; so, we cannot reject the null hypothesis; that residuals are not serial correlation. **(Appendix C.4.5).**

The Chi square P-value of Breusch-Pagan-Godfrey heteroscedasticity test equal 0.3192 is not statistically significant value at 5% level, so we cannot reject the null hypothesis; that residuals are not heteroscedasticity. **(Appendix C.4.6).**

According to the above results, the researcher concludes that the residuals are normally distributed, not autocorrelated and homoscedastic; so, this result means that the estimated regression makes sense and is acceptable to predictive purposes and forecasting.

4.2.4.2. Forecasting:

Table (4.14): Result of Forecasting Sample 2004-2013 of Nova Air:

Root Mean Squared Error	**15604.13**
Theil Inequality Coefficient	**0.001995**
Bias Proportion	0.00000

Source: Prepared by the researcher.

As shown in table (4.14), the root mean squared error

139

is equal to 15604.13, while Theil Ineqauality coefficient equal to 0.001995, which is close to zero, that means, the predictive power of this model is very strong. Bias proportion is equal to zero, that means, there is no gap between the actual total cost, and the predictive total cost, and they are moving closely. (Appendix C.4.7).

Graph (4.4): Forecasting Sample 2004-2013 of Nova Air:
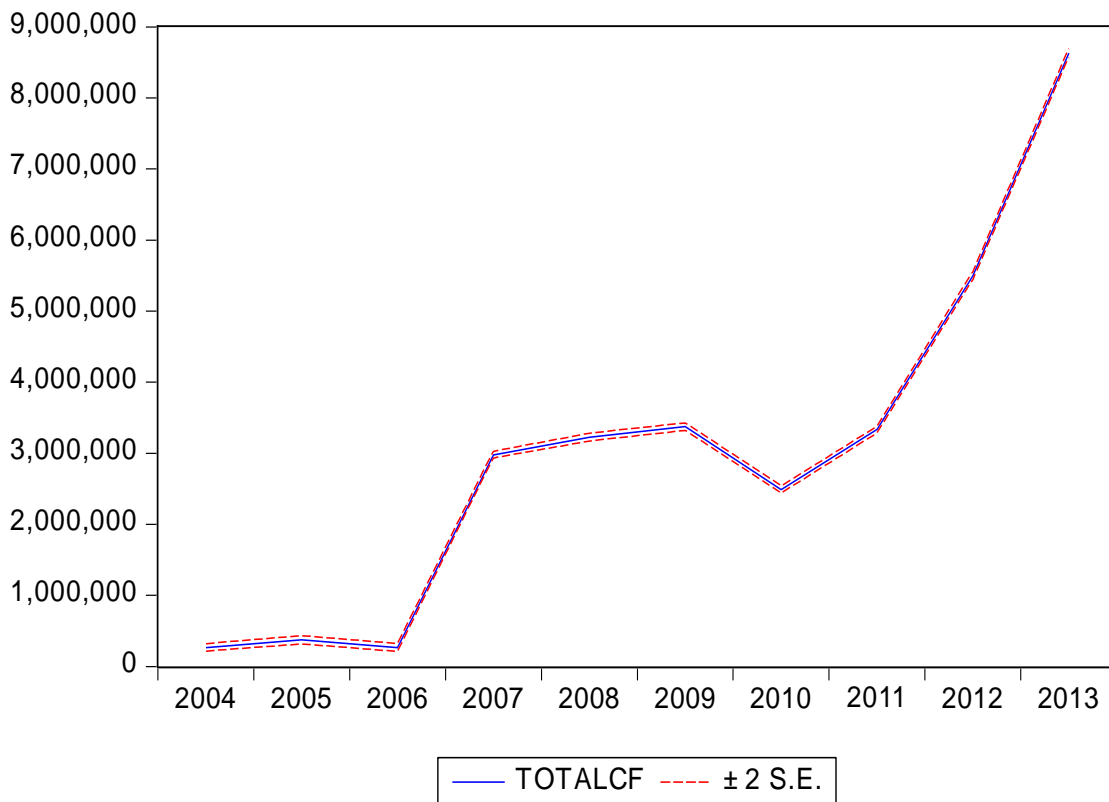
Source: Prepared by the researcher.

As shown in graph (4.4), the total cost value has been forecast and is passing throw 50% confidence interval with minimum square error; which is very small intervals; so, the

forecasting of the total cost is significant, and the ability of forecasting model is very strong.

4.2.5. Mid Airlines:

4.2.5.1. Evaluation of the Model:

According to the data in appendix (A); table (A.5), the estimated value of the total cost (CNLRM) is represented as follows:

Table (4.15): Significant Value of the Total Cost (TC) Regression Model (4.2.7) of Mid Airlines:

Variable	Coefficient	Prob.
C	229609.5	0.0000
PAX	5.772532	0.0000
LF	-242048.6	0.0000
FC	0.997609	0.0000

Source: Prepared by the researcher.

$$TC_t = 229609.5 + 5.772532PAX_t - 242048.6LF_t + 0.997609FC_t$$

(4.2.7)

Note that, the estimated model of the total cost of Mid Airlines uses sample size eight years (2004-2011) instead of

ten years, because their activity was stopped in since 2011. Also the researcher did not include the Freight (FRT) in the model, because Mid Airlines did not work in this field.

As shown in the table (4.15); there is 100% of the explanatory variables: passenger (PAX), load factor (LF) and fuel cost (CF); are statistically significant at level 5%; to influence the dependent variable; total cost (TC), **(Appendix C.5.1)**.

As expected, 5.772532 is the partial regression coefficient of passenger says that; if the number of passenger increases by one passenger per annum, the total cost increases by about 5.772532 US dollars. The partial coefficient -242048.6 of load factor says that; if the load factor increases by one revenue average, the total cost decreases by 242048.6 US dollars. The partial coefficient 0.997609 of fuel cost says that; if the fuel cost increases by one US dollar, the total cost increases by 0.997609 US dollars.

The intercept value of about 229609.5, mechanically interpreted means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean total cost would be about 229609.5 US dollars.

R-squared	0.999996

Prob(F-statistic)	0.000000
Durbin-Watson stat.	3.215065
Jarque-Bera-Normality Prob.	0.636362
Breusch-Godfrey Serial Correlation Chi-square Prob.	0.038
Breusch-Pagan-Godfrey Heteroscedasticity Chi-square Prob.	0.2755

Table (4.16): Tests Results of Goodness of Fit of the Estimated Regression Model (4.2.7) of Mid Airlines:

Source: Prepared by the researcher.

The R-squared value of about 0.999996 is statistically significant value (more than 60%), means that about 99.9996% of variation in the total cost is explained by passenger, freight, load factor and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (3.215065) is found to be 3 (less than 4), so there is no first-order autocorrelation, either positive or negative. Also, the R-squared (0.999996) is less than Durbin-Watson statistic, which means this model is not spurious. (Appendix C.5.1).

The probability of F-statistic equal to 0.000000 is statistically significant at level 5%, means that the independent variables: passenger, freight, load factor and fuel cost are jointly significant to influence the total cost.

The P-value of Jarque-Bera normality test equal to

0.636362 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis, that residuals are normally distributed. (Appendix C.5.4).

The Chi square P-value of Breusch-Godfrey serial correlation LM test equal to 0.038 is t statistically significant at 5% level; so, we can reject the null hypothesis; that residuals are serial correlation. (Appendix C.5.5).

The number of runs in Run Test equal 7 runs at an interval of (2.43376 , 7.56624) is statistically significant at 95% confidence interval; so, we cannot reject the null hypothesis of randomness, and the residuals are not serial correlation. The researcher concluded that the residuals were not autocorrelated and were contradicting the result when using the method of Breusch-Godfrey serial correlation LM test due to the very small sample used.

The Chi square P-value of Breusch-Pagan-Godfrey heteroscedasticity test equal to 0.2755 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are not heteroscedasticity. (Appendix C.5.6).

According to the above results, the researcher concludes that the residuals are normally distributed, not autocorrelated and homoscedastic; so, this result means that the estimated regression makes sense and is acceptable to

144

predictive purposes and forecasting.

4.2.5.2. Forecasting:

Table (4.17): Result of Forecasting Sample 2004-2011of Mid Airlines:

Root Mean Squared Error	1039.812
Theil Inequality Coefficient	0.000505
Bias Proportion	0.00000

Source: Prepared by the researcher.

As shown in table (4.17), the root mean squared error is equal to 1039.812, while Theil Ineqauality coefficient equal to 0.000505, which is close to zero, that means, the predictive power of this model is very strong. Bias proportion is equal to zero, that means, there is no gap between the actual total cost and the predictive total cost, and they are moving closely. (Appendix C.5.7).

Graph (4.5): Forecasting Sample 2004-2011of Mid Airlines:
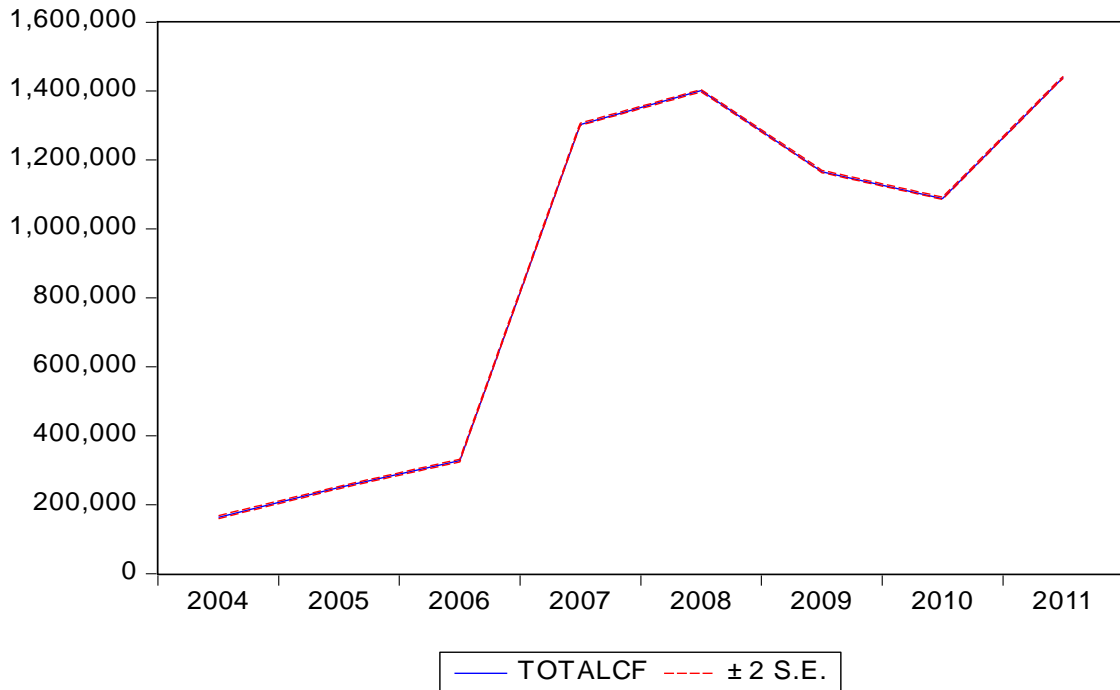
As shown in graph (4.5), the total cost value has been forecast and is applicable on 50% confidence interval; so, the forecasting of the total cost is significant, and the ability of forecasting model is very strong.

4.2.6. Badr Airlines:

4.2.6.1. Evaluation of the Model:

According to the data in appendix (A); table (4.6), the estimated value of the total cost (CNLRM) is represented as follows:

Table (4.18): Significant Value of the Total cost (TC) regression Model (4.2.8) of Badr Airlines:

Variable	Coefficient	Prob.
C	5667100	0.0000
PAX	53.89931	0.0226
FRT	−0.063447	0.6302
LF	−8774699	0.002
FC	1.085344	0.0000

Source: Prepared by the researcher.

$$TC_t = 5667100 + 53.89931\text{PAX}_t - 0.063447\text{FRT}_t - 8774699\text{LF}_t + 1.085344 FC_t$$

(4.2.8)

As shown in the table (4.18), there is 75% of the explanatory variables: passengers (PAX), load factor (LF) and fuel cost (FC); are statistically significant at level 5%, to influence the dependent variable; total cost (TC), **(Appendix C.6.1)**.

As expected, 53.89931 is the partial regression coefficient of passenger says that; if the number of passenger increases by one passenger per annum, the total cost increases by 53.89931 US dollars. The partial coefficient −0.063447 of freight says that; if the freight increases by hundred kgs, the total cost decreases by 6.4 US dollars. The partial coefficient −8774699 of load factor says that; if the load factor increases by one revenue average, the total cost

147

decreases by 8774699 US dollars. The partial coefficient 1.085344 of fuel cost says that; if the fuel cost increases by one US dollar, the total cost increases by 1.085344 US dollars.

The intercept value of about 5667100, mechanically interpreted means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean total cost would be about 5667100 US dollars.

Table (4.19): Tests Results of Goodness of Fit of the Estimated

R-squared	0.995109
Prob(F-statistic)	0.000006
Durbin-Watson stat.	1.976693
Jarque-Bera-Normality Prob.	0.510203
Breusch-Godfrey Serial Correlation Chi-square Prob.	0.4906
Breusch-Pagan-Godfrey Heteroscedasticity Chi-square Prob.	0.4406

Regression Model (4.2.8):

Source: Prepared by the researcher.

The R-squared value of about 0.995109 is statistically significant value (more than 60%), means that about 99.5% of variation in the total cost is explained by passenger, freight, load factor and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (1.976693)

148

is found to be 2, so there is no first-order autocorrelation, either positive or negative. Also the R-squared (0.995109) is less than Durbin-Watson statistic, which means this model is not spurious. (Appendix C.6.1).

The probability of F-statistic equal to 0.000006 is statistically significant at level 5%, means that the independent variables: passenger, freight, load factor and fuel cost are jointly significant to influence the total cost.

The P-value of Jarque-Bera normality test equal to 0.510203 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are normally distributed. (Appendix C.6.4).

The R-squared P-value of Breusch-Godfrey serial correlation LM test equal to 0.4906 is not statistically significant at 5% level; so, we cannot reject the null hypothesis; that residuals are not serial correlation. (Appendix C.6.5).

The R-squared P-value of Breusch-Pagan-Godfrey heteroscedasticity test equal to 0.4406 is not statistically significant value at 5% level; so, we cannot reject the null hypothesis; that residuals are not heteroscedasticity. (Appendix C.6.6).

According to the above results, the researcher concludes that the residuals are normally distributed, not autocorrelated and homoscedastic; so, this result means that

the estimated regression makes sense and is acceptable to predictive purposes and forecasting.

4.2.6.2. Forecasting:

Table (4.20): Result of Forecasting Sample 2004-2013 of Badr Airlines:

Root Mean Squared Error	418794.4
Theil Inequality Coefficient	0.017145
Bias Proportion	0.00000

Source: Prepared by the researcher.

As shown in table (4.20), the root mean squared error is equal to 418794.4, while Theil Ineqauality coefficient equal to 0.017145, which is close to zero, that means, the predictive power of this model is very strong. Bias proportion is equal to zero, that means, there is no gap between the actual total cost and the predictive total cost, and they are moving closely. (Appendix C.6.7).

Graph (4.6): Forecasting Sample 2004-2013 of Badr Airlines:
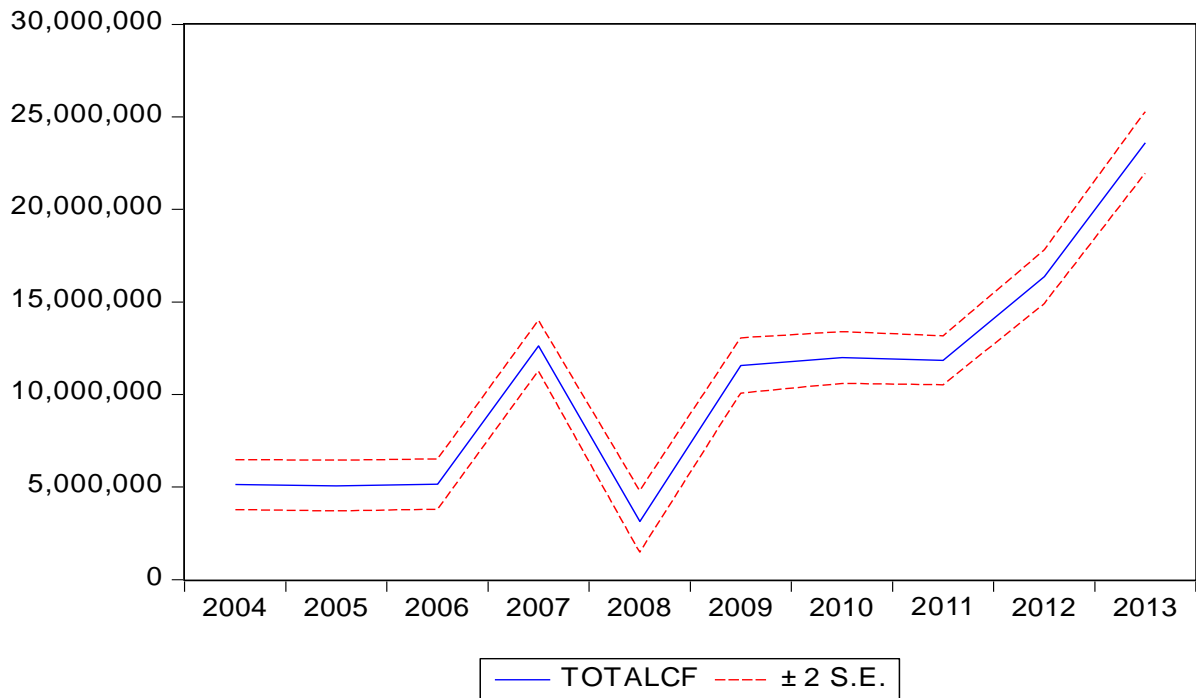
Source: Prepared by the researcher.

As shown in graph (4.6), the total cost value has been forecast and is passing throw 50% confidence interval; so, the forecasting of the total cost is significant and the ability of forecasting model is satisfactory.

Table (4.21): Tests Results of Goodness of Fit of the Estimated CNLRM for the Sudanese Airlines:

An Airline	R-Squared	Prob. (F-Statistic)	Durbin-Waston Stat.	Jarque-Bera-Normality Prob.	Breusch-Godfrey Serial Correlation Chi-Square Prob.	Breusch-Pagan-Godfrey Heterosced a-sticity Chi-Square Prob.
Sudan Airways	0.990847	0.000002	2.332225	0.613088	0.2347	0.6117
Marsland Aviation	0.990896	0.000027	1.625457	0.659185	0.1463	0.864
Sun Air	0.999844	0.0000	1.88048	0.611736	0.8706	0.9278
Nova Air	0.99996	0.0000	1.53558	0.87321	0.4533	0.3192
Mid Air	0.999996	0.0000	3.215065	0.636362	0.038	0.2755
Badr Airlines	0.995109	0.000006	1.976693	0.510203	0.4906	0.4406

As shown in the table (4.21), the Classical Normal Linear Regression Model (CNLRM) is acceptable to the predictive purpose of forecasting the function of total cost of each airline, with a high statistically significant value of R-squared (99%), and statistically significant values of F-statistic probability between (0.000 - 0.00027) at level 5%. Additionally, the residuals were Normally distributed; P-values of Jarque-Bera Normality Test between (0.510203 - 0.87321) are not statistically significant values at 5% level, and also the residuals were not autocorrelation (not serial correlation); most of R-squared P-values of Breusch-Godfrey Serial Correlation; LM Test between (0.1463 - 0.8706) are not statistically significant values at 5% level, also the residuals were homoscedastic; R-

squared P-values of Breusch-Pagan-Godfrey Heteroscedasticity Test between (0.2755 - 0.9278); are not statistically significant values at 5% level.

Table (4.22): Result of Forecasting Sample 2004-2013 of Airlines:

An Airline	Root Mean Squared Error	Theil Inequality Coefficient	Bias Proportion
Sudan Airways	615993.2	0.034086	0.0000
Marsland Aviation	385280.7	0.023852	0.0000
Sun Air	22469.18	0.003792	0.0000
Nova Air	15604.13	0.001995	0.0000
Mid Air	1039.812	0.000505	0.0000
Badr Airlines	418794.4	0.017145	0.0000

Source: Prepared by the researcher.

As shown in the table (4.22), the Theil Ineqauality coefficients values between (0.000505 - 0.034086) are close to zero and zero Bias Proportions for all airlines in sample research. These results mean that the estimated of CNLRM models make sense; with strong power for prediction and forecast.

4.2.7. Forecasting of the Airlines Total Cost:

By using the Simple Exponential Smoothing (smoothing parameter equals 0.7), the researcher forecasted the data of the explanatory variables, and then used the results to forecast the total cost of each airline from 2014 to 2018, by using CNLRM which is represented in the following table:

Table (4.23): Forecasting of the Airlines Total Cost (US $) From 2014-2018:

Year	Sudan Airways	Marsland Aviation	Sun Air	Nova Air	Mid Air	Badr Airlines
2014	18496993.66	13498856.00	2266256.06	14879891.47	1344394.04	3457151.67
2015	8952143.90	6248957.40	745327.31	4710811.03	564044.86	5004115.50
2016	6088688.97	4073987.82	289048.68	1660086.90	329940.11	5468204.65
2017	5229652.49	3421496.95	152165.09	744869.66	259708.68	5607431.40
2018	4971941.55	3225749.68	111100.02	470304.49	238639.25	5649199.42

Source: Prepared by the researcher.

The forecast table above (4.23), shows that the total cost of the subject sample airlines, except Bader Aviation, shall gradually decrease during the next five years. This is clearly noted in the airlines activities as reflected in their current actual status in 2016, due to the decreasing number of their fleet that minimizes their activities &, henceforth, their total cost, as

154

reflected by Mid Air & Marsland Aviation that are now closed down and are out of business.

The subject forecast above shows an increase in the total cost of Badr Airlines only, as reflected in the current actual status of 2016, due to its remarkable increasing fleet number and operational activities and routes.

4.3. ANALYSIS OF DATA BY USING PANEL DATA REGRESSION MODELS:

The researcher analyzed the data of five Sudanese airlines for the same period (2004-2013) for a total of 50 balanced long Panel Data observations. So, Mid Airlines was not included in the panel analysis because its sample size was eight years (2004-2011) instead of ten years.

4.3.1. Pooled OLS Regression Model:

Here, the researcher simply pooled all 50 observations and estimated a grand regression, neglecting the cross-section and time series nature of data. Consider the following model:

$$TC_{it} = \beta_0 + \beta_1 PAX_{it} + \beta_2 FRT_{it} + \beta_3 LF_{it} + \beta_4 FC_{it} + u_{it} \quad (4.3.1)$$

$$i = 1, 2, \dots, 5 \quad , \quad t = 1, 2, \dots, 10$$

Where i is ith subject and t is the time period for the variables

155

Table (4.24): Significant Value of the Total Cost (TC) Pooled OLS Regression Model (4.3.2):

Variable	Coefficient	Prob.
C	3320956	0.0000
PAX	0.864041	0.0000
FRT	0.228435	0.0743
LF	-3530482	0.0000
FC	1.161783	0.0000

Source: Prepared by the researcher.

$$TC_{it} = 3320956 + 0.864041\text{PAX}_{it} + 0.228435\text{FRT}_{it} - 3530482\text{LF}_{it} + 1.161783FC_{it} \quad (4.3.2)$$

As shown in the table (4.24), there is 75% of the explanatory variables: passengers (PAX), load factor (LF) and fuel cost (FC); are statistically significant at level 5%, to influence the dependent variable; total cost (TC). (Appendix C.7).

As expected, 0.864041 which is the partial regression coefficient of passenger says that; if the number of passenger increases by one passenger per annum, the total cost increases by 0.864041 US dollars. The partial coefficient 0.228435 of freight says that; if the freight increases by hundred kgs, the total cost increases by 23 US dollars. The partial coefficient -3530482 of load factor says that; if the load factor increases by one revenue average, the total cost decreases by 3530482 US

156

dollars. The partial coefficient 1.161783 of fuel cost says that; if the fuel cost increases by one US dollar; the total cost increases by 1.161783 US dollars.

The intercept value of about 3320956, mechanically interpreted means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean of total cost would be about 3320956 US dollars.

Table (4.25): Tests Results of Goodness of Fit of the Estimated

R-squared	0.941944
Prob(F-statistic)	0.000000
Durbin-Watson stat.	1.023638

Pooled OLS Regression Model (4.3.2):

Source: Prepared by the researcher.

The R-squared value of about 0.941944 is statistically significant value (more than 60%), means that about 94% of variation in the total cost is explained by passenger, freight, load factor and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (1.023638) is found to be 1 (more than zero); so, there is no first-order autocorrelation, either positive or negative. Also, the R-squared (0.941944) is less than Durbin-Watson statistic, which means this model is not spurious, and suggesting that there is no

157

autocorrelation or partial correlation in the data. (Appendix C.7).

The probability of F-statistic (182.5269) equal to 0.000000 is statistically significant at level 5%, means that the independent variables: passenger, freight, load factor and fuel cost are jointly significant to influence the total cost.

In spite of all good results, there is a major problem of this model, that, it does not distinguish between various airlines, nor does it tell us whether the response of the total cost to the explanatory variables over time is the same for all the airlines.

4.3.2. The Fixed Effects Least Squares Dummy Variable (LSDV):

Here the researcher pooled all 50 observations, but allowed each cross-section unit (airline) to have its own (intercept) dummy variable. Consider the following model:

$$TC_{it} = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \alpha_5 D_{5i} + \beta_2 PAX_{it} +$$
$$\beta_3 FRT_{it} + \beta_4 LF_{it} + \beta_5 FC_{it} + u_{it} \quad (4.4.3)$$

Where:

$D_{2i} = 1$, for airline 2, 0 otherwise.

$D_{3i} = 1$, for airline 3, 0 otherwise.

And so on..

The researcher was treating airline 1 (Sudan airways) as the base (reference) category, so the α_1 is the intercept

value of Sudan airways and other α coefficients represent the difference from the intercept value of the first airline Sudan Airways. The airlines arrangement in the model as follows:

Airline1= Sudan Airways (the reference).

Airline 2= Marsland Aviation.

Airline 3= Sun Air.

Airline 4= Nova Air.

Airline 5= Badr Airlines.

Table (4.26): Significant Value of the Total Cost (TC) Fixed Effects Least Squares Dummy Variable (LSDV) Model (4.3.4):

Variable	Coefficient	Prob.
C	1482810	0.0332
PAX	0.946771	0.0000
FRT	-0.00083	0.9938
LF	-1617728	0.015
FC	1.154968	0.0000
D_2	1151300	0.015
D_3	-600230	0.2145
D_4	151433.4	0.7973
D_5	3109517	0.0000

$$TC_{it} = 1482810 + 1151300D_{2i} - 600230D_{3i} + 151433.4D_{4i} + 3109517D_{5i} + 0.946771PAX_{it} - 0.00083FRT_{it} - 1617728LF_{it} + 1.154968FC_{it} + u_{it}$$

(4.3.4)

As shown in the table (4.26), there is 75% of the explanatory variables: passengers (PAX), load factor (LF) and fuel cost (FC); are statistically significant at level 5%, to influence the dependent variable; total cost (TC), (Appendix C.8).

As expected, in Sudan airways, 0.946771 is the partial regression coefficient of passenger says that; if the number of passenger increases by one passenger per annum; the total cost increases by 0.946771 US dollars. The partial coefficient -0.00083 of freight says that; if the freight increases by thousand kgs, the total cost decreases by 0.83 US dollars. The partial coefficient -1617728 of load factor says that; if the load factor increases by one revenue average, the total cost decreases by 1617728 US dollars. The partial coefficient 1.154968 of fuel cost says that; if the fuel cost increases by one US dollar; the total cost increases by 1.154968 US dollars.

The intercept value in Sudan airways is about 1482810, mechanically interpreted, means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean of total cost would be about 1482810 US dollars.

As these regressions show, the mean of total cost in Sudan Airways is about 1482810 US dollars, that of total cost in Marsland Aviation is higher about 1151300 US dollars than the mean total cost of Sudan Airways as benchmark category, with actual mean about

2634110 US dollars.

By contrast, the total cost of Sun Air is lower about 600230 US dollars, for an actual mean total cost 882580 US dollars.

But, the total cost of Nova Air is higher about 151433.4 US dollars, for an actual mean total cost 1634243.4 US dollars.

Also, the total cost of Badr Airlines is higher about 3109517 US dollars, for an actual mean total cost 4592327US dollars.

Table (4.27): Tests Results of Goodness of Fit of the Estimated

R-squared	0.974698
Prob(F-statistic)	0.000000
Durbin-Watson stat.	2.209698

Fixed Effects Least Squares Dummy Variable (LSDV) Model (4.3.4):

Source: Prepared by the researcher.

The R-squared value of about 0.974698 is a statistical significant value (more than 60%), means that about 98% of variation in the total cost is explained by passenger, freight,

load factor and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (2.209698) is found to be 2, so, there is no first-order autocorrelation, either positive or negative. Also, the R-squared (0.974698) is less than Durbin-Watson statistic, which means that this model is not spurious, and suggesting that there is no autocorrelation or partial correlation in the data. **(Appendix C.8)**.

The probability of F-statistic (197.4250) equal to 0.0000 is statistically significant at level 5%, means that the independent variables: passenger, freight, load factor and fuel cost are jointly significant to influence the total cost. **(Appendix C.8)**.

According to above result, it seems that the Fixed Effect Least Squared Dummy Variable (LSDV) model is better than the Pooled OLS model; so, to check this result, the researcher used the Wald Test, that depends on the F-test approach (Restricted least squares).

The F-statistic of Wald test equals 13.26867 with a probability value equal to 0.0000, which is a high statistical significant value at 5% level, **(Appendix C.12)**. So, we can reject the null hypothesis; that is to say, all the dummy variables are equal to zero, which are represented in the Pooled OLS model, and accept the alternative hypothesis which says that the Fixed Effect Least Squares Dummy variable (LSDV) is appropriate.

4.3.3. The Fixed Effect Within-Group (WG) Estimator:

The researcher used this model for each airline and then pooled all 50 mean-corrected values and ran an OLS regression. Letting tc_{it} , pax_{it} , frt_{it} , lf_{it} and fc_{it} represent the mean-corrected values and ran the following regression:

$$tc_{it} = \beta_2 pax_{it} + \beta_3 frt_{it} + \beta_4 lf_{it} + \beta_5 fc_{it} + u_{it} \quad (4.3.5)$$

Where,

$$i = 1, 2, \dots, 5 \qquad t = 1, 2, \dots 10$$

Table (4.28): Significant Value of the Total Cost (TC) Fixed Effect Within-Group (WG) Model (4.3.6):

Variable	Coefficient	Prob.
C	2245214	0.0001
PAX	0.946771	0.0000
FRT	-0.00083	0.9938
LF	-1617728	0.015
FC	1.154968	0.0000

Source: Prepared by the researcher.

163

$$tc_{it} = 0.946771pax_{it} - 0.00083frt_{it} - 1617728lf_{it} + 1.154968fc_{it} + u_{it}$$

$$(4.3.6)$$

As shown in Table (4.28), the researcher observed that the slope coefficients of the TC, PAX, FRT, LF and FC were identical with slope coefficients of (LSDV) model, because mathematically, the two models are identical. (Appendix C.9).

Table (4.29): Tests Results of Goodness of Fit of the Estimated Fixed Effects Within-Group (WG) Estimator Model (4.3.6):

R-squared	0.974698
Prob(F-statistic)	0.000000
Durbin-Watson stat.	2.209698

Source: Prepared by the researcher.

The researcher obtained the estimates of the intercepts using the WG method by subtracting from the mean value of the dependent variable the mean values of the explanatory variables for the airline times the estimated slope coefficients from the WG estimators.

By using the following equation:

$$\hat{a}_i = \overline{TC}_i - (0.946771) * \overline{PAX}_i - (-0.00083) * \overline{FRT}_i - (-1617728) * \overline{LF}_i - (1.154968) * \overline{FC}_i$$

$$(4.3.7)$$

Also, we can study the effect of an airline on the total cost to calculate the intercept values of five entities that are given in the regression result, as shown in the following table:

Table (4.30): **The Cross-Section Random Effects** Represent Effect of Airline on Total Cost in US Dollar:

Firm	Airline	Effect
1.	Sudan Airways	−762404
2.	Marsland Aviation	388895.5
3.	Sun Air	−1362634
4.	Nova Air	−610971
5.	Badr Airlines	2347113

Source: Prepared by the researcher.

The intercept value is 2245214. By using the differential intercept values of the five entities that are given in table (4.30), Firm number 1 (Sudan Airways) has intercept value which is 762404 US dollars lower than the common intercept value of 2245214; the actual value of the intercept Sudan Airways is 1482810 US dollars. On the other hand, the intercept value of firm number 2 (Marsland Aviation) is higher by 388895.5 US dollars than the common intercept value; the actual

intercept value for Marsland Aviation is 2634110 US dollars. The other intercepts values for the other airlines, similarly, are shown in the following table:

Table (4.31): **The Intercept Values** Represent an Actual Mean Total Cost of Sudanese Airlines in US Dollar:

Airline	Intercept
Sudan Airways	**1,482,810**
Marsland Aviation	2,634,110
Sun Air	882,580
Nova Air	1,634,243
Badr Airlines	4,592,327

Source: Prepared by the researcher.

4.3.4. The Random Effects Model (REM):

The researcher applied the following model:

$$TC_{it} = \beta_1 + \beta_2 PAX_{it} + \beta_3 FRT_{it} + \beta_4 LF_{it} + \beta_5 FC_{it} + w_{it} \quad (4.3.8)$$

Where,

$$w_{it} = \varepsilon_i + u_{it}$$

$$i = 1, 2, \dots, 5 \qquad t = 1, 2, \dots 10$$

The composite error term w_{it} consists of two components:

ε_i : Which is the cross-section or individual-specific error component.

166

$\boldsymbol{u_{it}}$: Which is the combined time series and cross-section error component and is sometimes called idiosyncratic term, because it varies over cross-section as well as time.

Table (4.32): Significant value of the Total Cost (TC) Random Effects Model (REM) (4.3.9):

Variable	Coefficient	Prob.
C	3320956	0.0000
PAX	0.864041	0.0000
FRT	0.228435	0.0113
LF	-3530482	0.0000
FC	1.161783	0.0000

Source: Prepared by the researcher.

The estimated value of the total cost by random effects model is represented as follows:

$$TC_{it} = 3320956 + 0.864041PAX_{it} + 0.228435FRT_{it} - 3530482LF_{it} + 1.161783FC_{it} + w_{it}$$

$$(4.3.9)$$

As shown in the table (4.32), all the differential intercept coefficients are individually highly statistically

significant: PAX, FRT, LF and FC; are statistical significant at level 5%, to influence the dependent variable; total cost (TC). (Appendix C.10).

As expected, 0.864041 is the partial regression coefficient of passenger says that; if the number of passenger increases by one passenger per annum, the total cost increases by 0.864041 US dollars. The partial coefficient 0.228435 of freight says that; if the freight increases by hundred kgs, the total cost increases by 22.8 US dollars. The partial coefficient −3530482 of load factor says that; if the load factor increases by one revenue average, the total cost decreases by 3530482 US dollars. The partial coefficient 1.161783 of fuel cost says that; if the fuel cost increases by one US dollar, the total cost increases by 1.161783 US dollars.

The intercept value of about 3320956, mechanically interpreted, means that; if the values of passenger, freight, load factor and fuel cost were fixed at zero, the mean of total cost would be about 3320956 US dollars.

Table (4.33): Tests Results of Goodness of Fit of the Estimated

R-squared	0.941944
Prob(F-statistic)	0.000000
Durbin-Watson stat.	1.023638

Random Effects Model (REM) (4.3.9):

The R-squared value of about 0.941944 is statistically significant value (more than 60%), means that about 94% of variation in the total cost is explained by passenger, freight, load factor and fuel cost; that means the goodness of fit of the regression line is very high. Durbin-Watson statistic (1.023638) is found to be 1 (more than zero); so, there is no first-order autocorrelation, either positive or negative. Also, the R-squared (0.995109) is less than Durbin-Watson statistic, which means this model is not spurious, and suggesting that there is no autocorrelation or partial correlation in the data. (Appendix C.10).

The probability of F-statistic (182.5269) equals 0.0000 is statistically significant at level 5%, means that the independent variables: passenger, freight, load factor and fuel cost are jointly significant to influence the total cost. (Appendix C.10).

By comparing the results of the random effect REM and pooled OLS regression; the researcher found no difference between the two; so, there was some doubt on the results. According to this result, it seems that the fixed effect within group (WG) is better than the random effects model (REM); so to check this result, the researcher used Hausman Test; the test statistic was developed by Hausman that has an asymptotic χ^2 distribution.

The Chi-square statistic value for 4 degrees of freedom of Hausman Test equals 53.074696 with probability value equal to 0.0000, which is a highly statistically significant value at 5% level, **(Appendix C.11)**. So, we can reject the null hypothesis; that is to say, that the REM is appropriate, and we accept the alternative hypothesis which says that the Fixed Effects within group (WG) is appropriate.

Table (4.34): Tests Results of Goodness of Fit of the Estimated
Panel Data Regression Models:

Panel Data Regression Models	R-Squared	Prob. (F-Statistic)	Durbin-Waston Stat.
OLS	0.941944	0.0000	1.023638
LSDV	0.974698	0.0000	2.209698
WG	0.974698	0.0000	2.209698
REM	0.941944	0.0000	1.023638

Source: Prepared by the researcher.

By comparing the results of the Panel Regression Models, as shown in table (4.34): Pooled OLS Regression Model, Fixed Effects Least Squares Dummy Variable (LSDV), Fixed Effect within–Group (WG) Estimator and Random Effects Model (REM), the researcher found that there was no difference between (REM) and Pooled (OLS) Regression Model. In the other hand, (LSDV) and (WG) Models were identical. So according to Hausman Test and Wald Test, the researcher concludes that the Fixed effects model is appropriate.

The next chapter provides conclusions and recommendations regarding the future of Sudan domestic air transport.

CHAPTER FIVE

CONCLUSIONS & RECOMMENDATIONS

5 - 0   INTRODUCTION.

5 - 1   CONCLUSIONS.

5 - 2   RECOMMENDATIONS.

5 - 3   FUTURE STUDIES.

CHAPTER FIVE

CONCLUSIONS & RECOMMENDATIONS

5. <u>INTRODUCTION:-</u>

This chapter displays conclusions of this research, according to the data obtained from Sudan Civil Aviation Authority, to determine a cost function for each airline during (2004-2013); and finally furnish recommendations and future studies.

5.1. <u>CONCLUSIONS:-</u>

- The Classical Normal Linear Regression Model (CNLRM) is acceptable to the predictive purpose of forecasting the function of total cost of each airline, with a high statistically significant value of R-squared (99%), and statistically significant values of F-statistic probability between (0.000 - 0.00027) at level 5%. Additionally, the residuals were Normally distributed; P-values of Jarque-Bera Normality Test between (0.510203 - 0.87321) are not statistically significant values at 5% level, and also the residuals were not autocorrelation (not serial correlation); most of R-squared P-values of Breusch-Godfrey Serial Correlation; LM Test between (0.1463 - 0.8706) are not statistically significant values at 5% level, also the residuals were homoscedastic; R-squared P-values of Breusch-Pagan-Godfrey Heteroscedasticity Test between (0.2755 - 0.9278); are not statistically significant values at 5% level

. The Theil Ineqauality coefficients values between (0.000505 - 0.034 086) are close to zero and zero Bias Proportions for all airlines in sample research. These results mean that the estimated regression models make sense; with strong power for prediction and forecast.

- The researcher used the Classical Normal Linear Regression Model (CNLRM) to estimate the total costs of five airlines, namely Sudan Airways, Marsland Aviation, Sun Air, Nova Air and Badr Airlines, for ten years duration from 2004 to 2013. The final results of the highest mean total costs per annum were 5, 667, 100 US dollars for Badr Airlines, followed by Sudan Airways; which registered (4,861,494 US dollars) per annum and Marsland Aviation which registered (3,141,858 US dollars) per annum. In the other hand, the lowest mean total costs were of Nova Air (352,633.7 US dollars) per annum and Sun Air (93, 500.7 US dollars) per annum.

- For the estimated model of total cost of Mid Air the researcher used a sample size of eight years (2004 – 2011) duration instead of 10 years ; because this Airline stopped its activities in 2011, when its mean total cost was only (229,609.5 US dollars) per annum.

- By comparing the results of the Panel Regression Models: Pooled OLS Regression Model, Fixed Effects Least Squares Dummy Variable (LSDV), Fixed Effect within–Group (WG) Estimator and Random Effects Model (REM), the researcher found that there was no difference between (REM) and Pooled (OLS) Regression Model. In the other hand, (LSDV) and (WG) Models were identical. By applying Wald Test , the F-statistic equals 13.26867 with probability value equal to 0.000

0, is a high statistical significant value at 5% level. So, we can reject the null hypothesis; that is to say, all the dummy variables are equal to zero, that are represented in Pooled OLS Regression Model, and accept the alternative hypothesis which says that the Fixed Effect Least Squares Dummy Variable (LSDV) is appropriate. Also, by applying Hausman Test the Chi-square statistic value for 4 degrees of freedom equals 53.0747 with probability value equal to 0.0000, is a high statistical significant value at 5% level. So, we can reject the null hypothesis; that is to say, that the Random Effects Model (REM) is appropriate, and accept the alternative hypothesis which says that the Fixed Effects Within Group (WG) is appropriate. Finally, according to Hausman Test and Wald Test, the researcher decided that the Fixed Effects Model was appropriate to estimate the total costs of airlines domestic services.

- The estimated total costs by using the Fixed Effects Model of five Sudanese Airlines were different from (CNLRM) estimator. In (WG) Model the highest mean total cost of Badr Airlines was about (4,592,327 US dollars) per annum, followed by Marsland Aviation whose mean total cost was about (2,634,110 US dollars) per annum, and Nova Air whose mean total cost was about (1,634,243 US dollars) per annum. In the other hand, the lowest mean total cost registered for Sudan Airways was about (1,482,810 US dollars) per annum and for Sun Air it was (882,580 US dollars) per annum.

- Also, the study concludes that, there was no statistical significant of freight (FRT) at level 5% to influence the total cost (TC) in two compa

nies; Sun Air and Nova Air, because there were no activities in this field for eight years in Sun Air and four years in Nova Air. Although, there was a negative significant of freight (FRT) at level 5%, to influence the total cost (TC) in Marsland Aviation which stopped loading freight for about four years during the sample size research. Although there were no continuous freight activities performed by some of these airlines, the researcher observed that there were no statistical significant of freight to influence the total costs in Badr Airlines and Sudan Airways. These results were due to the market competitions during certain seasons in some important routes. Moreover, the undercut fare rates policies practiced by some airlines restricted offering some services during flights, such as free food, drinks, etc… Also, during some seasons the airline used to raise the fare rates to the maximum feasible, so as to compensate for these reductions or, almost, to approach a breakeven. Due to the very high market global competition, it becomes vitally and critically important for the airline to struggle for existence in the sky. So, in the results of Panel Regression Models, there was no statistical significant of freight (FRT) at level 5% to influence the total cost (TC).

- In the subject airlines, there were negative statistical significant of load factor (LF) at level 5%, to influence the total cost (TC). That means when the Load Factor increases, the total cost decreases. Such cases indicate that the airlines policy was not running after high profit gains during that season; but it was just trying to break even, by offering such low fare rates of undercuts for the sake of sky existence..! If such particular airline is not financially capable and strongly managed, it wil

l not be able to commercially exist in aviation field.

- The researcher found a high correlation between the passengers and l oad factor in Sudan Airways cost analysis statistical data. Accordingl y, he did not consider evaluation of passenger's data analysis by the (CNLRM), because the passengers coefficient was of a lower statistica l significant value than, statistical significant value of load factor coef ficient, which statistically should not be considered in this case. This unique case of a high correlation shows that, Sudan Airways total cos ting depends on the load factor without any considerations to the pass engers, which indicates that overloading the aircraft could have often been practiced. As it seems to be illegal and violent, strict warnings a nd penalties should be circulated to all airlines concerned to strictly a bide to safety and air law regulations to immediately cease overloadin g the aircraft.

- The Past situation records in the research showed a progressive increase in passenger/cargo demand, as there was stability in the status of the country and its economy and the number of airlines fleet.

- The Present situation records in the research showed drastic deterioration in aviation activities due to instability of the political situation of the country and the adverse economical development caused by regional wars and unrest in most of Sudan states that led to stoppage and bankruptcy of aviation companies.

- From the forecast result for the period 2014-2018 the researcher concluded that the total cost of the subject sample airlines, except

Bader Aviation, shall gradually decrease during the next five years. This is clearly noted in the airlines activities as reflected in their current actual status in 2016, due to the decreasing number of their fleet that minimizes their activities &, henceforth, their total cost, as reflected by Mid Air & Marsland Aviation that are now closed down and are out of business. The subject forecast above shows an increase in the total cost of Badr Airlines only, as reflected in the current actual status of 2016, due to its remarkable increasing fleet number and operational activities and routes.

5.2. RECOMMENDATIONS:-

During data collection and statistical analysis of this research the r esearcher had clearly noted the negative attitudes and malpractices that have still been the most adverse reasons that caused serious, sorrowful drawbacks in Sudan Aviation field activities and services. Unfortunately, such negative matters shall not lead to any future development in aviation services, if not f ully attended to, with utmost care, awareness and knowhow. For such alarmi ng situation, *the researcher recaps the following mandatory recommendati ons so as to see Sudan Flag hovering in the sky all over the globe:*

On the account of its vast area of strategic, geographical location, variable climate, diversified, agricultural, livestock, industrial resources of oi l, gold and other precious mineral resources, Sudan is considered to be a suit able environment for aviation industry. As it is divided into many states, Sud an need to be linked with both regional and international markets to obtain th e maximum benefit of these resources.

Air transport as a service is not a goal in itself; it is rather a service

complementary to other services available to those travelling for business, shopping, tourism, holidays, education, medical treatments, etc...It has proved instrumental in keeping Sudan as a unified entity, and it can play a vital role in cases of emergencies and disasters; *so,*

- *It is highly recommended* to grant the national airlines companies privileges and facilities as encouraging incentives for, like comprehensive exemptions from input custom fees charges on aviation equipment, tools, machines, spare parts, transport facilities, operations and air ground handling equipment, vehicles, computers and communication appliances as required. Also tax free exemption for the first ten years from the date of operation.

- *It is also recommended* that SCAA mobilize its various administration potentials to provide the necessary infrastructure to airports, airstrips, communication facilities and air navigation aids, so as to link Sudan with the regional and international network airports.

- *It is also recommended* to urge SCAA to continue constructing new modern airports, and to rehabilitate activities for the existing airports and airstrips.

- *It is highly recommended* that the government should economically encourage the airlines by determining a fixed rate of foreign currency to purchase their operational needs, so as to stop violations or selfish greediness that leads them to corruption in black market fare rates charges, etc...Such an acceptable practice shall definitely have positive repercussions on the increase of air traffic, and, consequently refresh the national economy that attracts more investors.

- To guarantee trusted actual aircraft inputs and outputs statistical data on daily basis, it is **strongly recommended** that SCAA Traffic Directorate Data Bank shall always keep comprehensive current details of each aircraft. This shall help economists and researchers in aviation field to do further studies for evaluation and development in this unique field.

- Beside the financial capability, each airline investor should be aviation minded, educated in aviation business; NOT JUST, a greedy, money maker, or an ignorant money slave, running after wealth, and ignoring safety. So, **Education specialty is strongly recommended** and mandatory, for the sake of aircraft serviceability, airworthiness and safety.

- *Fleet type standardization is highly recommended* for Sudan to minimize the risk of facing aircraft grounding and technical delays and cancellations due to shortage of aircraft components or spares. Fleet standardization shall make it easier and faster for each airline to get whatever is needed from any sister company, on loan basis, or by direct purchase through a pool agreement with the neighbouring countries or national companies. It also facilitates availability of qualified trained technical staff and crew, etc.

- As long as standardization can be applicable to the operating airlines in Sudan, it becomes utterly urging to **highly recommend** establishing a unified aviation body in Sudan to standardize the salary structure, benefits, fleet type, staff, crew and premises insurance.

For statistical analysis purpose the researcher highly recommends:

- To use the Classical Normal Linear Regression Model (CNLRM) for predictive purpose and forecasting the function of the total cost of each airline.

- To use the Fixed Effect Models as he concluded that they are more appropriately preferable to statistically estimate the total cost of airline domestic services than the two remaining models of the Panel Regression Models.

5.3. <u>FUTURE STUDIES:-</u>

- Determinants of Domestic Air Travel Demands in Sudan by using the Regression Analysis.
- Statistical Estimation of An Airline cost Function of International Services in Sudan.
- Determinants of the Demand for International Air Travel to and from Sudan by using the Regression Analysis.

REFERENCES & SOURCES

1. REFERENCES.

2. E- SOURCES.

REFERENCES & SOURCES

1. REFRENCES:-

1) Aderamo, A. J. (2010). Demand for Air Transport in Nigeria. *Journal Economics* 1(1):23-31.

2) Adrian, C. (1994). *A dictionary of Econometrics*. Edward Elgor, Cheltenham, UK., pp.178-179.

3) AFRAA, (1999). *African Airlines Association*. Accra, Ghana.

4) Agung, I.G. (2009). *Time Series Data Analysis Using Eviews*. University of North Carolina at Chapel Hill, John Wiley& Sons Pte Ltd.

5) Baikgoki, O.A. and Daw, O.D. (2013). The Determinations of Domestic Air Passenger Demand in the Republic of South Africa. *Mediterranean Journal of Social Science,* 4(13): 389-396.

6) Balcilar, M. (2007). *Forecast Accuracy*. Department of Econometrics, EMU.

7) Baltagi, B.H. (1995). *Econometric Analysis of Panel Data*. John Wiley and Sons, New York, pp.3-6.

8) Ba-Fail, A.O., Abed, S.Y. and Jasimuddin, S.M. (2000). The Determinants of Domestic Air Travel Demand in the Kingdom of Saudi Arabia. *Journal of Air Transportation World Wide*, 5(2), 72-86.

9) Bhadra, D., and Kee, J. (2008). Structure and Dynamics of the Core US Air Travel Markets. A Basic Empirical Analysis of Domestic Passenger Demand. *Journal of Air Transport Management*, 14(1), 27-39.

10) Bhadra, D., and Welles, M., (2005). Air travel by state: Its Determinations and Contributions in the United State. *Public Works Management and Policy*, 10, 119-137.

11) Blanchard, O.J., (1967). Comment. *Journal of Business and Econometric Statistics,* vol. 5, pp. 499-451.

12) Breusch, T.S., (1978). Testing for Autocorrelation in Dynamic Linear Models. *Australian Econometric Papers, vol 17*, pp. 334-355.

13) Breusch, T.S., and Pagan, A. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica, vol. 47*, pp. 1287-1294.

14)C.A.A., (2012). *Sudan Civil Aviation Planning Directorate*, Khartoum, Sudan.

15) Cameron, A.C. and Trivedi, P.K. (2005). *Microeconometrics Methods and Applications*. Cambridge University Press. PP. 679-740.

16) Castelli, L., Preseti, R. and Ukovich, W. (2002). *An Airline-Based Multilevel Analysis of Airfare Elasticity for Passenger Demand*, Deel University of Trieste.

17) Cipan, J. (2004). *Evaluating Forecast Accuracy*, Department of Econometrics, University of Vienna.

18) Cramer H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton, Chap. 17.

19) Darag, J. and Hanly, M. (2001). The Determinants of the Demand for International Air Travel to and from the UK. *Paper presented at the 9th World Conference on Transport Research,* Edinburgh, Scotland, May, 2001.

20)Davidson, J. (2000). *Econometric Theory*, Blackwell publishers, UK, p. 10.

21)Demirsoy, C. and Reeven, P.A. (2012). *Analysis of Stimulation Air Transport Demand in Turkey. What are the main Drivers*, M.Sc. Thesis, School of Economics, Erasmus University Rotterdam.

22)D.G.,C.A.A. (2002). *Khartoum New International Airport (KNIA) Feasibility Study*, Civil Aviation Khartoum, Sudan, 6-12.

23)Dobruszkes, F., Lennert, M., and Hamme, G. V. (2011). An Analysis of the Determinants of Air Traffic volume for European Metropolitan. **Journal of Transport Geography, 19**, 755-762.

24)Draper, N. and Smith, H. (1998). *Applied Regression Analysis,* 3d ed., john Wiley & Sons, New York, pp. 207-208.

25)Durbin, S. and Watson, G.S. (1951). Testing for Serial Correlation in Least- Squares Regression, **Biometrika, vol. 38**, pp. 159-171.

26)Farrar, D.E. and Glauber, R.R. (1967)." Multicollinearity in Regression Analysis: The problem Revisited", **Review of Econometrics and Statistics, vol. 49**, pp. 92-107.

27) Frawley, G. (2000). The International Directory of Civil Aircraft, *Air life Publishing Ltd,* England 207-225.

28) Fridström, L. and Thune-Larsen, H. (1989). An Econometric Air Travel Demand Model for the Entire Conventional Domestic Network: The case of Norway. *Transportation Research,* 23(3), 213-223.

29) Frisch, R. (1934). Statistical Confluence Analysis by Means of Complete Regression Systems, Institute of economics, *Oslo University, publ. 5.*

30) Geary, R.C. (1970). Relative Efficiency of Count Sign Changes for Assessing Residual Autoregression in Least Squares Regression. *Biometrika, vol. 57,* pp. 123-127.

31) Glejser, H. (1969). A New Test of Heterocedasticity. *Journal of American Statistical Association vol. 64,* pp. 316-323.

32) G.M. Marsland Airlines Statistics Section (2012), Khartoum, Sudan.

33) Graham, A. (2000). Demand for Leisure Air Travel and Limits to Growth. *Journal of Air Transport Management,* 6, 109-118.

34)Granger, C.W.J. and Newbold, P. (1974). Spurious Regressions in Econometrics, *Journal of Econometrics, vol. 2*, pp. 111-120.

35)Godfrey, L.G. (1978). Testing for Multiplicative Heterocedasticity. *Journal of Econometrics, vol. 8*, pp. 227-236.

36)Godfrey, L.G. (1978). Testing Against General Autoregressive and Moving Average Error Models When the Regressor Includes Lagged Dependent Variables, *Econometrica, vol. 46*.

37)Goldfeld, S.M. and Quandt, R.E. (1972). *Nonlinear Methods of Econometrics*, North-Holland, Amsterdam, chapter 3.

38)Greene, W.H. (1951). *Econometric Analysis*, Fifth Edition. New York University, Upper Saddle River, New Jersey.

39)Grosche, T., Rothlauf, F. and Heinzl, A. (2007). Gravity Models for Airline Passenger Volume Estimation. *Journal of Air Transport Managerment,* 13,175-183.

40)Gujarati, D.N. and Porter, D.C. (2009). *Basic Econometrics*, Fifth Edition, Uni McGraw-Hill Companies.ted State

Military Academy, West Point and University of Southern California.

41) Harrison, M.J. and McCabe, B.P. (1979). A Test of Heterocedasticity Based on Ordinary Least Squares Residuals, *Journal of American Statistical Association, vol. 74*, pp. 494-499.

42) Hausman, J.A. (1978). Specification Tests in Econometrics. *Econometrica,6, vol 46,* pp. 1251-1271.

43) Heij et al. (2004). *Econometric Methods with Applications in Business and Economics*. Oxford University press, Oxford, p. 197.

44) Hill, R.C. , Judge, G.G. and Griffiths, W.E. (1982). *Introduction to the Theory and Practice of Econometrics.* John Wiley & Sons, New York, pp.300-621.

45) Hyndman, R.J. (2014). *Forecasting: Principles and Practice*. University of Western Australia, pp. 174-222.

46) *ICAO FAQ, (2011). Freedoms of the Air International Civil Aviation Organisation.* United Nations.

47) Inglada, V., Rey, B., Alvarez, A. R. and Millans, P. C. (2005). Liberalisation and Efficiency in International Air

Transport. Unpublished manuscript, Department of Economics, Carlos III University, Madrid, Spain.

48) Jackson, R. (2004). *Civil Aircraft.* Amber Books Ltd, London 119-313.

49) Jarque, C.M. and Bera, A.K. (1987). A Test for Normality Observations and Regression Residuals. ***International Statistical Review, vol. 55,*** *pp. 163-172.*

50) Johnston, J. and Dinardo, J. (1997). *Econometric Methods*, Fourth Edition, University of California, Irvine, McGraw-Hill Companies.

51) Kendall, M.G., Buckland, W.R. (1971). *Dictionary of Statistical Terms*, Hafner Publishing Company, New York, p. 8.

52) Khadaroo, J. and Seetanah, B. (2008). The Role of Transport Infrastructure in International Tourism Development: A Gravity Model Approach. ***Journal of Air Transport Management***, (5), 831-840.

53) Kleinbaum D.G., Kupper L.L. and Muller, K.E. (1988). *Applied Regression Analysis and Other Multivariate Methods*, 2d ed., PWS-Kent, Boston, Mass, p. 210.

54) Kmenta, J. (1986). *Element of Econometrics*, 2ed., Macmillan, New York, pp. 625-633.

55) Kumar, T.K. (1975). *Multicollinearity in Regression Analysis*, Review of Econometrics and Statistics, Vol. 57, pp. 366-368.

56) Maddala, G.S. (1992). *Introduction to Econometrics*, Second Edition. University of Florida, Macmillan Publishing Company.

57) Malinvaud, E. (1996). *Statistical Methods of Econometrics.* Rand McNally, Chicago, p.75

58) Marazzo, M., Scherre, R., and Fernandes, E. (2010). Air Transport Demand and Economic Growth in Brazil: Atime series analysis. **Transportation Research Part E**, **46**, 261-269.

59) Michael, D. (1978). *Econometric Models, and Applications.* Prentice Hall, Englewood Cliffs, N.J., chap. 3

60) NASA, Kennedy Space Center (1982). *STS and Cargo Glossary, Acronyms and Abbreviation*, Vandenberg Air Force Base.

61) O'connell, J. F. and Williams, G. (2004). Passengers' Perceptions of Low Cost Airlines and Full Service Carriers. Unpublished manuscript, Cranfield University, UK.

62) O'Hagan, J. and McCabe, B. (1975). Tests for the Severity of Multicollinearity in Regression Analysis A comment, *Review of Econometrics and Statistics, vol. 57*, pp. 368-370

63) Oum,T.H, Park, J. H. and Zhang, A. (1996). The Effect of Airline Codes haring Agreements on Firm Conduct and International Air Fares. *Journal of Transport Econometric and policy*, 4(2), 187-202.

64) Oum,T.H, Yan, J. and Yu.C. (2008). Ownership Forms Matter for Airport Efficiency: A Stochastic Frontier Investigation of Worldwide Airports. *Journal of Urban Economics* 64, 422-435.

65) Park, R.E. (1966). Estimation with Heterocedastic Errors Terms, *Econometrica, vol. 34*, p.888

66) Petronas Marketing (2012). *Impact of Liberalization Policy on low Cost Carrier & Passengers Flow* (fuel Overview), Sudan.

67) Rajiv, D.B. and Johnston H. (1993). An Empirical study of Cost Drivers in the U.S. Airline Industry, Unpublished manuscript, Department of Economics, University of Miiesota.

68) Rawlings, J.O., Pantula, S.G. and Dickey D.A. (1932). *Applied Regression Analysis: A Research Tool*, Second Edition. North Carolina State University, Springer, USA.

69) Rendall, D. (1999). *Aircraft Recognition Guide,* Harper Collins Publisher, London 10-489.

70) Ross, S.M. (2000). I*ntroduction to Probability and Statistics for Engineers and Scientists,* Second Edition. Harcourt Academic Press, New York, pp. 193-194.

71) S.A.C. (2012). *Sudan Airways Central Planning Directorate*, Khartoum, Sudan.

72) S.A.C. (2012). *Sudan Airways Research Study Directorate*, Unpublished Manuscript, Khartoum, Sudan.

73) S.C.A.A. (2012). *Sudan Civil Aviation Integrated Statistics Centre,* Khartoum, Sudan.

74) Seraj, Y. A., Abdullah, O. B. and Sajjad, M. J., (2001). An Econometric Analysis of International Air Travel Demand

in Saudi Arabia, *Journal of Air Transport Management*, 7(3), 143-148.

75) Sc hafer, A. (1998). The Global Demand for Motorized Mobility. *Transportation Research*, **32**(6), 455-477.

76) Sivrikaya, O. and Tunc, E. (2013). Demand Forecasting for Domestic Air Transportation in Turkey. *The Open Transportation Journal*, 7, 20-26.

77) Sultan, A.M., Siddig, M.K. (2006). *Air Transport Company Feasibility Study*, Civil Aviation, Khartoum, Sudan, 5-12.

78) Sultan, A.M. (2012). Aviation Prospect in Sudan, Unpublished, Khartoum, Sudan, Issue 3, 1-15.

79) Sultan, M.A., Mudawi, M.H. and Abdellatif, A.H. (2016). Past, Present & Future of Airlines Domestic Services in Sudan. *IOSR Journal of Mathematics (IOSR-JM).12*(3), Ver. IV, 74-85.

80) Sultan, M.A., Mudawi, M.H. and Abdellatif, A.H. (2016). Estimation of Sudanese Airlines Domestic Services Cost Function. *IOSR Journal of Mathematics (IOSR-JM). .12*(3), Ver. VI, 42-50.

81)Theil, H. (1971). *Principles of Econometrics*, John Wiley & Sons, New York, pp. 43-45.

82)Tintner, G. (1965). *Econometrics,* John Wiley & Sons, New York.

83)White, H. (1980). A Heterocedasticity Consistent Covariance Matrix Estimator and A Direct Test of Heterocedasticity, *Econometrica, vol. 48*, pp. 817-818.

84)Yule, G.U. (1926).Sampling and Nature of Time Series, *Journal of the Royal Statistical Society, vol. 89*, pp. 1-64.

85)Yule, G.U. and Kendall, M.G. (1953). *An Introduction to the Theory of Statistics.*Charles Griffin & Company, London, p.455

2. E-SOURCES:-

1) Aircraft Operating Cost Calculator:

(http://www.freightmetrics.com).

2) Airlines Information: (http://www.airlines-inform.com).

3) Badr Airlines: (http://www.badrairlines.com).

4) Bank of Sudan: (http://www.cbos.gov.sd).

5) Marsland Aviation: (http://www.marsland.com).

6) Mid Airlines: (http://www.midairlines.com).

7) Nile Bakri Aviation Co. Ltd.: (http://nilebakri.com).

8) Nova Airways: (http://www.novaairways.com).

9) Sudan Air: (http://www.sudanair.com).

10)   Sun Air: (http://www.sunair.com).