KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY- KUMASI

[College of Science/Institute of Distance Learning]

THESIS TOPIC:

FORECASTING INDUSTRIAL/COMMERCIAL ELECTRICITY CONSUMPTION USING THE LOGISTIC MODEL

(Case study of the Electricity Company of Ghana Limited)

By

[Emile Kpakpo Adotey, (B.Sc. Computer Science, EC-Council CEH)]

A Thesis submitted to the Department of Mathematics, Kwame Nkrumah University of Science and Technology in partial fulfillment of the requirements for the degree of [MSc. INDUSTRIAL MATHEMATICS]

[June, 2013]

DEDICATION

My first thanks goes to the Almighty God, then to my parents and relatives. Not forgetting my classmates in MSc. Industrial Mathematics, Institute of Distance Learning Accra Centre 2011 - 2013 year group.

DECLARATION

I hereby declare that this submission is my own work towards the MSc. Industrial Mathematics degree, and to the best of my knowledge, it contains no material previously published by another person nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in the text.

EMILE KPAKPO ADOTEY (20249934)

--- ------------------------- -------------------

Student Name & ID Signature Date

Certify by:

DR. OSEI FRIMPONG

--- ------------------------- -------------------

Supervisor(s) Name Signature Date

Certified by:

--- ------------------------- -------------------

Head of Dept. Name Signature Date

Certified by

--- ------------------------- -------------------

Dean, IDL Signature Date

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my supervisor, Dr. Osei Frimpong of the Mathematics Department (Kwame Nkrumah University of Science and Technology). Who offered constructive suggestions and checked my thesis when I was working on it. I am also pleased to acknowledge the helpful recommendations and valuable perspectives offered by my friends and classmates in MSc. Industrial Mathematics, Institute of Distance Learning Accra Centre 2011 - 2013 year group.

Emile Kpakpo Adotey

ABSTRACT

This thesis presents an electricity forecasting model based on the Logistic equation. The proposed model is applied to electricity consumption data of only industrial and commercial consumers. This data is obtained from an Automatic Meter Reading (AMR) system implemented by the Electricity Company of Ghana Limited. The system enables the company to read the meters of industrial and commercial customers remotely without having to send a meter reader to the customer's premises. As a result of the implementation of the AMR system, the Electricity Company of Ghana Limited has the capacity to feed enormous amount of energy consumption data into a central AMR database at the head office.

The development of the model involved using the analytical solution of the Logistic differential equation. The solution curve has monthly energy consumption as the dependent variable and time as the independent variable. The parameters in the solution curve are computed using linear regression analysis. This is done by fitting the linear form of the solution to historical electricity consumption data. The carrying capacity of the Logistic equation which is referred to as optimal asymptote in this thesis is computed using Fibonacci Search Technique. The optimal asymptote and the parameters that are computed are further substituted into the solution curve of the Logistic equation. This final solution is then used to compute future consumptions.

Finally the model was applied and a comparison made between forecast consumption and historical consumption to establish the validity of the model.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ECG	Electricity Company of Ghana Limited
BSP	Bulk Supply Point
SLT	Special Load Tariff
CT-Non SLT	Commercial Tariff – Non Special Load Tariff
AMR	Automatic Meter Reading
ARIMA	Autoregressive Integrated Moving Average
GDP	Gross Domestic Product
VAL	Variable Asymptote Logistic
GWh	Giga Watt Hour
KWh	Kilo Watt Hour
SSR	Sum of Squared Residuals
MAPE	Mean Absolute Percentage Error

CHAPTER 1

INTRODUCTION

This introductory chapter deliberates on the usage of electricity consumption data at the Electricity Company of Ghana Limited. The discussion also focuses on the scope of the electricity consumption forecasting model that will be developed. Major challenges that this research seeks to address and the system that is serving as the basis for the model are discussed.

1.0 BACKGROUND

The Electricity Company of Ghana Limited is a 100% government shares limited liability company with the mandate of distributing electricity in the southern half of the country. The company has two main classes of energy consumers; these are domestic consumers and commercial/industrial consumers. The second group is further classified based on the level of energy consumption. These consumption classes are:

- Bulk Supply Point Consumption (BSP)
- Boundary Consumption (Boundary)
- Special Load Tariff Consumption (SLT)
- Commercial Tariff – Non Special Load Tariff Consumption (CT-Non SLT)

These classes of consumers constitute a very critical group to the company. As a result of their level of energy consumption and the important role that they play in the economic development of the country, their relevance cannot be overemphasized .The commercial/industrial energy consumption data constitutes the basis for this research work. In other words the data that is collected from the above four classes of commercial/industrial customers is the data that will be analyzed throughout this research work.

In the past the company relied on personnel who are deployed to customer premises to read meters. After which the energy consumption data is sent to the respective regional offices for customer bill to be generated. This approach constituted a phase in metering technology development which had a lot of challenges. Challenges such as the accuracy and reliability of energy consumption data that is presented by meter reading personnel. Despite the enormous benefits that the company derived from those technologies, there were major bottlenecks. The company used available technologies until eventually smart meters started emerging on the metering market. These meters brought a revolution on to the metering market as a result of the functionalities and capabilities that they offered. The company started switching over to the usage of these smart meters on a pilot basis but was still stuck to the old ways of deploying meter reading personnel to customer premises to take meter readings. This meant that very little benefits were derived from the deployment of these smart meters relative to the benefits that could have been derived from their deployment.

1.0.1 THE AUTOMATIC METER READING SYSTEM

In an attempt to inject a high level of efficiency and the application of modern technology in the collection of energy consumption data, the Electricity Company of Ghana Limited embarked on a project which involves the deployment of smart meters for industrial/commercial consumers. This deployment is limited to the southern half of the country and was accompanied by the implementation of an Automatic Meter Reading (AMR) System. The smart meters are programmable energy meters that can be programmed to function in a specified manner. It has the capability of reading and storing energy consumption data, voltages, currents and power factors. The meter has a communication modem attached to it, which enables communication between the meter and a communication point at the company’s head office. With the aid of the

Automatic Meter Reading System and the communication infrastructure that has been put in place, it is possible to remotely read energy consumption data from the meters. The meter readings are stored in a central database at the company's head office. The Automatic Meter Reading System as the name implies is a computer software and hardware implementation which has given the company the leverage to read enormous amount of meter data.

1.1 PROBLEM STATEMENT

As a result of the implementation of the Automatic Meter Reading System (AMR), and the deployment of smart meters for commercial/industrial consumers across the southern half of the country, the company now has the leverage to remotely read enormous amount of energy consumption data into a central database at the company's head office. These readings are done at a relatively low cost. However, all that the company does with the data at the moment is to extract the readings on a monthly basis for the generation of customer bills. The company does not take advantage of other meaningful pieces of information that can be generated from the huge data available.

In the past, prior to the implementation of the Automatic Meter Reading System, there were challenges with the accuracy, reliability and availability of energy consumption data. However, with the implementation of the Automatic Meter Reading System, there is a twenty four hour availability of reliable and accurate data. However, the major concern as expressed in the previous paragraph is the underutilization of the data. The data is underutilized because there are no well defined, accurate and reliable concepts and techniques that can be used to analyze the data that is available.

1.2 OBJECTIVIES

From previous discussions it has been identified the leverage that the Automatic Meter Reading System offers. This is leverage in terms of the quantity and quality of data that can be made available in a central database of the Electricity Company of Ghana Limited. It is prudent and relevant that a well defined mathematical model be develop for the purposes of analyzing the data. The model should be tailored at analyzing the data in order to generate results that are targeted at meeting the forecasting needs of the company's management team.

Forecast in the sense that it should be possible to analyze energy consumption trend of a customer within a specified time range. Based on that analysis it should be possible to forecast future energy consumption trend of the customer. As a result of the urgent need and importance of a model that can be used to analyze the data, this research work has been embarked on and targeted at developing a mathematical model based on the logistic differential equation. The model will be used to analyze available energy consumption data with the intention of being able to forecast future energy consumption trend. It should be emphasized that the model is targeted at only industrial/commercial consumers of the company.

In specific terms it should be expected that upon the completion of this research work, the following would be achieved:

1. Develop a mathematical model based on the Logistic differential equation.
2. Compute model constants and parameters by analyzing historical energy consumption data using Linear Regression and Fibonacci Search Technique.
3. Developed a computer algorithm that implements the model using the C Sharp (C#) programming language.

4. Forecast future energy consumption trend using the developed model.

1.3 JUSTIFICATION

As stated in the objectives, this research work is targeted at developing a mathematical model for forecasting industrial/commercial energy consumption. In other words domestic consumers will not be considered in this research. The data that will be used is limited to the consumption data that is read remotely from smart meters across the southern half of the country into the company's central database. These readings as stated much earlier is done using the Automatic Meter Reading System. The limitation is as a result of the fact that the Automatic Meter Reading System has been implemented for only industrial/commercial consumers.

As a result of the unavailability of a model that can be used to forecast with high level of accuracy, the future energy demand of consumers, this research cannot be overemphasized. The model will make it relatively easier to plan ahead of time, the nature of distribution networks that must be implemented. The model will equally help determine the capacity of transformers and switch gears that must be put in place in order to cater for future energy demand at optimized cost to the company.

Based on the challenges enumerated under the problem statement, the proposed model can serve as a solution that will help address those challenges. The model will eventually help minimize the commercial and technical losses that the company is currently faced with, which at the moment stands at about 23%.

Major management challenges such as the ability to forecast future energy consumption, and the identification of consumption trend could easily be surmounted supposing the model that is developed is applied in the analysis of the available data. Forecasting energy consumption forms a major factor that informs the nature of distribution networks that must be implemented and the capacity of transformers that must be installed.

1.3.1 REASONS FOR USING THE LOGISTIC MODEL

Over the years, a lot of growth curve models have been developed for forecasting electricity consumption. Nevertheless, it has been relatively difficult to state which model is the most effective. The Logistic model which involves the fitting of the logistic curve to historical energy consumption data, and also involves the use of Fibonacci search technique to establish optimal asymptote, has proved to be very effective method for forecasting electricity consumption. This model has also proved to be very effective method for forecasting many social and technological patterns. It offers some level of superiority as a result of its inherent recognition for physical limits to growth if conditions are unchanged. Many of the models proposed for forecasting electricity consumption uses economic factors. However, the Logistic model uses mathematical equation that describes the energy consumption data as a time series with a relatively higher degree of accuracy.

The most commonly used model for forecasting the electricity consumption of countries and institutions use economic models and data. However, as a result of the unstable and stochastic nature of macro and micro economic policies, it can be stated that forecasts that are made using economic models and data are relatively unreliable. Economic models as stated above have proved to be quite unreliable, despite the fact that it has provided accurate forecast in some cases.

Due to the relatively higher degree of accuracy of the Logistic model, it was chosen for this research.

1.4 METHODOLOGY

The development of the model that will achieve the stated objectives will begin by the creation of the database that will host the data that will be analyzed. As stated in the background of the research, the data that will be used is the data read remotely from smart meters across the southern half of the country into a central database at the company's head office.

Relevant and needed data in the central database will be extracted, transformed and loaded into the database that will be created. This means all the data analysis that will be done will involve the use of data in the new database that will be created. The data that will be used for the analysis will span the period January 2012 to February 2013. The data will however continue to grow on a monthly basis in the central database. The model when encoded into software will handle an estimated two thousand (2000) customers when it goes into production. These customers as was stated earlier are industrial/commercial consumers whose energy consumption data is read on a monthly basis.

Upon the completion of the database development and the loading of data, the next stage of the research will involve the identification of all the assumptions that will form the basis for the model development. The assumptions will take into consideration the model to be developed and the data that will be analyzed.

The next phase will involve the identification of the various variables and parameters that will be used in the model. These variables and parameter will be related to energy consumption, time and factors that influence energy consumption.

The main stage of the research will be the formulation of the mathematical equation that will form the model, and the computation of parameters and constants.

Developing the equations will first involve writing of the logistic equation in its differential equation form. The equation will describe the rate of change of energy consumption with respect to time. The parameters and variables that have been identified will be used in formulating the differential equation. A solution curve to the differential equation will be found followed by the computation of unknown constants and the optimal asymptote value. The optimal asymptote value is the projected ceiling or expected maximum energy consumption of a consumer. The value will be computed using the Fibonacci search algorithm.

The constants in the solution curve to the differential equation will be computed by fitting the linear form of the solution curve to the historical energy consumption data of a consumer. The fitting of the solution curve to the historical data will involve the use of linear regression analysis. It should be stated that the constants and the optimal asymptote value will vary for each consumer whose data is will be analyzed. Future energy consumption of a consumer will finally be extrapolated using the particular solution curve that is obtained.

CHAPTER 2

RELEVANT LITERATURE

This chapter presents previous and related works that involved the use of the logistic model. Further presentation is also done on various electricity forecasting models that have been developed by various researchers. The models are presented qualitatively.

2.0 QUALITATIVE ANALYSIS OF THE LOGISTIC EQUATION

According to Mohamed Z. and Bodger P.S. (2003), the Logistic Equation has been applied in a wide range of areas ranging from biological, technological and economic fields. Studies such as Pearl's biological growth curve, Fisher and Pry's technological substitution model and Mansfield's modeling of the rate of imitation are some of the areas where the logistic equation has been applied. The logistic equation nonetheless has also had extensive usage in electricity consumption forecasting for various countries and economies around the world.

2.0.1 THE EXPONENTIAL GROWTH MODEL

(http://en.wikipedia.org/wiki/Exponential_growth/ on January 19, 2013), the Exponential growth equation is given as

$$\frac{dE}{dt} = kE \qquad (2.01)$$

Where E is the dependent variable, t is the independent variable and k is a constant of proportionality.

From Equation 2.01 the rate of growth of the dependent variable E is proportional to its value thereby forming a first order linear differential equation.

(https://www.math.duke.edu/education/ccp/materials/diffeq/logistic/logi1.html on January 19, 2013), from a qualitative analysis point of view of *Equation (2.01)*, it implies that if

$$E = 0$$

then

$$dE/dt = 0$$

This gives the constant function

$$E(t) = 0$$

This is the equilibrium solution of the differential equation. On the other hand, if

$$E(t) \neq 0$$

at an initial time t_0, then at $t = t_0$

$$\frac{dE}{dt} = kE(t_0) \neq 0$$

If $k > 0$ and $E(t_0) > 0$ then

$$\frac{dE}{dt} = kE(t_0) > 0$$

E increases as its value increases thereby producing the graph shown in **Figure 2.1.** Using different values of $E(t_0)$ as the initial condition will produce different functions $E(t)$ as shown in **Figure 2.2.**

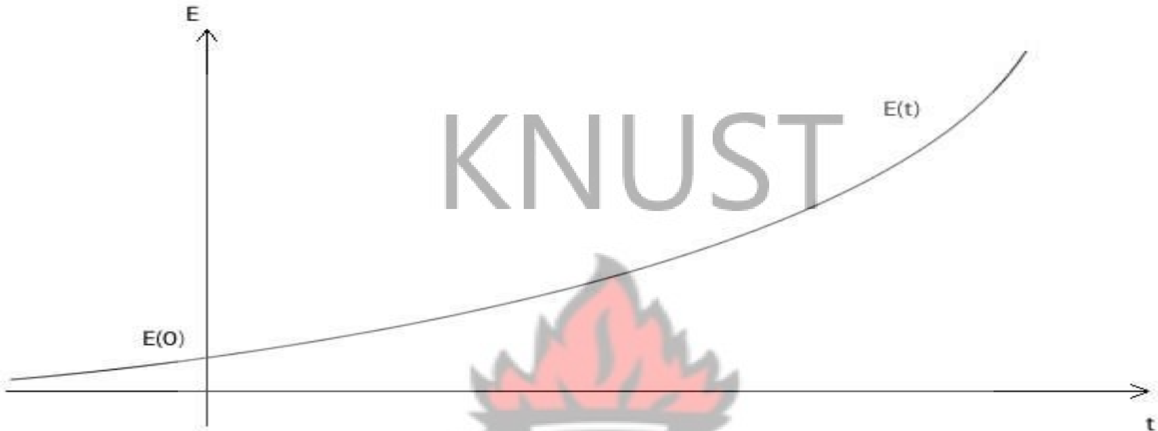

Figure 2.1: Solution Curve for $\frac{dE}{dt} = kE$

Figure 2.2: Solution Curve for $\frac{dE}{dt} = kE$ with Different Initial Conditions

From **Figure 2.2** if $k > 0$ and $E(t_0) < 0$ then $dE/dt < 0$ thereby producing a curve with a negative gradient below the *t – axis* which is a replication of similar curves above the *t – axis* in a flip manner.

2.0.2 THE LOGISTIC MODEL

The Exponential growth model does not take into consideration the fact that there could be a limit to the growth of the curve. However in practical or real world application of the equation, it must be taken into consideration that there is a limit to growth. This leads to the Logistic equation proposed by Pierre-Francois Verhulst (1838) which is written as

$$\frac{dE}{dt} = k\left(1 - \frac{E}{E_m}\right)E \qquad (2.02)$$

Where E is the dependent variable, t is the independent variable and k is a constant of proportionality; E_m is the carrying capacity which is a limit to the growth of the dependent variable. Qualitative analysis of the logistic equation is done for three conditions where dE/dt is zero, positive and negative. For the first condition where dE/dt is zero we obtain a quadratic function which is shown in **Figure 2.3** with the curve crossing the E-axis at $E = 0$ and $E = E_m$.

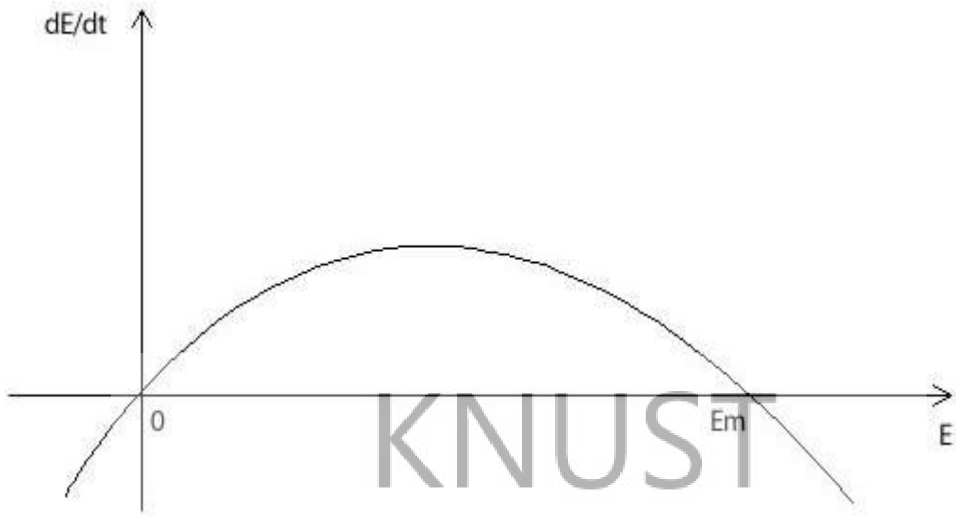

Figure 2.3 Graph of the Logistic Equation on an axis of $\frac{dE}{dt}$ *and* E

From **Figure 2.3** $dE/dt = 0$ at $E(t) = 0$ *and* $E(t) = E_m$ which serves as equilibrium solutions as shown in **Figure 2.4**

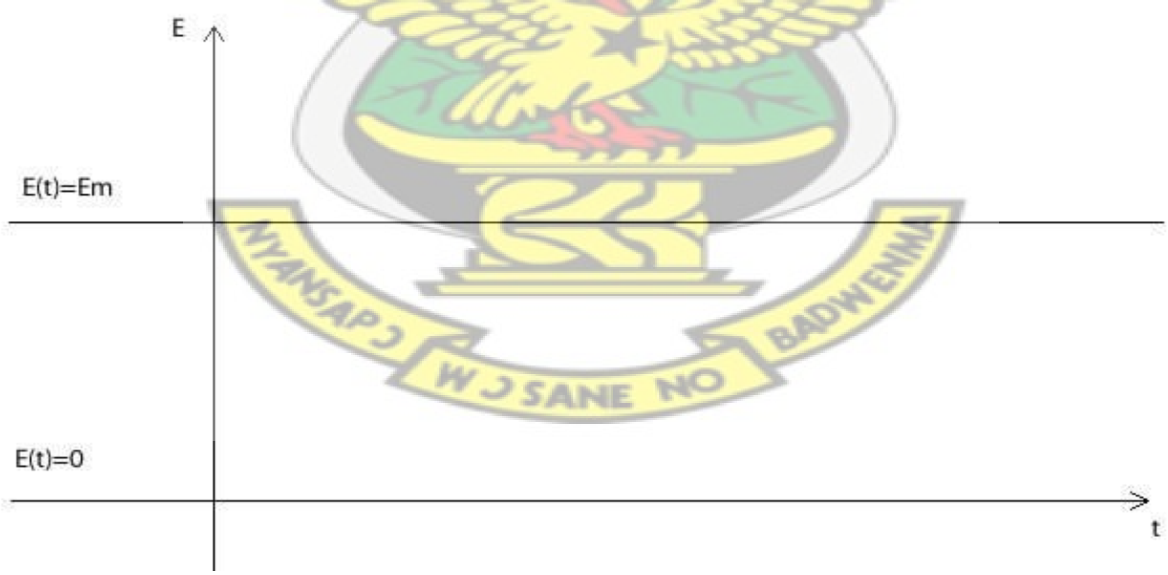

Figure 2.4 Equilibrium Solution of the Logistic Differential Equation

(http://tutorial.math.lamar.edu/Classes/DE/EquilibriumSolutions.aspx on January 19, 2013), if the dependent variable $E(0)$ lies between 0 and E_m, then $dE/dt > 0$ and as $E(t)$ increases, it levels off at $E = E_m$ and as E decreases it levels off at $E = 0$ as shown in **Figure 2.5**. On the other hand if $E(0) > E_m$ then $dE/dt < 0$ and $E(t)$ will be decreasing until it eventually levels off at $E = E_m$ as shown in **Figure 2.5**

Figure 2.5 Solutions of the Logistic Differential Equation approaching the Equilibrium Solutions

Finally, if $E(0) < 0$ then $dE/dt < 0$ and $E(t)$ decreases but does not level off at any particular value. For this instance dE/dt becomes more and more negative as $E(t)$ decreases.

The logistic differential equation produces different solution curves depending on the initial condition given. **Figure 2.6** shows different solution curves for different initial conditions. From the graph shown, $E = 0$ and $E = E_m$ are equilibrium solutions, $E(t)$ increases if $0 < E < E_m$ and $E(t)$ decreases if $E > E_m$ or < 0 . From the qualitative analysis of the logistic differential equation it can be seen that the exact value of $E(t)$ at any given time t depends of the value of $E(0)$, k , and E_m.

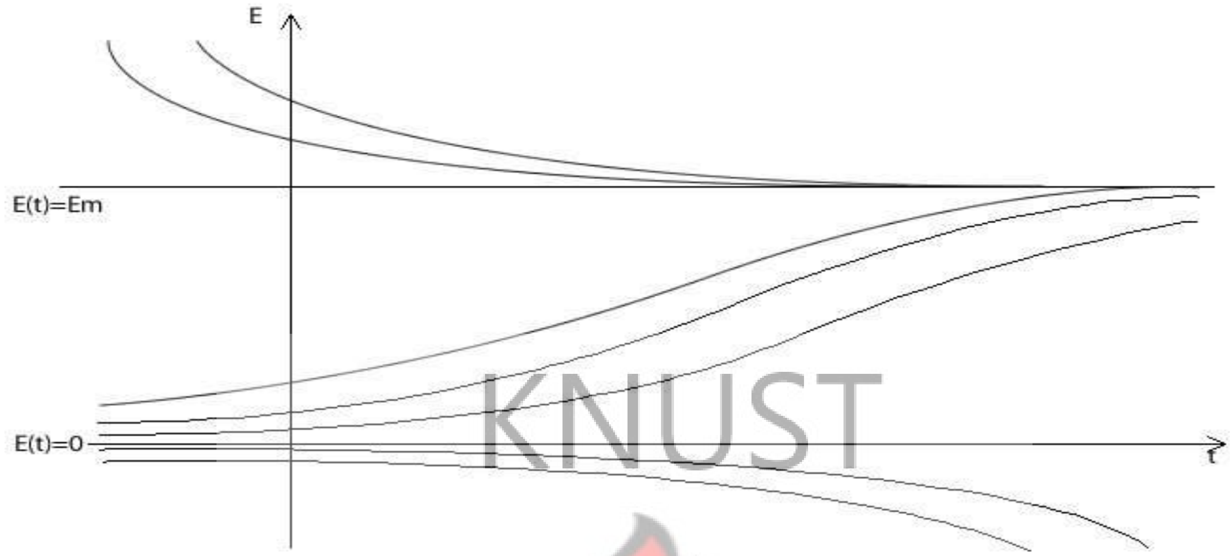

Figure 2.6 Solutions of the Logistic Differential Equation approaching the Equilibrium Solution $E = E_m$ and moving away from the Equilibrium Solution $E = 0$.

2.1 APPLICATIONS OF THE LOGISTIC EQUATION

In the subsequent section, we discuss some applications of the Logistic Model.

2.1.1 Population Growth Modeling

One popular area of application of the logistic equation is the modeling of population growth which was originally done by Pierre-Francois Verhulst (1838), where the rate of population growth is expressed as being proportional to the existing population and the available resources, all other things being equal (http://en.wikipedia.org/wiki/Logistic_function on January 21, 2013). The logistic equation for population growth which is also called Verhulst equation is sometimes called Verhulst-Pearl equation following its rediscovery in 1920. It was also derived by Alfred J. Lotka (1925) and was called the law of population growth.

(https://www.math.duke.edu/education/ccp/materials/diffeq/logistic/logi1.html on January 19, 2013), in modeling the population growth, let *P* represent the population size and *t* represent time and the equation written as a differential equation in the form

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{K}\right) \tag{2.03}$$

Where the constant r defines the growth rate and *K* is the carrying capacity.

From the Verhulst equation, the unimpeded growth rate is modeled by the first term +*rP*. The value of the rate *r* represents the proportional increase of the population *P* in one unit of time. As the population grows, the second term, which multiplied out is $-rP^2/K$, becomes larger than the first as some of the population *P* interfere with each other by competing for some critical resource, such as food or living space. This antagonistic effect is called the *bottleneck*, and is modeled by the value of the parameter *K*. The competition diminishes the combined growth rate, until the value of *P* ceases to grow (this is called *maturity* of the population). It is important to stress that the carrying capacity is asymptotically reached.

2.1.2 Modeling the growth of tumors

(http://en.wikipedia.org/wiki/Logistic_function on January 30, 2013), an application area of the logistic curve is in the field of medicine, where the logistic differential equation is used to model the growth of tumors. Denoting with *X(t)* the size of the tumor at time *t*, its dynamics are governed by the equation

$$X' = r\left(1 - \frac{X}{K}\right)X \tag{2.04}$$

which is of the type

$$X' = F(X)X, \tag{2.05}$$

where X' is the rate of proliferation of the tumor.

If chemotherapy is started with a log-kill effect, the equation may be revised to be

$$X' = r\left(1 - \frac{X}{K}\right)X - c(t)X \tag{2.06}$$

where *c(t)* is the therapy-induced death rate. In idealized case of very long therapy, *c(t)* can be modeled as a periodic function (of period T) or (in case of continuous infusion therapy) as a constant function, and one has

$$\frac{1}{T}\int_0^T C(t)dt > r \rightarrow \lim_{t\rightarrow+\infty} X(t) = 0 \tag{2.07}$$

i.e. if the average therapy-induced death rate is greater than the baseline proliferation rate then there is the eradication of the disease. This is an oversimplified model of both the growth and the therapy and does not take into consideration phenomenon such as clonal resistance.

2.1.3 Fisher and Pry's Technological Substitution Model

Fisher and Pry (1971) proposed the technological substitution model which asserts that, technological evolution consists mainly of substituting a new form of satisfaction for an old one. The model describes the modeling of competing technologies in the sense that it focuses on the time pattern of technological acceptance in a social environment. For two competing technologies in this case, it has been noticed that the acceptance of each technology or innovation over time assumes the S-shaped curve (the logistic-type pattern).

The assumptions under which the logistic model for technological substitution is obtained are:

1. The number N of individuals (firms, products, services, etc.) in the total population is constant and the nature of these individuals does not change.
2. The technology is disseminated when contact by two individuals occurs; the frequency c of the contact in the unit of time between any two individuals is the same for all pair of individuals and constant over time.
3. The probability p that the technology will be transmitted during contact between those who know about the technology and those who do not know about the technology is constant over time and is the same for all relevant pairs.
4. The technology cannot be lost once acquired.

Let $N_a(t)$ be the number of technologically affected individuals at time t, from the second assumption the number of contacts between technologically affected and unaffected individuals in the period from t to $t + \Delta t$ is equal to

$$cN_a(t)\big(N - N_a(t)\big)\Delta t \tag{2.08}$$

Applying the two last assumptions, the number of new technologically affected individuals in the Δt period is equal to

$$\Delta N(t) = cpN_a(t)\big(N - N_a(t)\big)\Delta t \quad (2.09)$$

Let *f(t)* denote the fraction $N_a(t)/N$ technologically affected individuals in the total population. From the above equation we obtain the logistic differential equation

$$\frac{df}{dt} = bf(1-f) \ \ where\ b = cpN \quad (2.10)$$

With the logistic function as its solution

$$f(t) = \frac{1}{1+\exp(-a-bt)} \quad (2.11)$$

where *a* is a constant of integration with value given by the initial fraction *f(0)*.

For the purpose of estimation of parameters *a* and *b,* the above equation is frequently rewritten in the following form:

$$\ln\left(\frac{f(t)}{1-f(t)}\right) = a + bt \quad (2.12)$$

The above presented model is the Fisher-Pry model. The modification of that model was proposed by Blackman (1973). Blackman assumes that there exists an upper limit F of the market share by the new technology. Therefore, the solution of the market share in the model is equal to

$$\ln\left(\frac{f(t)}{1-f(t)}\right) = a + bFt \tag{2.13}$$

For *F=1,* the Blackman model turns to be the Fisher-Pry model.

Floyd, Bright J.R. (1968) and also Linstone and Sahal (1976) proposed another modification of the Fisher-Pry model by adding to the Fisher-Pry solution the second, analogous term but with the limit *F* and without the natural logarithm. The form of that model is the following:

$$\ln\left(\frac{f(t)}{1-f(t)}\right) + \frac{F}{F-f(t)} = a + bt \tag{2.14}$$

The Fisher-Pry and the Blackman models give overestimations of the forecast and the Floyd gives underestimations of the forecast. Sharif and Kabir (1976) proposed a generalized version of the model for technological substitution. They suggest that the linear combination of the Blackman's and Floyd's models can give correct results. This leads to the following generalized model:

$$\ln\left(\frac{f(t)}{1-f(t)}\right) + \frac{\sigma F}{F-f(t)} = a + bt \tag{2.15}$$

The Sharif-Kabir model can be rewritten in the differential equation form

$$\frac{df}{dt} = \frac{bf(F-f)}{F\big(F-f(1-\sigma)\big)} f(1-f) \tag{2.16}$$

For σ=1 we have Floyd's model; for σ=0 and F=1 we have the Fisher-Pry model.

2.1.4 Pearl's Biological Growth Curve

Johns Hopkins biologist Raymond Pearl (1924) proposed a statistical-mathematical view on nature. During the 1920s, when emphasis was put on laboratory experiments on population sequences and cycles and on single-species aggregations or groups, Pearl combined a mathematical point of view with an experimental style. With his studies on *drosophila* (fruit flies), Pearl built a program in comparative demography. Biostatistics became his method of a new biology, not of individuals, but of groups. In a systemic attempt to apply demographic techniques to animal populations, he borrowed from social statistics the established tools of life tables, death rates, and life expectancies. When the predicament of population and food supply came up again in the aftermath of the First World War, Pearl in conjunction with his laboratory studies, started to engage in the problem of human population growth. Together with the mathematician Lowell J. Reed he began analyzing rates of population growth. The degree of mathematization he was seeking was allowed for through the information available in the accumulated body of statistics, on fertility, growth, disease, and mortality.

Pearl's "search for quantitative measures and correlations under different experimental conditions" aimed at establishing a law of population growth, expressed as a mathematical equation which would conform to experimental observations as well as to assumptions about how populations behaved, to indicate future trends reasonably accurately. Pearl and Reed (1920) presented their equation in the nineteen twenties. They called the corresponding graph the "logistic curve", recalling the work of Belgian Pierre-Francois Verhulst (1838), who more than eighty year earlier had described population growth in a similar way and in 1845 had called his curve of growth over time “logistique”. Pearl and Reed adopted the term “logistic” for their smooth, S-shaped curve towards a stable upper limit. The curve eventually contributed to population ecology one of its simplest mathematical models. **Figure 2.7** shows the logistic curve and its first derivative curve.

In Pearl’s words, the curve demonstrated that “plainly all growth, including that of population, is fundamentally a biological matter”. In his times, Pearl’s results led to a controversy over the basic assumptions that went with the law and the curve; among them, the implication of symmetry, with the point of inflection exactly halfway through the curve. To many contemporaries, it seemed an idealization that the forces in the first half were exactly as strong and distributed as in the second half. However, a generalization of the equation to free the curve from its restrictive symmetry would imply that the curve could be made to fit almost any data, which the consequence that it could not be considered a calculating device with prognostic value, and hardly a “law” in the strong sense. Pearl and Reed “had assumed that, which had to be proved, yet they presented their curve as being empirically fitted, not logically derived.” Pearl’s book on *The Biology of Population Growth* demonstrates how the sparse population data of different countries were accumulated and arranged along the curve line, and then extrapolated to fit the logistic world population growth curve. Pearl had come full circle, having first assumed logistic growth in trying to find a curve to fit his initial data, he now believed that the empirical evidence proved the truth of the logistic ‘law’, even though large parts of the curves were extrapolated.”

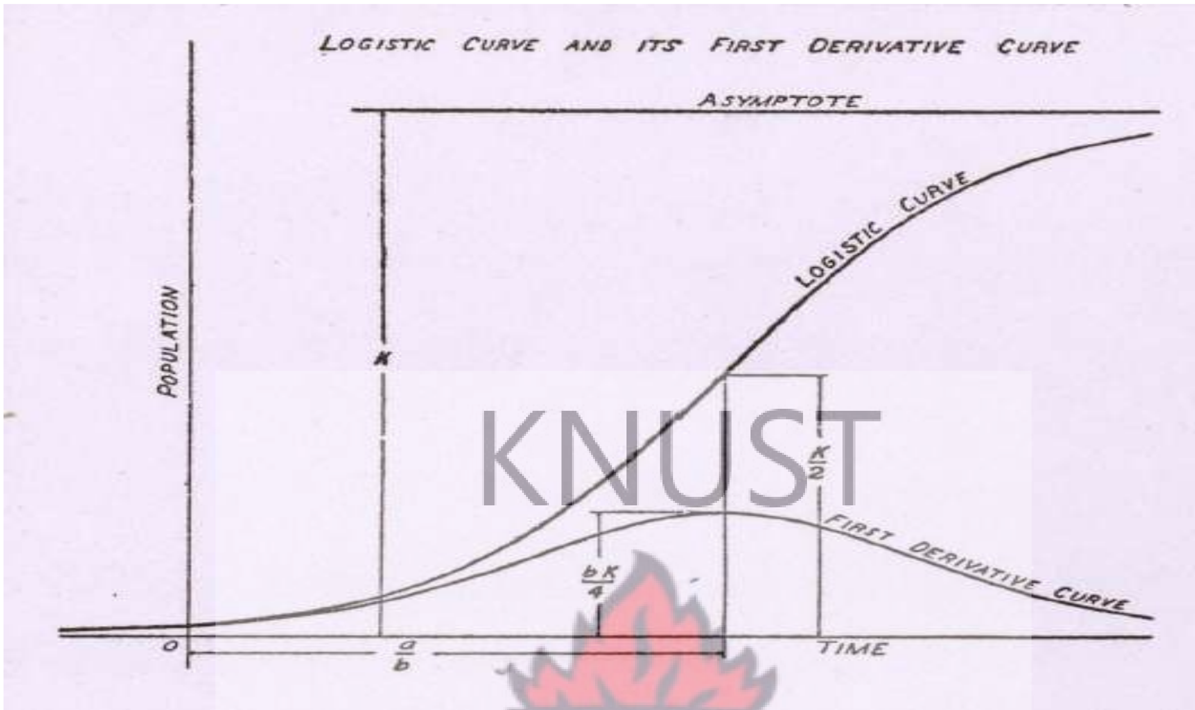

Figure 2.7 The Logistic Curve and its First Derivative

There are five factors that must be taken into account in mathematical modeling of the population growth of biological species as was proposed by Pearl. They are:

1. The finite limit of land area of habitation.
2. The upper limiting asymptote of population constrained by (1).
3. The lower limiting asymptote of the population (= 0).
4. The epochal or cyclic character of growth, successive cycles being additive.
5. The general S-shape of the growth curve.

In order to satisfy the five constraints, Pearl proposed the following general form of the logistic equation

$$y = \frac{k}{1 + Mexp(F(x))} \tag{2.17}$$

where

$$F(x) = a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n \tag{2.18}$$

M is a positive constant, and

k is the upper asymptote.

A third order equation of *F(x)* has been generally used with the assumption that less than five arbitrary constants are used in any practical problem.

2.1.5 Modeling the Rate of Imitation

Mansfield (1961) examined the pattern or rate at which different firms follow an innovation. This helped in building a model for the rate of imitation. Mansfield's approach and assumptions were different from those of Pearl and Fisher and Pry, but the final form of his model is a logistic growth curve. The final form of his model is given by

$$M_t = \frac{n}{1 + \exp[-(h + qt)]} \tag{2.19}$$

where

n is the total number of firms which could possibly introduce an innovator,

M_t is the total number of firms which have done so at time *t*, and *h* and *q* are constants.

2.2 ELECTRICITY CONSUMPTION FORECASTING MODELS

Forecasting electricity consumption has been undertaken using many theoretical methods including growth curves, multiple linear regression methods that use economic, social, geographic and demographic factors, and Box-Jenkins autoregressive integrated moving average (ARIMA) techniques.

Six forecasting models developed for electricity consumption is briefly discussed in this section as was presented by Mohamed Z. etal (2003). First the Logistic model based on the growth curve is discussed. In a second model, the influence of selected economic and demographic variables on electricity consumption is investigated. The study uses population, price of electricity and gross domestic product (GDP). The resulting combined models are discussed which involves using multiple linear regression analysis. The third model used an ARIMA technique in developing electricity forecasting models. Fourthly, two other models, Harvey models and Harvey Logistic models, based on growth curves are discussed.

Finally, the Variable Asymptote Logistic (VAL) model for electricity consumption is discussed. The saturation levels of the logistic curve are estimated using the Fibonacci search technique.

In all the model equations to follow, *Y* represents the annual electricity consumption in Giga Watts Hour (GWH).

2.2.1 LOGISTIC MODEL

The proposed Logistic model is,

$$Y = \frac{F}{1 + \exp(C_0 + C_1 t)} \tag{2.20}$$

Where F is the asymptotic value in GWH, t is time in years and C_0 and C_1 are constants.

2.2.2 HARVEY LOGISTIC AND HARVEY MODELS

The Harvey Logistic model is based on the Logistic model. The proposed Harvey Logistic Model is

$$ln\, y_t = 2\, ln\ Y_{t-1} + \delta + \gamma + \varepsilon_t \tag{2.21}$$

where $y_t = Y_t - Y_{t-1}$, ε_t is a disturbance term and δ and γ are constants to be found by regression and Y_t is the energy consumption at t time

The Harvey model is based on general modified exponentials. The proposed Harvey model is

$$ln\ y_t = \rho\, ln\ Y_{t-1} + \delta + \gamma + \varepsilon_t \tag{2.22}$$

Where $\rho = \frac{k-1}{k}$, $\delta = \ln\left(k\beta\alpha^{1/k}\gamma\right)$, and ρ , β and γ are parameters to be estimated. The value of k determines the form of the exponential function. When k = -1, it is Logistic and when k = 1 it is a simple modified exponential.

2.2.3 THE COMBINED MODEL

This model uses multiple linear regression analysis using economic and demographic variables which is the reason for its name and is given as:

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + u \qquad (2.23)$$

Where X_1 is GDP, X_2 is electricity price, X_3 is population and u is the error term. This model basically uses economic factors to forecast electricity consumption.

2.2.4 AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) MODEL

Autoregressive Integrated Moving Average (ARIMA) models are generally written as ARIMA (p, d, q), where p represents the order of the autoregressive (AR) part, d denotes the degree of first differencing (I) involved and q denotes the order of the moving average (MA) part. The autoregressive (AR) part of the model with order p is of the form:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ \dots \dots + \phi_p Y_{t-p} + e_t \qquad (2.24)$$

The moving average (MA) part of the model consists of the past errors as the explanatory variable. A moving average model of order q is of the form

$$Y_t = c + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ \dots \dots + \theta_p e_{t-q} + e_t \qquad (2.25)$$

Where ϕ and θ are the maximum likelihood estimates of the respective models and e_t is the error series. The Box – Jenkins methodology for modeling time series consists of identification, estimation, testing and forecasting.

2.2.5 VAL MODEL

In the Variable Asymptote Logistic (VAL) model, the saturation level *F* of the Logistic model is estimated using economic and demographic variables $(X_1 \ldots X_n)$ and used as a variable asymptote *F(X)*. The proposed VAL model takes the form

$$Y = \frac{F(X)}{1 + \exp(a_0 + a_1 t)} \tag{2.26}$$

$$F(X) = C_0 + \sum_{i=1}^{n} (C_i X_i) \tag{2.27}$$

Where *F(X)* is the saturation level expressed as a function of *n* variables and C_0 and C_i are the parameters obtained from the explaining variables.

CHAPTER 3

METHODOLOGY

In this chapter, the mathematical model is developed. The steps involved in the development of the model are presented. The application of the model that is developed in this chapter is demonstrated in the next chapter.

3.0 MODEL DEVELOPMENT

The model is based on three mathematical concepts. The main concept involves the use of the Logistic growth model. The second concept involves the use of the Fibonacci search technique for the computation of the optimal asymptote value (carrying capacity) of the logistic equation. This is done using the historical energy consumption data obtained. The third concept involves linear regression and is used to compute the constants in the logistic equation. These constants are computed by treating the energy consumption data as a time series. The final solution that is obtained after the computation of carrying capacity and constants is used to extrapolate future energy consumption.

3.0.1 MODEL ASSUMPTIONS

The assumptions which form the basis of the model are:

1. No new factors influencing the rate of growth different from those during historical period shall come into play in the forecast region. One main factor being the tariff class of consumers

2. The forecast portion of the curve cannot be used to assess the adequacy of the curve to describe the historical growth.

3. The value of the optimal asymptote (carrying capacity) is revised whenever additional data is available.

4. The historical data used is assumed to have some form of logistic growth.

3.0.2 MODEL VARIABLES AND PARAMETERS

Independent Variable

The only independent variable is time which is represented by "t" for all the various equations that will be used for the modeling.

Dependent Variable

The dependent variable in the model is the monthly energy consumption which is represented as "E" in the model.

Parameters and Constants

The parameters and constants used in the model are:

1. "E_{max}" representing the expected maximum monthly energy consumption of a consumer whose consumption is being forecasted. In the logistic equation, this parameter is used as the carrying capacity and referred to as optimal asymptote in this research. This is computed by applying the Fibonacci Search Technique to the historical energy consumption data of a consumer.
2. k_1 and k_2 are constants that are computed by regression analysis. In computing these constants the Logistic growth curve is transformed into a linear equation and fitted to the historical energy consumption data.

3.0.3 MATHEMATICAL FOMULATION

The fundamental equation in the formulation of the model is the Logistic growth model which is represented in its differential equation form as

$$\frac{dE}{dt} = k_1 E\left(1 - \frac{E}{E_{max}}\right) \tag{3.01}$$

The solution to the differential equation above is found using the analytical approach as demonstrated below. Re-arranging *Equation (3.01)* gives the equation

$$\frac{E_{max} dE}{E(E_{max} - E)} = k_1 dt \tag{3.02}$$

Expressing the coefficient of *dE* in *Equation (3.02)* as partial fraction gives

$$\frac{E_{max}}{E(E_{max} - E)} = \frac{A}{E} + \frac{B}{(E_{max} - E)} \tag{3.03}$$

where A and B are Constants. From *Equation (3.03)* we obtain *A=1* and *B=1*

Equation (3.03) reduces to

$$\frac{E_{max}}{E(E_{max} - E)} = \frac{1}{E} + \frac{1}{(E_{max} - E)} \tag{3.04}$$

The differential *Equation (3.02)* becomes

$$\frac{E_{max} dE}{E(E_{max} - E)} = \left(\frac{1}{E} + \frac{1}{(E_{max} - E)}\right) dE = k_1 dt \tag{3.05}$$

Consequently,

$$\int \frac{dE}{E} + \int \frac{dE}{(E_{max} - E)} = \int k_1 dt$$

$$ln|E| - ln|E_{max} - E| = k_1 t + k_2$$

$$ln\left|\frac{E}{(E_{max} - E)}\right| = k_1 t + k_2 \qquad (3.06)$$

Further simplification gives

$$E = \frac{E_{max}}{1 + e^{-(k_1 t + k_2)}} \qquad (3.07)$$

Equation 3.07 is used to forecast energy consumption for a given time *t*. This equation is used after the computation of the parameters and constants E_{max} , k_1 *and* k_2 . Once these parameters and constants are computed the forecast for time $t_1,\ t_2,\ t_3,\ t_4\ \ldots\ t_n$ can be computed.

Pseudocode For The Logistic Equation

The algorithms that are used to perform various computations and tasks have their detailed implementation with the C Sharp (C#) programming language in the appendix. However, the algorithms in their pseudocode form are provided in order to give simplified illustration wherever the algorithm is defined the first time.

The detailed implementation of the Logistic Equation as given in *Equation 3.07* is presented in **Appendix A1.** The pseudocode for the Logistic equation is given below:

procedure logisticEquation (Em , k1 , k2 , t)

{ Comment - Em : Optimal Asymptote also known as the carrying capacity}

{ Comment - k1 : Constant one , k2 : Constant two , t : Time in months which is given as natural number 1, 2, 3, . . . n}

{ Comment - E : Computed result of the Logistic Equation}

begin

Compute E using Equation 3.07

return E

end

3.0.4 EQUILLIBRIUM SOLUTION OF MODEL

Using the Logistic differential equation given in Equation 3.01, a qualitative analysis is done in order to establish the expected consumption growth behavior. Rewriting Equation 3.01 for a consumer, we have:

$$\frac{dE_c}{dt} = k_c E_c \left(1 - \frac{E_c}{E_{max}^c}\right)$$

The equilibrium solution is obtained by solving the equation

$$k_c E_c \left(1 - \frac{E_c}{E_{max}^c}\right) = 0$$

Resulting in the following two equations

$$E_c = 0 \quad and \quad \left(1 - \frac{E_c}{E_{max}^c}\right) = 0$$

Consequently, we obtain two solutions which is given as

$$E_c = 0 \quad \text{and} \quad E_c = E^c_{max}$$

Ideally, a consumer whose consumption is being forecasted is expected to range between zero consumption and the maximum projected consumption which is given as E^c_{max}.

3.1 COMPUTATION OF MODEL PARAMETERS AND CONSTANTS

The constants are first computed using linear regression analysis followed by the computation of the optimal asymptote (E_{max}) using the Fibonacci Search Technique. The constants need to be computed first because the computation of the optimal asymptote will require the use of the constants. The computation is done in four steps, which involve a series of processes starting from the retrieval of required data from the database, followed by the computation of values that will serve as input for the computation of the constants and parameter. The four steps are enumerated below:

3.1.1 STEP 1: RETRIEVAL OF DATA FOR ANALYSIS

The set of historical energy consumption data which will form the basis for the computation is obtained and is represented as

$$H = \{E_t \,|E_t \in R, t = 1,2,3, \dots ,n\} \tag{3.08}$$

where E_t is the historical monthly energy consumption and t is the time in months.

In order to be able to identify the very important and much needed trend that exists in the historical consumption, smoothening is done on the data. The weighted moving average algorithm is used in smoothening out the historical consumption data before the remaining computation is done. The weighted moving average is given as

$$E_t^{wma} = \frac{tE_t + (t-1)E_{t-1} + \cdots + 2E_2 + E_1}{t + (t-1) + \cdots + 2 + 1} \tag{3.09}$$

$t = 1, 2, 3, \ldots, n$ and E_t^{wma} is the weighted moving average at t

<u>Pseudocode For Data Retrieval</u>

The pseudocode for the algorithm that extracts data from the database is given below with the detailed implementation in **Appendix A2.** The pseudocode for the Weighted Moving Average is incorporated in the pseudocode for data retrieval as defined below.

procedure dataExtraction (Mtr , d1 , d2)

{ Comment - Mtr : Meter number of a consumer whose data is to be extracted from the central database }

{ Comment - d1 : Starting date , d2 : Ending date – available data between the starting and ending date is extracted from the database }

{ Comment - C : Array of Energy Consumption data extracted from the database}

begin

{ Comment - The computation in **Step 1** is done using the procedure **dateRange (d1, d2)** }

Step 1 : Compute the range of months between the starting and ending date and store it in a one dimensional array

Step 2 : Extract energy reading from the database for each month within the range of months obtained in **Step 1**.

Step 3 : Compute monthly energy consumption from the data obtained from **Step 2**.

Step 4 : Smoothing of **Step 3** data using weighted moving average algorithm given as Equation 3.09.

return C

end

Pseudocode For Date Range Generation

The detailed implementation of the pseudocode below, which is used to obtain the date range is presented in **Appendix A3**. This algorithm is used to generate the date of the last day of each month within the provided date range. For example, if the starting date is 23-03-2013 and the ending date is 20-07-2013, this algorithm will generate the following dates: {31-03-2013, 30-04-2013, 31-05-2013, 30-06-2013, and 31-07-2013}. These dates are necessary in order to be able to extract the energy readings for specific months, since each month's reading in the database is stored against the last date of that month.

procedure dateRange (d1 , d2)

{ Comment - d1 : Starting date , d2 : Ending date }

{ Comment - D : One dimensional array of dates within the date range d1 and d2 }

begin

{ Comment - The computation in **Step 1** is done using the procedure **arraySize (d1, d2)** }

Step 1 : Compute of the number of months within the date range provided

{ Comment - The computation in **Step 2** is done using the procedure **lastDayDate (M, Y)** }

Step 2 : Generate last day's date for each month in the date range obtained in **Step 1**

return D

end

Pseudocode For Array Size Determination

The detailed implementation of the pseudocode which follows below is given in **Appendix A4**. This algorithm produces an integer value which is a count of the number of months in the date range that is passed to the procedure that has just been defined.

procedure arraySize (d1 , d2)

{ Comment - d1 : Starting date , d2 : Ending date }

{ Comment - S : Number of months within the date range d1 to d2 }

begin

Count the number of months within the date range d1 to d2

return S

end

Pseudocode For Generating The Last Date Of A Given Month

Appendix A5 has the detailed implementation of the pseudocode below. The algorithm computes the date of the last day of a given month.

procedure lastDayDate (M , Y)

{ Comment - M : Integer value of month }

{ Comment - Y : Year }

{ Comment - D : Date in the format "Year – Month – Day" }

begin

Compute the last date in the given month M

return D

end

3.1.2 STEP 2: COMPUTATION OF AUTOCORRELATION COEFFICIENT

In applying the projection technique to a time series analysis, it is assumed that the data values given in step 1 are related to each other at one or more time periods apart. The extent of the relationship can be measured by taking two data sequences from the time series, one lagging the other by one or more time periods and calculating the autocorrelation coefficient between the two sequences. The value of the autocorrelation coefficient ranges from -1 to +1 and the closer it is to +1, the better the relationship between the data sequence for the purposes of this research. For the time series with *n* data points given in step 1, the autocorrelation coefficient (r_k) for two data points *k* periods apart is given as

$$r_k = \frac{\sum_{t=1}^{n-k}\left(E_t - \overline{E}\right)(E_{t+k} - \overline{E})}{\sum_{t=1}^{n}(E_t - \overline{E})^2} \tag{3.10}$$

$$-1 \leq r_k \leq 1\,, r_k \in R$$

$$1 \leq k < n\,, k \in Z$$

where E_t is the energy consumption for the t^{th} month and $\overline{E}$ is the average value of the historical energy consumption time series.

The average value of the time series is given as

$$\overline{E} = \frac{1}{n}\sum_{t=1}^{n} E_t \tag{3.11}$$

<u>Pseudocode For Computing Autocorrelation Coefficients</u>

The pseudocode for the autocorrelation correlation coefficient computation is given below with the detailed implementation in **Appendix A6**. The historical consumption data that is extracted from the database is passed to this algorithm for the autocorrelation coefficient computation.

procedure correlationComputation (H)

{ Comment - H : Array of historical energy consumption data extracted from database }

{ Comment - A : Array of autocorrelation coefficient }

{ Comment - N : Maximum period interval for which autocorrelation coefficient is computed }

begin

k := 1

{ Comment - The computation in **Step 1** is done using the procedure **averageConsumption (H)** }

Step 1 : Compute the average consumption for the historical energy consumption data

Step 2 : Compute the denominator of Equation 3.10

Step 3 :

while k <= N

begin

Compute numerator of Equation 3.10

Compute autocorrelation coefficient for period k using equation 3.10

k := k + 1

end

return A

end

<u>Pseudocode For Computing Average Consumption</u>

Appendix A7 has the detailed implementation for the pseudocode below. This algorithm is used to compute the average consumption of a given historical energy consumption time series.

procedure averageConsumption (H)

{ Comment - H : Array of historical energy consumption data extracted from database }

{ Comment - Av : Computed average consumption }

begin

Compute the average consumption using Equation 3.11

return Av

end

3.1.3 STEP 3: COMPUTATION OF CONSTANTS *k1* AND *k2*

The linear form of the Logistic equation (*Equation 3.06*) which is stated below is applied to the historical energy consumption data given is step 1 by regression analysis to compute k_1 and k_2

$$ln\left|\frac{E}{(E_{max} - E)}\right| = k_1 t + k_2$$

The constants k_1 and k_2 are computed using the linear regression equations given below

$$k_1 = \frac{n\sum tE - (\sum t)(\sum E)}{n(\sum t^2) - (\sum t)^2} \quad (3.12)$$

$$k_2 = \frac{\sum E - k_1 \sum t}{n} \quad (3.13)$$

Pseudocode For Constants Computation

The detailed implementation of the algorithm for computing the constants k_1 and k_2 is given in **Appendix A8.** The pseudocode is given below.

procedure constantsComputation (H)

{ Comment - H : Array of historical energy consumption data extracted from database }

{ Comment - K : One dimensional array to hold constants k1 and k2 }

begin

Step 1 : Compute k1 using Equation 3.12

Step 2 : Compute k2 using Equation 3.13

K := { k1, k2}

return K

end

3.1.4 STEP 4: COMPUTATION OF OPTIMAL ASYMPTOTE

The computation of the optimal asymptote is done using the Fibonacci Search Technique which is an optimization technique. The objective function which will be minimized is the Sum of Squared Residual (SSR). The process is enumerated below:

3.1.4.1 SSR AND MAPE COMPUTATION

When the model is fitted to historical energy consumption data set as in the case of the linear form of the Logistic equation, the model is assessed on the basis of how well it fits the historical energy consumption data. This is referred to as the *goodness of fit* and how well it could be used to estimate future consumption values is referred to as the “forecasting accuracy”. The “goodness of fit” is computed as the sum of squared residuals (SSR) which is given as

$$SSR = \sum_{t=1}^{n} \left(\overline{E_t} - E_t\right)^2 \tag{3.14}$$

and the forecasting accuracy is computed as the mean absolute percentage error (MAPE) which is given as

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \left(\frac{E_t - \overline{E_t}}{E_t} \times 100 \right) \right| \tag{3.15}$$

where,

E_t is the actual consumption value,

$\overline{E_t}$ is the corresponding predicted values, and

n is the number of data points used.

<u>Pseudocode For Sum Of Squared Residuals and Mean Absolute Percentage Error</u>

The pseudocode for Equation 3.14 and 3.15 is given below with the detailed implementation in **Appendix A9 and A10** respectively.

procedure sumOfSquaredResiduals (H)

{ Comment - H : Two dimensional array with two rows, first row contains predicted energy consumption data and second row contains actual energy consumption data }

{ Comment - SSR : Computed value of Sum of Squared Residuals }

begin

Compute Sum of Squared Residuals using Equation 3.14

return SSR

end

procedure meanAbsolutePercentageError (A , P)

{ Comment - A : Array of historical energy consumption data extracted from database }

{ Comment - P : Array of computed forecast consumption }

{ Comment - M : Computed value of Mean Absolute Percentage Error }

begin

Compute M using Equation 3.15

return M

end

3.1.4.2 FIBONACCI NUMBER GENERATION

The application of the Fibonacci search technique involves one variable where the assumption of unimodality holds. While locating the optimal asymptote (E_{max}), the sum of the squared residuals (SSR) is minimized between the fitted curve and the actual data. The Fibonacci search technique involves the use of Fibonacci numbers and is generated by the expression

$$F_k = F_{k-1} + F_{k-2} \quad for\ k > 1 \quad with\ F_0 = 1\ and\ F_1 = 1 \tag{3.16}$$

The first *k* Fibonacci numbers are generated, where the number of terms *k* determines the number of iterations that will be performed in the computation of the optimal asymptote. The value of *k* is computed from the relation

$$\frac{1}{F_k} < \frac{r}{100} \tag{3.17}$$

Where r % is the interval of uncertainty and F_k is the k^{th} Fibonacci term.

<u>Pseudocode For Fibonacci Number Generation</u>

The detailed implementation of the Fibonacci number generation algorithm is given in **Appendix A11**. The pseudocode is given below:

procedure fibonacciNumbers (I)

{ Comment - I : Interval of uncertainty }

{ Comment - k : The lowest integer value for which Equation 3.17 applies }

{ Comment - F : Array of Fibonacci numbers }

begin

Step 1 : Iteratively compute the value of F_k from Equation 3.16 until the condition of Equation 3.17 is satisfied

Step 2 : Terminate the computation once the condition of Equation 3.17 is satisfied

return F

end

3.1.4.3 FIBONACCI SERCH TECHNIQUE

In computing the optimal asymptote (E_{max}) we proceed as follows:

I. A range of values of the energy consumption data E between which the optimal asymptote (E_{max}) could be located is established. The lower limit represented as E_L is taken as the maximum monthly energy consumption value recorded. It is obtained from the known historical energy consumption data. The upper limit E_U is taken as 100 times the value of E_L. If no solution of E is found between the selected range, then it can be concluded that the historical energy consumption data has little influenced on the optimal asymptote.

II. Two equidistant points E^I and E^{II} from each end of the interval is established as shown in **Figure 3.1**

$$E^I = E_L + d_1 \tag{3.18}$$

$$E^{II} = E_U - d_1 \tag{3.19}$$

The distance d_1 is defined as

$$d_1 = \frac{F_{k-2}}{F_k} \times L \tag{3.20}$$

with

$$L = E_U - E_L \tag{3.21}$$

where,

F_{k-2} is the $(k-2)_{th}$ Fibonacci number,

F_k is the k_{th} Fibonacci number,

(k-1) is the number of calculations to be performed.

Two calculations of SSR are made at the point E^Is and E^{II}, namely SSR_1 and SSR_2 respectively,

where

$$SSR_1 = \sum_{t=1}^{n} \left(\overline{E_t} - E_t\right)^2 \tag{3.22}$$

with

$$\overline{E_t} = \frac{E^I}{1 + e^{-(k_1 t + k_2)}} \tag{3.23}$$

and

$$SSR_2 = \sum_{t=1}^{n} (\overline{E_t} - E_t)^2 \tag{3.24}$$

with

$$\overline{E_t} = \frac{E^{II}}{1 + e^{-(k_1 t + k_2)}} \tag{3.25}$$

III. A new interval E_L' and E_U' is established based on the relative magnitudes of SSR_1 and SSR_2 that have been calculated in the previous step. If $SSR_1 < SSR_2$, then the region E^{II} and E_U is discarded and the new upper limit is chosen as $E_U' = E^{II}$ and the lower limit is maintained as $E_L' = E_L$. Alternatively if $SSR_1 > SSR_2$, then the lower region E_L and E^I is discarded such that $E_U' = E^U$ and $E_L' = E^I$. This narrows the interval in which the optimal asymptote would lie based on the minimum SSR.

IV. The new value of the narrowed interval L' and a new value for the distance d_2 is then calculated as follows:

$$L' = E_U' - E_L' \tag{3.26}$$

$$d_2 = \frac{F_{k-3}}{F_{k-1}} \times L' \tag{3.27}$$

The new value of E^I and E^{II} and their corresponding SSR_1 and SSR_2 are computed.

V. Steps III and IV are repeated, replacing old values of d and L with new values each time until all the *(k-1)* iterations are performed.

At the final iteration E^I and E^{II} will be very close together as the interval L is narrowed significantly by the Fibonacci search technique at each step. At this point the optimal asymptote E_{max} is chosen as the value of E with the lowest SSR.

Upon the completion of this stage, *Equation 3.07* can now be used to compute the forecast consumption. The weighted moving average algorithm is then further used in smoothening out the forecast consumption data in order to identify the required future consumption trend.

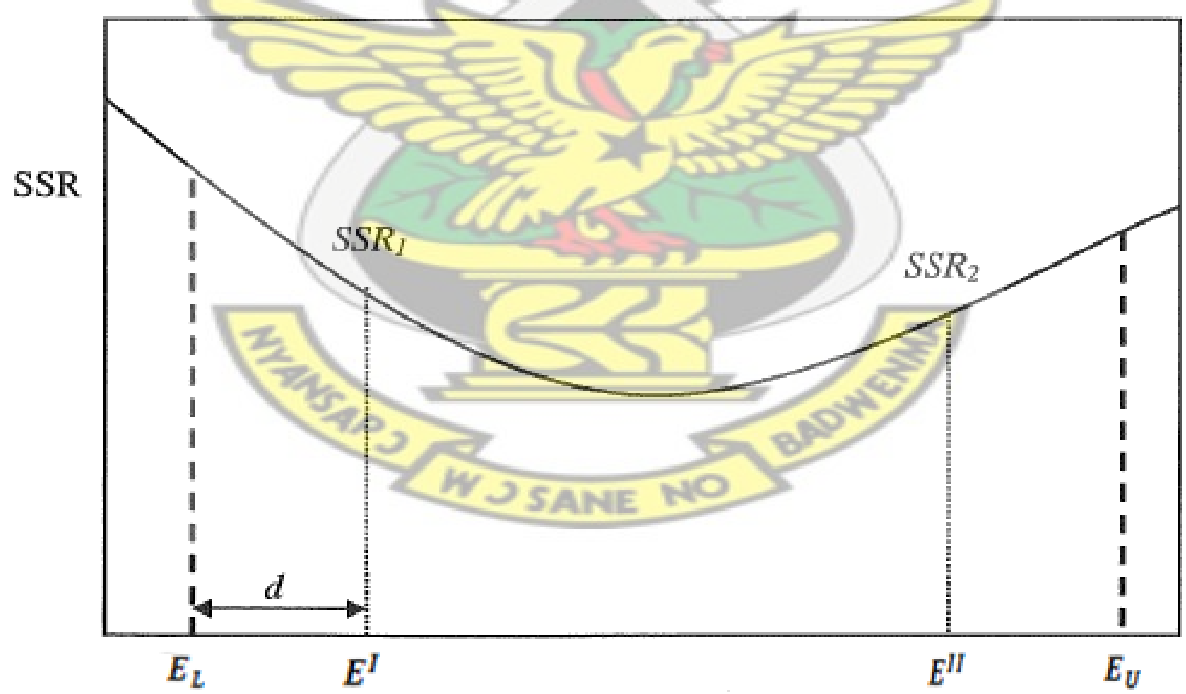

Figure 3.1 Minimizing SSR using the Fibonacci Search Technique

Pseudocode For Fibonacci Search Technique

The detailed implementation of the Fibonacci Search Algorithm for computing optimal asymptote is given in **Appendix A12** with the pseudocode given below:

procedure optimumAsymptote (H , I)

{ Comment - H : Array of historical energy consumption data extracted from database }

{ Comment - I : Interval of uncertainty }

{ Comment - Op : Optimal Asymptote }

begin

{ Comment - The **Step 1** is done using the procedure **fibonacciNumbers (I)** }

Step 1 : Compute the Fibonacci Numbers to be used

{ Comment - The **Step 2** is done using the procedure **constantsComputation (H)** }

Step 2 : Compute the constants k1 and k2 to be used

Step 3 : Loop through the historical energy consumption data to retrieve the maximum consumption

Step 4 : Compute the energy consumption range within which the optimal asymptote will be located

Step 5 : Apply the Fibonacci Search Technique until the required number of iterations are completed

Step 6 : Obtain the Energy consumption that resulted in the minimum Sum of Squared Residuals

{ Comment - The energy consumption in **Step 6** is returned as the Optimal Asymptote }

return Op

end

CHAPTER 4

DATA ANALYSIS AND RESULTS

After the formulation of the model in the previous chapter, it is necessary to put the model into practice in order to get a clear understanding of how the model works. This chapter focuses on using the model to analyze some sample data which has been extracted from the Automatic Meter Reading database. Computations at the various stages of the model are performed and analysis done to show the potential of the model.

4.0 ANALYSIS OF CUSTOMER ENERGY CONSUMPTION

The consumption pattern of a single customer is analyzed in this section and used to forecast future consumption for thirty-six months. The thirty-six months include the period of the historical consumption data thereby enabling a comparison of the historical consumption and the forecast consumption.

4.0.1 DATA COLLECTION

The historical energy consumption data used for the analysis is available in the Automatic Meter Reading central database where it is extracted for the analysis. This data is the historical energy consumption data of a selected consumer. The model can however be extended to the total energy consumption of a group of consumers.

The data in **Table 4.1** is the historical energy consumption data of a selected consumer. The data ranges from January 2012 to February 2013, a total of fourteen months (14).

Months	Time (t)	Consumption (KWH)
JAN 2012	1	1010.000
FEB 2012	2	1493.060
MAR 2012	3	5000.029
APR 2012	4	6002.968
MAY 2012	5	10003.005
JUN 2012	6	12991.445
JUL 2012	7	14009.799
AUG 2012	8	17089.809
SEP 2012	9	17001.024
OCT 2012	10	22008.991
NOV 2012	11	22991.282
DEC 2012	12	23998.697
JAN 2013	13	25740.014
FEB 2013	14	25970.236

Table 4.1 Historical Energy Consumption Data for a Customer

4.0.2 COMPUTATION OF AUTOCORRELATION COEFFICIENTS

The autocorrelation coefficients range from -1 to 1. When the autocorrelation coefficients are high, i.e. close to ± 1, it implies that the data points are closely correlated at the specified time period. However, for the purposes of this research, the preferred option is when the autocorrelation coefficient is close to +1. This means that it is statistically significant to make a forecast to the next time period based on the present data. A summary of the autocorrelation coefficient r_k values for time lags of $k = 1, 2, 3, 4, \ldots, 13$ months interval is given in **Table 4.2** below. The values give a fairly good idea of the level of correlation between the data points in the time series.

From the **Table 4.2**, it can be noticed that the autocorrelation coefficient is relatively high when the data points are one period apart. It kept reducing until the data points are four periods apart after which it eventually enters into the negative but reasonably close to the positives. As a result of the relatively good autocorrelation coefficient for the first few time intervals, the forecasting model can be applied to the historical consumption data in **Table 4.1** above should give a reasonably good forecast. The various formulas and processes used in the computation of the autocorrelation coefficient are illustrated after the table below.

Period_(k)	Correlation_Coefficient_(rk)
1	0.8082
2	0.5915
3	0.3890
4	0.1696
5	-0.0129
6	-0.1399
7	-0.2696
8	-0.3515
9	-0.4376
10	-0.4306
11	-0.3731
12	-0.2927
13	-0.1505

Table 4.2 Autocorrelation Coefficient for Historical Energy Consumption Data

In computing the autocorrelation coefficient, the average consumption from **Table 4.1** is computed using the formula below which was previously stated in chapter 3. The meaning of the symbols used in these equations is as stated in chapter three.

From the **Table 4.1**, n = 14 from which calculation is done to obtain

$$\overline{E} = \frac{1}{n}\sum_{t=1}^{n} E_t$$

$$\overline{E} = 14665.026 \text{ KWH}$$

The autocorrelation coefficient from one month up to 13 months period apart is computed from the formula:

$$r_k = \frac{\sum_{t=1}^{n-k}\left(E_t - \overline{E}\right)\left(E_{t+k} - \overline{E}\right)}{\sum_{t=1}^{n}\left(E_t - \overline{E}\right)^2}$$

The result is illustrated in **Table 4.2** above. The maximum periods apart which is thirteen (13) months was chosen to be one month less than the total number of months in the historical energy consumption time series.

4.0.3 COMPUTATION OF CONSTANTS

The linear form of the logistic equation is fitted to the historical energy consumption data in **Table 4.1** from which the constants k_1 and k_2 are computed by regression analysis.

From **Table 4.1**

n = 14, $\sum tE$ = 2017949.917, $\sum t$ = 105.000, $\sum t^2$ = 1015.000, $\sum E$ = 205310.359

$$k_1 = \frac{n\sum tE - (\sum t)(\sum E)}{n(\sum t^2) - (\sum t)^2} \qquad k_1 = 2101.636$$

$$k_2 = \frac{\sum E - k_1 \sum t}{n} \qquad k_2 = -1097.245$$

After the computation of the constants k_1 and k_2 the values are rationalized by dividing k_1 by itself and dividing k_2 by k_1 to obtain $k_1 = 1.00$ and $k_2 = -0.522$. The values needed to be rationalized because the original values of the constants have an unwanted effect on the model results due to their large values.

4.0.4 COMPUTATION OF OPTIMAL ASYMPTOTE

Using an interval of uncertainty of 0.010 % produces $k = 19$ which requires the generation of the first 20 terms of the Fibonacci numbers as shown in **Table 4.3** below.

Record Number	Fibonacci Number (Fk)
0	1
1	1
2	2
3	3
4	5
5	8
6	13
7	21
8	34
9	55
10	89
11	144
12	233
13	377
14	610
15	987
16	1597
17	2584
18	4181
19	6765

Table 4.3 First 20 Terms of the Fibonacci Numbers

The energy consumption interval within which the optimal asymptote (E_{max}) could be located is computed. The lower limit E_L is the maximum consumption value from the historical energy consumption data in **Table 4.1**.

From **Table 4.1** the maximum consumption is 25970.236 KWH, which implies

E_L= 25970.236 KWH

E_U = 100 *E_L

E_U = 2597023.600 KWH

The Fibonacci numbers in Table 4.3 and the consumption range E_L to E_U is used to compute the optimal asymptote (E_{max}) using the Fibonacci Search technique as modeled in chapter 3.

From the iteration process of the Fibonacci Search technique, the optimal asymptote is computed and given as:

$$E_{max} = 37162.591 \text{ KWH}$$

From the Logistic equation, the forecast consumption for 36 months starting from the first month of the historical consumption is shown in **Table 4.4** below.

Time (Months)	Consumption (KWH)
JAN 2012	1111.000
FEB 2012	1642.366
MAR 2012	5500.032
APR 2012	6603.265
MAY 2012	11003.306
JUN 2012	14290.590
JUL 2012	15410.779
AUG 2012	16748.013
SEP 2012	18701.126
OCT 2012	21568.811
NOV 2012	22531.456
DEC 2012	23518.723
JAN 2013	25225.214
FEB 2013	27201.706
MAR 2013	28433.176
APR 2013	28979.588
MAY 2013	29461.713
JUN 2013	29890.269
JUL 2013	30273.713
AUG 2013	30618.813
SEP 2013	30931.046
OCT 2013	31214.894
NOV 2013	31474.060
DEC 2013	31711.628
JAN 2014	31930.191
FEB 2014	32131.942
MAR 2014	32318.748
APR 2014	32492.211
MAY 2014	32653.710
JUN 2014	32804.443
JUL 2014	32945.452
AUG 2014	33077.647
SEP 2014	33201.831
OCT 2014	33318.710
NOV 2014	33428.909
DEC 2014	33532.987

Table 4.4 Thirty-Six (36) Months Forecast Energy Consumption for Customer

The optimal asymptote of the Logistic Equation has significant effect on the forecasts made. Once the historical data is used to get the best fit, i.e. to get the constants of the regression equation, it is the upper limit that determines the accuracy of the forecast. While the Fibonacci Search Technique used to compute the asymptote has been proven to be an effective method, the value of the asymptote will depend on the extent of the data used in determining its value. The more the available historical consumption data, the more accurate the asymptote value that is obtained.

The graph in **Figure 4.1** illustrates the historical consumption and the forecast consumption plotted on the same set of axis. From the forecast consumption curve as illustrated with a blue curve in **Figure 4.1**, it can be observed that the consumption increased steadily from January 2012 until March 2013 after which the graph slowed down in its growth. Between January 2012 and February 2013, the historical consumption is modeled using the developed model after which the actual forecast is done from March 2013 onwards.

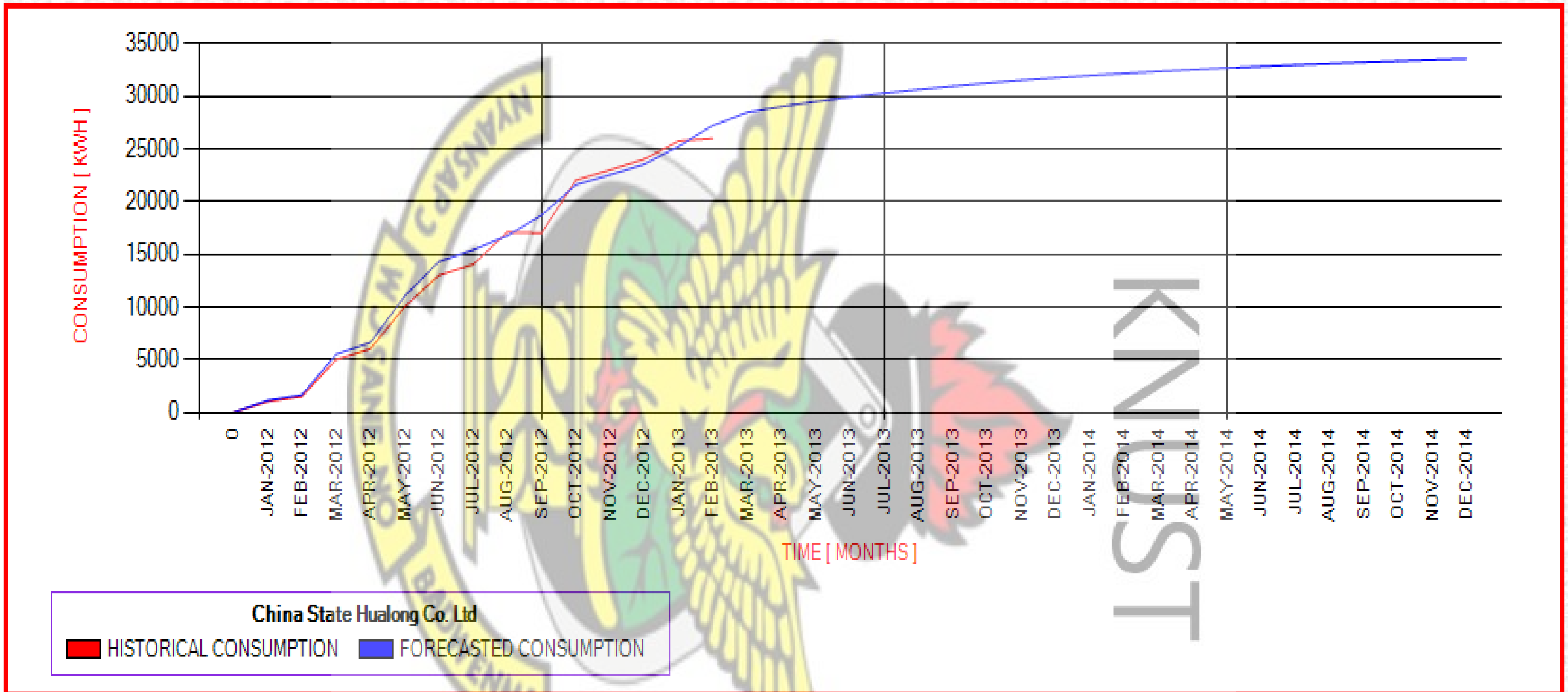

Figure 4.1 Graph of Time against Energy Consumption for Customer

4.0.5 COMPARISON OF HISTORICAL AND FORECAST CONSUMPTION DATA

The data provided in **Table 4.5** is a comparison of the forecast data produced by the model against the actual historical data. The comparison is done from January 2012 to February 2013. The first column has the period, the second column has the forecast consumption generated by the model and the third column has the historical consumption data. The fourth column has the difference in value between the forecast and the historical consumption while the fifth column depicts the percentage difference. The percentage difference is computed relative to the historical consumption for each period. **Figure 4.2** is a graph that spans the period for which historical data is available. It graphically illustrates a comparison of the historical and forecast data as given in **Table 4.5**. It could be observed from the table that the forecast value at maximum is 10% more than the historical value and at minimum 2% less than the historical value.

Time (Months)	Forecast - F (KWH)	Historical - H (KWH)	(F - H) (KWH)	(F - H)*100 / H (% Difference)
JAN 2012	1111.000	1010.000	101.000	10.00
FEB 2012	1642.366	1493.060	149.306	10.00
MAR 2012	5500.032	5000.029	500.003	10.00
APR 2012	6603.265	6002.968	600.297	10.00
MAY 2012	11003.306	10003.005	1000.301	10.00
JUN 2012	14290.590	12991.445	1299.145	10.00
JUL 2012	15410.779	14009.799	1400.980	10.00
AUG 2012	16748.013	17089.809	-341.796	-2.00
SEP 2012	18701.126	17001.024	1700.102	10.00
OCT 2012	21568.811	22008.991	-440.180	-2.00
NOV 2012	22531.456	22991.282	-459.826	-2.00
DEC 2012	23518.723	23998.697	-479.974	-2.00
JAN 2013	25225.214	25740.014	-514.800	-2.00
FEB 2013	27201.706	25970.236	1231.470	4.74

Table 4.5 Comparison of Historical and Forecast Energy Consumption Data

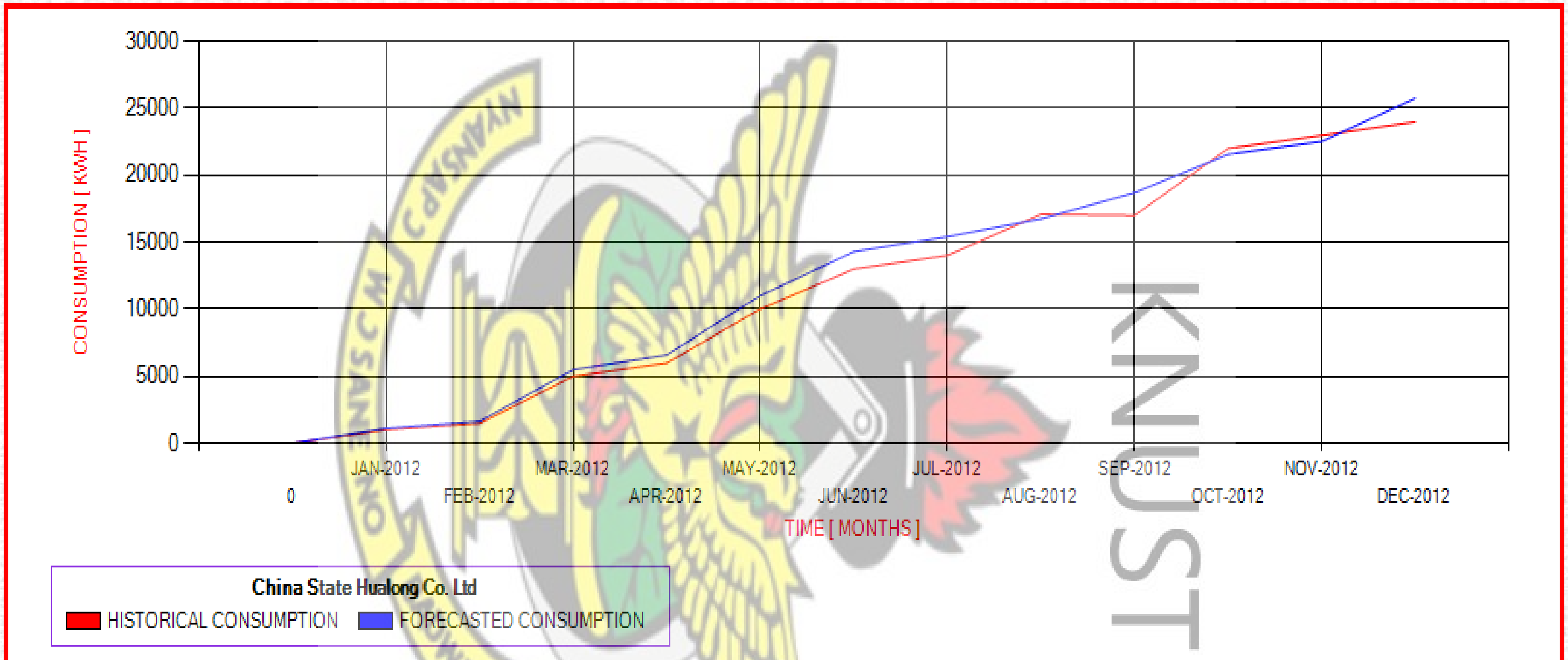

Figure 4.2 Comparison Graph for Historical and Forecast Energy Consumption Data

4.0.6 QUALITATIVE ANALYSIS OF THE FORECAST DATA

From ***Equation 3.15*** the Mean Absolute Percentage Error (MAPE) which is a measure of the forecasting accuracy is determined. Using the forecast and historical consumption given in **Table 4.5**, the computation is done and given as:

$$\text{MAPE} = 6.77$$

From the value obtained for the MAPE, the forecast consumption on average is expected to have a value of ± 6.77 % of the historical consumption.

In Section 3.4 of chapter three, the equilibrium solution of the Logistic differential equation that is ***Equation 3.01*** was determined. From the equilibrium solution that was determined we have

$$E = 0 \quad and \quad E = E_{max} = 37162.591$$

Based on the equilibrium solution obtained in the analysis, the forecast consumption is consequently expected to fall in the range of $0 - 37162.591$ KWH.

4.1 FURTHER ANALYSIS OF CUSTOMER ENERGY CONSUMPTION

Further analysis is done using a customer with historical energy consumption pattern slightly different from the one that was previously done. This is to demonstrate the effect of a different historical consumption patterns on forecast consumption. **Table 4.6** is the historical consumption of another customer with a historical consumption pattern different from that of **Table 4.1**.

Time (Months)	Time (*t*)	Consumption (KWH)
JAN 2012	1	1041.490
FEB 2012	2	1781.570
MAR 2012	3	4680.029
APR 2012	4	6522.968
MAY 2012	5	9483.005
JUN 2012	6	13191.445
JUL 2012	7	13809.799
AUG 2012	8	17329.809
SEP 2012	9	18321.024
OCT 2012	10	20448.991
NOV 2012	11	22991.282
DEC 2012	12	23998.697
JAN 2013	13	25340.014
FEB 2013	14	29130.236

Table 4.6 Historical Consumption Data for another Customer

The autocorrelation coefficient r_k values for time lags of $k = 1, 2, 3, 4, \dots, 13$ months interval is given in **Table 4.7** below. The values are an indication of the level of correlation between the data points in the time series.

From the **Table 4.7**, it can also be realised that the autocorrelation coefficient is relatively high when the data points are one period apart. It kept reducing until the data points are five periods apart, after which it eventually enters into the negative but relatively close to the positives as was noted in the previous application of the model. As a result of the relatively good autocorrelation coefficient for first few time intervals, the forecasting model can be applied to the historical consumption data in **Table 4.6** above. The computation process of the autocorrelation coefficient is illustrated after the table below.

Period_(k)	Correlation_Coefficient_(rk)
1	0.7818
2	0.5784
3	0.3820
4	0.1743
5	0.0045
6	-0.1244
7	-0.2644
8	-0.3333
9	-0.4079
10	-0.4235
11	-0.3782
12	-0.3068
13	-0.1825

Table 4.7 Autocorrelation Coefficient for Historical Energy Consumption Data

The average consumption from **Table 4.6** is computed as illustrated below. The meaning of the symbols used in these equations is as stated in chapter three.

From the **Table 4.6**, n = 14 from which calculation is done to obtain

$$\overline{E} = \frac{1}{n}\sum_{t=1}^{n} E_t$$

$$\overline{E} = 14862.169 \text{ KWH}$$

The values of the autocorrelation coefficient from one month up to 13 months period apart are computed as shown in **Table 4.7** using the previously stated autocorrelation formula.

From **Table 4.6**

n = 14, $\sum tE$ = 2054118.427, $\sum t$ = 105.000, $\sum t^2$ = 1015.000, $\sum E$ = 208070.359

$$k_1 = \frac{n\sum tE - (\sum t)(\sum E)}{n(\sum t^2) - (\sum t)^2} \qquad k_1 = 2169.630$$

$$k_2 = \frac{\sum E - k_1 \sum t}{n} \qquad k_2 = -1410.054$$

Following the computation of the constants k_1 and k_2 the values are once again rationalized to obtain $k_1 = 1.00$ and $k_2 = -0.650$.

The computation of the optimal asymptote and the forecast is done using the procedure as was done in the previous application under section 4.0.4. As was done in the previous example, the energy consumption interval within which the optimal asymptote (E_{max}) could be located is computed. The lower limit E_L is the maximum consumption value from the historical energy consumption data in **Table 4.6**.

From **Table 4.6** the maximum consumption is 29130.236 KWH, which implies

E_L= 29130.236 KWH

E_U = 100 *E_L

E_U = 2913023.600 KWH

The Fibonacci numbers in Table 4.3 and the consumption range E_L to E_U is used to compute the optimal asymptote (E_{max}) using the Fibonacci Search technique as was done in the previous application of the model.

After the required number of iterations using the Fibonacci Search technique, the optimal asymptote is obtained as:

$$E_{max} = 40157.101 \text{ KWH}$$

From the Logistic equation, the forecast consumption for 36 months starting from the first month of the historical consumption is shown in **Table 4.8** below.

The graph in **Figure 4.3** illustrates the historical consumption and the forecast consumption plotted on the same set of axis. The red curves representing the historical consumption curve while the blue curve representing the forecast curve.

Time (Months)	Consumption (KWH)
JAN 2012	1145.639
FEB 2012	1959.727
MAR 2012	5148.032
APR 2012	7175.265
MAY 2012	10431.306
JUN 2012	14510.590
JUL 2012	15190.779
AUG 2012	16983.213
SEP 2012	17954.604
OCT 2012	20040.011
NOV 2012	22531.456
DEC 2012	23518.723
JAN 2013	24833.214
FEB 2013	29646.223
MAR 2013	30162.210
APR 2013	30787.833
MAY 2013	31339.851
JUN 2013	31830.533
JUL 2013	32269.564
AUG 2013	32664.691
SEP 2013	33022.188
OCT 2013	33347.185
NOV 2013	33643.921
DEC 2013	33915.929
JAN 2014	34166.176
FEB 2014	34397.174
MAR 2014	34611.061
APR 2014	34809.670
MAY 2014	34994.582
JUN 2014	35167.166
JUL 2014	35328.616
AUG 2014	35479.976
SEP 2014	35622.162
OCT 2014	35755.984
NOV 2014	35882.159
DEC 2014	36001.325

Table 4.8 Thirty-Six (36) Months Forecast Energy Consumption for Customer

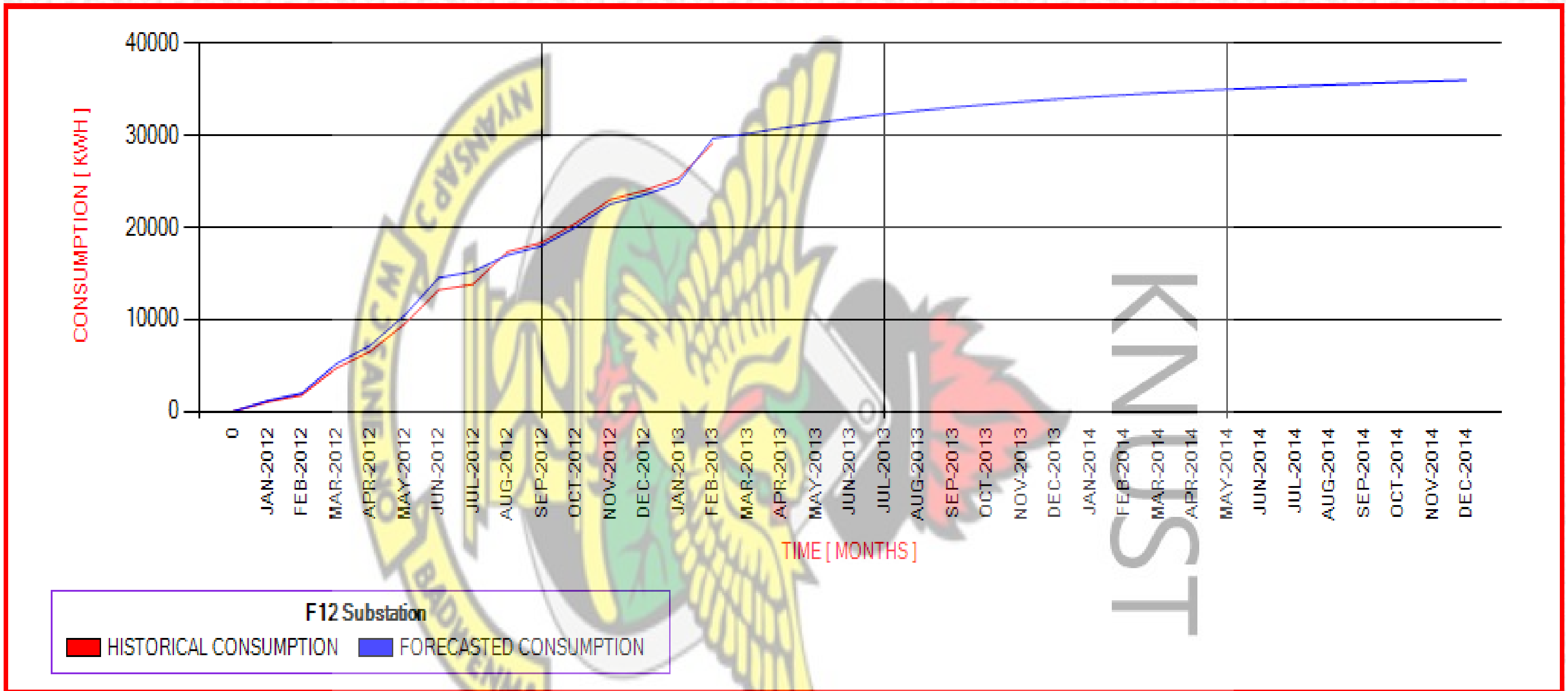

Figure 4.3 Graph of Time against Energy Consumption for another Customer

Table 4.9 is a comparison of the forecast and actual historical data. The comparison is done as it was performed in the previous example. In the last column where the percentage difference is computed, it can be observed that the percentage difference is 10% for the first seven months. Which means from January 2012 to July 2012 the forecast consumption is 10% above the actual historical consumption. From August 2012 to January 2013 the forecast consumption is 2% below the actual historical consumption. Finally the forecast consumptions in February 2013 is 1.77% above the actual historical consumption for that month. **Figure 4.4** is a graphical presentation of the historical and forecast values plotted on the same set of axis. The values used in plotting the graph are taken from the values in **Table 4.9**.

Time (Months)	Forecast - F (KWH)	Historical - H (KWH)	(F - H) (KWH)	(F - H)*100 / H (% Difference)
JAN 2012	1145.639	1041.490	104.149	10.00
FEB 2012	1959.727	1781.570	178.157	10.00
MAR 2012	5148.032	4680.029	468.003	10.00
APR 2012	7175.265	6522.968	652.297	10.00
MAY 2012	10431.306	9483.005	948.301	10.00
JUN 2012	14510.590	13191.445	1319.145	10.00
JUL 2012	15190.779	13809.799	1380.980	10.00
AUG 2012	16983.213	17329.809	-346.596	-2.00
SEP 2012	17954.604	18321.024	-366.420	-2.00
OCT 2012	20040.011	20448.991	-408.980	-2.00
NOV 2012	22531.456	22991.282	-459.826	-2.00
DEC 2012	23518.723	23998.697	-479.974	-2.00
JAN 2013	24833.214	25340.014	-506.800	-2.00
FEB 2013	29646.223	29130.236	515.987	1.77

Table 4.9 Comparison of Historical and Forecast Energy Consumption Data

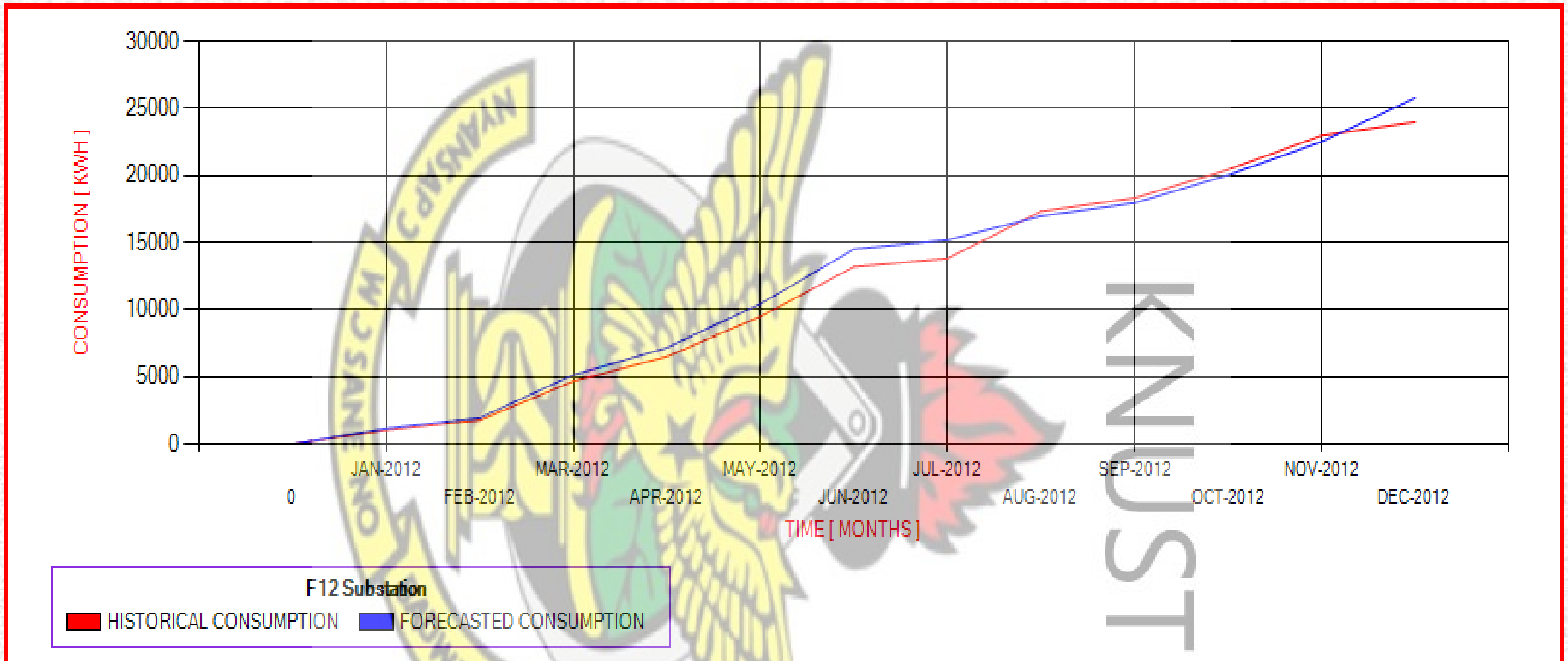

Figure 4.4 Comparison Graph for Historical and Forecast Energy Consumption Data

The Mean Absolute Percentage Error (MAPE) is computed using ***Equation 3.15*** by applying it to forecast and historical consumption data given in **Table 4.9.** The result obtained is given as:

$$\text{MAPE} = 5.98$$

From the value obtained for the MAPE, the forecast consumption on average is expected to have a value of ± 5.98% of the historical consumption.

The equilibrium solution of the Logistic differential equation given as ***Equation 3.01*** is obtained below:

$$E = 0 \quad and \quad E = E_{max} = 40157.101$$

The forecast consumption is expected to be in the range $0 - 40157.101$ KWH.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.0 CONCLUSION

This research was used to investigate the Logistic growth curve model for forecasting industrial and commercial electricity consumption. The Electricity Company of Ghana Limited was used as a case study and only the company's customer data were considered. The consumption data of customers were obtained through an Automatic Meter Reading Systems which enables the remote reading of customer's energy meter.

A mathematical model based on the Logistic differential equation was used as the basis for the development of the model. The rate of change of energy consumption was expressed in the form of the logistic differential equation, after which the analytical solution for the equation was obtained. The constants in the analytical solution were obtained by analyzing historical energy consumption data using Linear Regression. The carrying capacity of the Logistic equation referred to as the Optimal Asymptote in this research was obtained using the Fibonacci Search Technique. The Fibonacci Search Technique was equally applied to the historical energy consumption data that was used to compute the constants.

After the computation of constants and parameter in the analytical solution, electricity consumption forecast was done for the customer under consideration. The forecast consumption was compared with the actual historical consumption in order to ascertain the level of disparity of the forecast from the actual. Further qualitative analysis was done by the computation of the Mean Absolute Percentage Error which is a measure of the forecasting accuracy of the model.

5.1 RECOMMENDATION

This thesis was limited to the development of a mathematical model for forecasting electricity consumption of consumers with logistic growth in their historical consumption. The model generated a good forecast for consumers whose historical energy consumption had logistic growth. Therefore, it is recommended that any future research should concentrate on developing a forecasting model for consumers with non logistic growth.

REFERENCES

1. Aboagye-Sarfo P., "MATH 518: Statistical Models", Kwame Nkrumah University of Science and Technology, Kumasi, Faculty of Distance Learning, 2007.
   Unit 1; Unit 2.
2. Amponsah K.O, "MATH 507: Optimization Techniques 1", Kwame Nkrumah University of Science and Technology, Kumasi, Institute of Distance Learning, 2009.
   Unit 1, Session 3-1.
3. Bewley, R. and Fiebig, D.G. (1998), "A flexible logistic growth model with applications in telecommunications", *International Journal of Forecasting*, Vol. 4, pp. 177-192, 1988.
4. Bewley, R. and Griffiths, W.E. (2003), "The penetration of CDs in the sound recording market: issues in specification, model selection and forecasting", *International Journal of Forecasting,* Vol. 19, pp.111-121, 2003.
5. Bianchi, L., Jarrett, J. and Hanumara, RC. (1998), "Improving forecasting for telemarketing centers by ARIMA modeling with intervention", *International Journal of Forecasting,* Vol. 14, pp. 497-504, 1998.
6. Dennis G. Zill, Michael R. Cullen, "Differential Equations with Boundary – Value Problems", (Seventh Edition), Brooks/Cole, Cengage Learning, 2009.
   Chapter 3
7. Dontwi I.K., "MATH 534: Modeling With Differential Equations", Kwame Nkrumah University of Science and Technology, Kumasi, Faculty of Distance Learning, 2007.
   Unit 1.

8. Dontwi I.K., "MATH 515: Differential Equation", Kwame Nkrumah University of Science and Technology, Kumasi, Institute of Distance Learning, 2006.
Module 1, Page 7; Module 2, Page 26.

9. Koutsoyiannis A., "Theory of Econometrics", (Second Edition) Palgrave Macmillan, 1977. Chapter 2.

10. MathBits.com, "Regression Analysis Basics", Accessed September 10, 2012, http://resources.esri.com/help/9.3/arcgisdesktop/com/gp_toolref/spatial_statistics_toolbox/regression_analysis_basics.htm.

11. MathBits.com, "Correlation Coefficient", Accessed September 21, 2012, http://mathbits.com/MathBits/TISection/Statistics2/correlation.htm.

12. Mohamed Z. and Bodger P.S, "Analysis of the Logistic model for Predicting New Zealand Electricity Consumption", presented at the Electricity Engineer's Association (EEA) New Zealand 2003 Conference, Christchurch, New Zealand.

13. Thomas Malthus (1798), "An Essay on the Principle of Population", 1st edition, McMaster University, Canada.

14. Wikipedia, "Regression Analysis", Accessed September 10, 2012, http://en.wikipedia.org/wiki/Regression_analysis.

15. Wikipedia, "Logistic Function", Accessed January 30, 2013, http://en.wikipedia.org/wiki/Logistic_function.

16. Wikipedia, "Moving Average", Accessed May 01, 2013, https://en.wikipedia.org/wiki/Moving_average.

17. Wikipedia, "Stability Theory", Accessed October 18, 2012, http://en.wikipedia.org/wiki/Stability_theory

18. Wikipedia, "Nonlinear System", Accessed September 12, 2012, https://en.wikipedia.org/wiki/Nonlinear_system.

19. Wikipedia, "Differential Equation", Accessed September 12, 2012, http://en.wikipedia.org/wiki/Differential_equation.

APPENDIX

The source codes of the computer program that was used to implement the model are given in this appendix. Functions used to extract data from the database and those that perform the computations are presented.

A1. THE LOGISTIC EQUATION

This function implements the Logistic Equation and is used to compute energy consumption. This is done by passing the Optimal Asymptote, the Constants and Time as arguments to the function. The computation is done and the results returned.

```
// THIS FUNCTION EXECUTES THE LOGISTIC EQUATION
// eMax      : The carrying capacity of the logistic equation.
// const_1   : The first constant (k1).
// const_2   : The second constant (k2).
// comp_time : The number of a specific month in the range (t).

public double fn_LogisticEquation(double eMax, double const_1, double const_2,
                          int comp_time)
{
   double logistic_result = 0.000;
   logistic_result = eMax / (1 + Math.Exp(-(const_1 * Convert.ToDouble(comp_time) +
                        const_2)));
   return logistic_result;
}
```

A2. EXTRACTION OF DATA FROM DATABASE

This function is used to extract the historical energy consumption for a selected consumer from the central database. This function accepts as arguments the *Meter Number* of a consumer under consideration, the *Starting* and *Ending* date interval within which data is to be extracted. This function depends on the function in **Appendix A3** for its computation.

```
//THIS FUNCTION EXTRACTS CONSUMPTION DATA FROM THE DATABASE

public string[,] fn_DataExtraction(string meterNumber, string startingDate, string
                        endingDate)
{
   string msg = "";
   string currentDate="";
   DateTime currentDate_=new DateTime();

   string[] dateRange=fn_DateRange(startingDate,endingDate);
   string[,] dataSet=new string[3,dateRange.Length];

   for (int count=0; count < dateRange.Length; count ++)
   {
      currentDate = dateRange[count];
      currentDate_ = Convert.ToDateTime(currentDate);
      try
      {
         var currentReading = (from readingTable in db.tblMeterReadings where
         readingTable.ResetDate == currentDate_ && readingTable.MeterNumber ==
                        meterNumber orderby readingTable.KWH descending
                        select readingTable.KWH).Take(1);

         var currentReading_ = currentReading.ToArray();

         if (!currentReading_[0].Equals(""))
         {
            dataSet[0, count] = dateRange[count];
            dataSet[1, count] =
                  Convert.ToDouble(currentReading_[0])/1000).ToString("#0.000");
            dataSet[2, count] = "0.000";
         }
      }
      catch (Exception ex)
```

```
        {
            msg = ex.Message;
            dataSet[0, count] = dateRange[count];
            dataSet[1, count] = "0.000";
            dataSet[2, count] = "0.000";
        }
    }

    double consumption = 0.000;

    for (int count_3 = 1; count_3 <= dataSet.GetUpperBound(1); count_3++)
    {
        consumption = Convert.ToDouble(dataSet[1, (count_3)]) –
                    Convert.ToDouble(dataSet[1,(count_3 - 1)]);

        if (consumption <= 0)
            dataSet[2, count_3] = "0.000";
        else
            dataSet[2, count_3] = consumption.ToString("#0.000");
    }

    // Smoothing of consumptions using Weighted Moving Average is used to smoothing out
      the data
    double denomenator = 0.000;
    double numerator = 0.000;

    for (int count4 = 1; count4 <= dataSet.GetUpperBound(1); count4++)
    {
        denomenator = (count4 * (count4 + 1)) / 2;
        for (int count5 = count4; count5 >= 1; count5--)
        {
            numerator = numerator + count5 * Convert.ToDouble(dataSet[2, count5]);
        }
        dataSet[2, count4] = (numerator / denomenator).ToString("0.000");
        denomenator = 0.000;
        numerator = 0.000;
    }

    return dataSet;
}
```

A3. DATE RANGE CONVERSION

This function is used by the function in **Appendix A2**. It generates an array of dates which consists of the last date of the months within the specified date range. The Date Range is passed as argument to the function. Assuming the following date range is passed to the function:

FromDate = 2013-04-27, ToDate = 2013-07-24

The function returns an array with the following dates:

{ 2013-04-30, 2013-05-31, 2013-06-30, 2013-07-31 }

This range is needed because monthly energy consumption readings in the database are stored against the last day of the month. As a result these dates are needed in order to be able to extract specific monthly readings from the database for analysis.

This function also depends on the functions in **Appendix A4** and **Appendix A5**. The function in **Appendix A4** is used to compute the array size that is needed to store the date range. The function in **Appendix A5** is used to compute the last date of each month.

```
// THIS FUNCTION PRODUCES AND ARRAY OF LAST DATE OF THE MONTH FOR
THE DATE RANGE.

public String[] fn_DateRange(String FromDate, String ToDate)
{
   int index = 0;
   int size = fn_ArraySizeNeeded(FromDate, ToDate);
   String[] dateRange = new String[size + 1]; // One is added in order to make
                              allocation for the previous date before  the FromDate

   int FromDateYear = Convert.ToUInt16(FromDate.Substring(0, 4));
   int FromDateMonth = Convert.ToUInt16(FromDate.Substring(5, 2));
   int FromDateDay = Convert.ToUInt16(FromDate.Substring(8, 2));

   int ToDateYear = Convert.ToUInt16(ToDate.Substring(0, 4));
   int ToDateMonth = Convert.ToUInt16(ToDate.Substring(5, 2));
   int ToDateDay = Convert.ToUInt16(ToDate.Substring(8, 2));
```

```
if (FromDateMonth == 1)
{
   dateRange[index] = fn_LastDayDate(12, (FromDateYear - 1));
   index++;
}
else
{
   dateRange[index] = fn_LastDayDate((FromDateMonth - 1), FromDateYear);
   index++;
}

if (FromDateYear == ToDateYear)
{
   if (FromDateMonth == ToDateMonth)
   {
      dateRange[index] = fn_LastDayDate(FromDateMonth,FromDateYear);
      index++;
   }
   else
   {
      for (int count = FromDateMonth; count <= ToDateMonth; count++)
      {
         dateRange[index] = fn_LastDayDate (count,FromDateYear);
         index++;
      }
   }
}
else
{
   for (int count1 = FromDateYear; count1 <= ToDateYear; count1++)
   {
      if (count1 == FromDateYear)
      {
         for (int count2 = FromDateMonth; count2 <= 12; count2++)
         {
            dateRange[index] = fn_LastDayDate (count2, count1);
            index++;
         }
      }
      else if (count1 == ToDateYear)
      {
         for (int count2 = 1; count2 <= ToDateMonth; count2++)
         {
            dateRange[index] = fn_LastDayDate(count2, count1);
            index++;
```

```
            }
        }
        else
        {
            for (int count2 = 1; count2 <= 12; count2++)
            {
                dateRange[index] = fn_LastDayDate(count2, count1);
                index++;
            }
        }
    }
  }
  return dateRange;
}
```

A4. ARRAY SIZE COMPUTATION

This function accepts the date range that is passed to the function in **Appendix A3** as argument. It then computes the array size that is needed to store the date range that is generated by the function in **Appendix A3.**

```
// THIS FUNCTION DETERMINES THE ARRAY SIZE THAT NEEDS TO BE
ALLOCATED.

public int fn_ArraySizeNeeded(String FromDate, String ToDate)
{
   int index = 0;

   int FromDateYear = Convert.ToUInt16(FromDate.Substring(0, 4));
   int FromDateMonth = Convert.ToUInt16(FromDate.Substring(5, 2));
   int FromDateDay = Convert.ToUInt16(FromDate.Substring(8, 2));

   int ToDateYear = Convert.ToUInt16(ToDate.Substring(0, 4));
   int ToDateMonth = Convert.ToUInt16(ToDate.Substring(5, 2));
   int ToDateDay = Convert.ToUInt16(ToDate.Substring(8, 2));



   if (FromDateYear == ToDateYear)
   {
      if (FromDateMonth == ToDateMonth)
      {
         index++;
      }
      else
      {
         for (int count = FromDateMonth; count <= ToDateMonth; count++)
         {
            index++;
         }
      }
   }
   else
   {
      for (int count1 = FromDateYear; count1 <= ToDateYear; count1++)
      {
         if (count1 == FromDateYear)
```

```
        {
          for (int count2 = FromDateMonth; count2 <= 12; count2++)
          {
            index++;
          }
        }
        else if (count1 == ToDateYear)
        {
          for (int count2 = 1; count2 <= ToDateMonth; count2++)
          {
            index++;
          }
        }
        else
        {
          index += 12;
        }
      }
    }
    return index;
  }
```

A5. OBTAINING THE LAST DATE OF A MONTH

This function accepts Month and Year as arguments and generates the date of the last day in that month. The month and year are passed as integer values to the function.

```
//THIS FUCTION GENERATES THE LAST DAY'S DATE FOR A GIVEN MONTH AND
YEAR

public string fn_LastDayDate(int month, int year)
{
    int day = 0;
    string monthI = "0";

    int[] monthOfYear = new int[] { 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 };
    int[] days = new int[] { 31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31 };
    for (int count = 0; count < 12; count++)
    {
        if (month == monthOfYear[count])
        {
            day = days[count];
            if (month == 2 && year % 4 == 0)
            {
                day++;
            }
            break;
        }
    }

    if (month < 10)
        monthI = "0" + month.ToString();
    else
        monthI = month.ToString();

    return year+"-"+ monthI +"-"+day;

}
```

A6. AUTOCORRELATION COEFFICIENT COMPUTATION

```
// THIS FUNCTION IS USED TO COMPUTE THE CORRELATION COEFFICIENTS AT K PERIODS INTERVAL
// HERE THE DEFAULT K WHICH STARTS FROM 1 TO (n-1) IS USED
// "dataSet" is a two dimensional array with three rows
// First row contains dates
// Second row contains reading
// Third row contains consumption

public string[,] fn_CorrelationComputation(string[,] dataSet)
{
    // The correlation coefficient will be computed for the period interval k=1 to k=(n-
      1)
    // Where n is the data points that are being used
    // Note that the first data point in the array "dataSet" is not considered, its
      reading was only used to compute the
    // the first consumption in the actually specified date range

    // Consumption data in "dataSet" array starts form location 1 To
      dataSet.GetUpperBound(1)

    string[,] correlation_coefficient = new string[2,dataSet.GetUpperBound(1)];
    // Values in first column of array will not be used.
    correlation_coefficient[0, 0] = "0";
    correlation_coefficient[1, 0] = "0";

    double average_consumption = 0.000;
    double denominator = 0.000;
    double numerator = 0.000;
    int while_loop_control = 1;

    // The Average consumption is first computed
    average_consumption = fn_AverageConsumption(dataSet);

    // The denominator part of the equation is computed
    for (int count_1 = 1; count_1 <= dataSet.GetUpperBound(1); count_1++)
    {
        denominator = denominator + (Convert.ToDouble(dataSet[2, count_1]) -
                average_consumption) * (Convert.ToDouble(dataSet[2, count_1]) -
                average_consumption);
    }

    // The correlation coefficient are computed
    while (while_loop_control <= correlation_coefficient.GetUpperBound(1))
```

```
  {
    // The value of k in the correlation coefficient equation is equal to the value
      of "while_loop_control"      (k = while_loop_control)
    // The value of t in the correlation coefficient equation is equal to the value
      of "count_2" (t = count_2)
    // The value of n in the correlation coefficient equation is equal to the value
      of dataSet.GetUpperBound(1)  (n = dataSet.GetUpperBound(1))

    for (int count_2 = 1; count_2 <= (dataSet.GetUpperBound(1) - while_loop_control);
       count_2++ )
    {
      numerator = numerator + (Convert.ToDouble(dataSet[2, count_2]) -
      average_consumption) * (Convert.ToDouble(dataSet[2, (count_2 +
      while_loop_control)]) - average_consumption);
    }
    correlation_coefficient[0, while_loop_control] = while_loop_control.ToString();
    correlation_coefficient[1, while_loop_control] = (numerator /
                                   denominator).ToString("#0.0000");
    numerator = 0.000;

    while_loop_control++;
  }

  return correlation_coefficient;
}
```

A7. COMPUTATION OF AVERAGE CONSUMPTION

```
// THIS FUNCTION COMPUTES THE AVERAGE CONSUMPTION
// "dataSet" is a two dimensional array with three rows
// First row contains dates
// Second row contains reading
// Third row contains consumption

public double fn_AverageConsumption(string[,] dataSet)
{
   double average_consumption = 0.000;
   for (int count = 1; count <= dataSet.GetUpperBound(1); count++ )
   {
      average_consumption += Convert.ToDouble(dataSet[2,count]);
   }

   average_consumption = (average_consumption / dataSet.GetUpperBound(1));
   return average_consumption;
}
```

A8. COMPUTATION CONSTANTS

```
// THIS FUNCTION COMPUTES THE CONSTANTS K1 AND K2 BY REGRESSION
ANALYSIS
// "dataSet" is a two dimensional array with three rows
// First row contains dates
// Second row contains reading
// Third row contains consumption

public string[] fn_ConstantsComputation(string[,] dataSet)
{
  string[] constants = new string[2];
  double sum_of_time = 0.000;
  double sum_of_time_squared = 0.000;
  double sum_of_consumption = 0.000;
  double sum_of_time_consumption = 0.000;
  double numerator = 0.000;
  double denominator = 0.000;
  double constant_1 = 0.000;
  double constant_2 = 0.000;

  for (int count = 1; count <= dataSet.GetUpperBound(1); count++)
  {
    sum_of_time = sum_of_time + count;
    sum_of_time_squared = sum_of_time_squared + count * count;
                   sum_of_consumption = sum_of_consumption +
                   Convert.ToDouble(dataSet[2,count]);
    sum_of_time_consumption = sum_of_time_consumption + count *
                      Convert.ToDouble(dataSet[2, count]);
  }

  numerator = dataSet.GetUpperBound(1) * sum_of_time_consumption - sum_of_time *
          sum_of_consumption;
  denominator = dataSet.GetUpperBound(1) * sum_of_time_squared - sum_of_time *
            sum_of_time;
  constant_1 = numerator / denominator;
  constant_2 = (sum_of_consumption - constant_1 * sum_of_time) /
            dataSet.GetUpperBound(1);
  constants[0] = (constant_1 / constant_1).ToString("#0.000");
  constants[1] = (constant_2 / constant_1).ToString("#0.000");

  return constants;
}
```

A9. COMPUTATION OF SUM OF SQUARED RESIDUALS

```
// THIS FUNCTION COMPUTES THE SUM OF SQUARED RESIDUALS
// "dataSetConsumptions" is a two dimensional array with two rows
// First row of array "dataSetConsumptions" contains the predicted data
// Second row contains the actual data.
public double fn_Sum_Of_Squared_Residuals(string[,] dataSetConsumptions)
{
  double sum = 0.000;

  for (int count = 0; count <= dataSetConsumptions.GetUpperBound(1); count++)
  {
    sum = sum + (Convert.ToDouble(dataSetConsumptions[0, count]) –
        Convert.ToDouble(dataSetConsumptions[1, count])) *
        (Convert.ToDouble(dataSetConsumptions[0, count]) –
        Convert.ToDouble(dataSetConsumptions[1, count]));
  }

  return sum;
}
```

A10. COMPUTATION OF MEAN ABSOLUTE PERCENTAGE ERROR

```
// THIS FUNCTION COMPUTES THE THE MEAN ABSOLUTE PERCENTAGE ERROR
// This function receives two arrays of actual and forecast consumption.
// "actualConsumption" is a two dimensional array with three rows and has the consumption
   in the third row.
// "preditedConsumption" is a two dimensional array with two rows and has the consumption
   in the second row.
public double fn_MeanAbsolutePercentageError(string[,] actualConsumption, string[,]
                                            predictedConsumption)
{
    double ampe = 0.000;

    for (int count = 1; count <= actualConsumption.GetUpperBound(1); count++)
    {
        ampe = ampe + Math.Abs((((Convert.ToDouble(actualConsumption[2, count]) –
                    Convert.ToDouble(predictedConsumption[1, count]))/
                    Convert.ToDouble(actualConsumption[2, count]))*100));
    }

    return ampe / Convert.ToDouble(actualConsumption.GetUpperBound(1));
}
```

A11. FIBONACCI NUMBER GENERATION

```
public int[] fn_Fibonacci_Numbers(double intervalOfUncertainty)

{
   double val1 = intervalOfUncertainty / 100;
   double val2 = 0.000;
   int fib_val = 0;
   int fib_val_1 = 1;
   int fib_val_2 = 1;
   int counter = 2;
   while (true)
   {
      fib_val = fib_val_1 + fib_val_2;
      val2 = 1.00 / Convert.ToDouble(fib_val);
      if (val2 > val1)
      {
         counter++;
         fib_val_2 = fib_val_1;
         fib_val_1 = fib_val;
      }
      else
         break;
  }
  int[] fibonacci_numbers = new int[counter];
  fibonacci_numbers[0] = 1;
  fibonacci_numbers[1] = 1;
  for (int count = 2; count < counter; count++)
  {
     fibonacci_numbers[count] = fibonacci_numbers[count - 1] +
                         fibonacci_numbers[count - 2];
  }

  return fibonacci_numbers;
}
```

A12. FIBONACCI SEARCH TECHNIQUE

```
// THIS FUNCTION COMPUTES THE OPTIMAL ASYMPTOTE VALUE

public double fn_Optimum_Asymptote(string[,] dataSet, double intervalOfUncertainty)
{
  // The array "dataSet" contains the readings and consumption data

  double asymptote = 0.000;
  double consumption_lower_limit_1 = 0.000; // Value of El
  double consumption_lower_limit_2 = 0.000; // Value of E'
  double consumption_upper_limit_1 = 0.000; // Value of Eu
  double consumption_upper_limit_2 = 0.000; // Value of E"
  double sum_of_squared_residuals_lower_limit = 0.000; // Value of SSR1
  double sum_of_squared_residuals_upper_limit = 0.000; // Value of SSR2
  double limit_difference = 0.000; //Value of L
  double fibonacci_distance = 0.000; // Value of d
  int[] fibonacci_numbers = fn_Fibonacci_Numbers(intervalOfUncertainty); // F(k)=F(k-1)
                                                                  + F(k-2)
  int while_loop_control=(fibonacci_numbers.Length - 1); // Value of k which will be
                             used in the computation of [F(k-2)]/[F(k)]*L
  string[] constants=fn_ConstantsComputation(dataSet); // Value of k1, k2

  // First row of array "comsumptions" contains the predicted data
  // Second row contains the actual data.
  string[,] consumptions = new string[2, dataSet.GetUpperBound(1)];

  for (int count_1 = 0; count_1 < dataSet.GetUpperBound(1); count_1++)
  {
    consumptions[0, count_1] = "0";
    consumptions[1, count_1] = dataSet[2, (count_1 + 1)];
  }

  // The maximum historical consumption is selected.
  asymptote = Convert.ToDouble(consumptions[1, 0]);

  for (int count_2 = 1; count_2 <= consumptions.GetUpperBound(1); count_2++)
  {
    if (Convert.ToDouble(consumptions[1, count_2]) > asymptote)
    {
      asymptote = Convert.ToDouble(consumptions[1, count_2]);
    }
  }
  consumption_lower_limit_1 = asymptote;
  consumption_upper_limit_1 = asymptote * 100;
```

```
while (while_loop_control > 1)
{
  limit_difference = consumption_upper_limit_1 - consumption_lower_limit_1; // L=Eu-
                                                            El
  double dist = (Convert.ToDouble(fibonacci_numbers[while_loop_control - 2]) /
            Convert.ToDouble(fibonacci_numbers[while_loop_control]));
  fibonacci_distance = (Convert.ToDouble(fibonacci_numbers[while_loop_control - 2]) /
                  Convert.ToDouble(fibonacci_numbers[while_loop_control])) *
                  limit_difference; // d=[F(k-2)]/[F(k)]*L
  consumption_lower_limit_2 = consumption_lower_limit_1 +
                    fibonacci_distance; // E'=El+d

  consumption_upper_limit_2 = consumption_upper_limit_1 –
                    fibonacci_distance; // E"=Eu-d


  // Computation of predicted values with consumption_lower_limit_2 as the optimal
    asyptotes

  for (int count_3 = 0; count_3 <= consumptions.GetUpperBound(1); count_3++)
  {
    consumptions[0, count_3] = fn_LogisticEquation(consumption_lower_limit_2,
                    Convert.ToDouble(constants[0]),
                    Convert.ToDouble(constants[1]), (count_3 +
                    1)).ToString("0.000");
  }

  // Computation of SSR1
  sum_of_squared_residuals_lower_limit = fn_Sum_Of_Squared_Residuals(consumptions);

  // Computation of predicted values with consumption_upper_limit_2 as the optimal
    asyptotes

  for (int count_3 = 0; count_3 <= consumptions.GetUpperBound(1); count_3++)
  {
    consumptions[0, count_3] = fn_LogisticEquation(consumption_upper_limit_2,
                    Convert.ToDouble(constants[0]),
                    Convert.ToDouble(constants[1]), (count_3 +
                    1)).ToString("0.000");
  }

  // Computation of SSR2
  sum_of_squared_residuals_upper_limit = fn_Sum_Of_Squared_Residuals(consumptions);

  if (sum_of_squared_residuals_lower_limit < sum_of_squared_residuals_upper_limit)
```

```
        {
          // SSR1 < SSR2
          consumption_upper_limit_1 = consumption_upper_limit_2; // Eu=E"
          //"consumption_lower_limit_1" remains the same since upper limit is discarded
            (E" - Eu is discarded)
        }
        else if (sum_of_squared_residuals_lower_limit >
                sum_of_squared_residuals_upper_limit)
        {
          consumption_lower_limit_1 = consumption_lower_limit_2;
          //"consumption_upper_limit_1" remains the same since lower limit is discarded
            (El - E' is discarded)
        }
        else
        {
          // SSR1 = SSR2 NOT LIKELY TO HAPPEN
        }

        while_loop_control--;
      }


      // The maximum asymptote is chosen here
      if (sum_of_squared_residuals_lower_limit < sum_of_squared_residuals_upper_limit)
      {
        // SSR1 < SSR2
        asymptote = consumption_lower_limit_2;
      }
      else
      {
        // SSR1 >= SSR2
        asymptote = consumption_upper_limit_2;
      }

      return asymptote;
    }
```