Journal of the Nigerian Association of Mathematical Physics
Volume 29, (March, 2015), **pp 231 – 244**

ARX and ARMAX Model Identification for Prediction of Power Consumption in Residential Buildings

Reginald O. A. Osakwe[1], Vincent A. Akpan[2†], Sylvester A. Ekong[3]

[1]Department of Physics, The Federal University of Petroleum Resources, P.M.B. 1221 Effurun, Delta State, Nigeria.
[2]Department of Physics Electronics, Federal University of Technology, P.M.B. 704 Akure, Ondo State, Nigeria.
[3]Department of Physics & Energy Studies, Achievers University, P.M.B. 1030, Owo, Ondo State, Nigeria.

Abstract

This paper presents the formulation of an ARX and ARMAX model predictors for model identification and prediction of power consumption in residential buildings. A total of 18,000 data samples from the Power Holding Company of Nigeria were collected for different residential buildings in Akure, Ondo State – Nigeria, namely: 1). single flat housing unit (SFHU); 2). multiple flat housing unit (MFHU); 3). single flat housing unit based on climatic conditions (SFHUC); 4). multiple flat housing unit based on climatic conditions (MFHUC); 5). single flat housing unit based on square-feet (SFHUSF); and 6). multiple flat housing unit based on square-feet (MFHUSF). The performances of the two model predictors are validated by one-step and five-step ahead prediction methods. The results obtained from the application of these two model structures and predictors for the modeling and prediction of power consumption in residential building as well as the validation results show that the ARMAX model outperforms the ARX model with much smaller predictions error and good prediction capabilities with appreciable degree of accuracy and that the proposed ARMAX model structure can be used for power consumption predictions in real scenarios.

Keywords: ARX, ARMAX, mathematical modelling, model predictor, model structure, power consumption.

1.0 Introduction

The key reason for building energy modeling and simulation is to understand the building condition and check it against utility bills with a view to reducing the energy consumption of the buildings. Computer based modeling and simulation is a proven technique for evaluating building energy consumption [1]. Due to growing environmental concerns as well as economic and political requirements, the envisioned future power grid will increasingly rely on renewable sources such as wind and solar energies. For example, President Obama's "New Energy for America" calls for renewable energy to supply 10% of the nation's electricity by 2012, rising to 25% by 2025 [2].

This era of fossil fuel dependency and concerns about greenhouse gas emissions has increased interest in the use of policy and technology solutions to reduce and shift energy use. The residential sector accounted for about 22% of total primary energy consumption in the U.S. in 2009, indicating that there are major potential gains from implementing such solutions in residential settings [3]. The potential energy, cost, and emissions savings of such policies and technologies can be investigated by modeling their impacts on residential energy demand and the resulting interactions between this demand and the power grid, renewable generation, energy storage, and plug-in electric vehicles.

Energy consumption of buildings (both residential and commercial) has steadily increased, reaching figures between 20% and 40% in developed countries [4]. For example, building energy consumption accounts for one-third of the societal energy consumption in China, and has the largest energy-saving potential [5, 6]. In recent years, with the rapid development of

Corresponding author: Reginald O. A. Osakwe, E-mail: avincent@eng.auth.gr, Tel.: +2347031351144, 08034968173(S.A.E)

urbanization, economic development, people's incomes and living standards, China's building energy consumption has increased dramatically.

In the UK and many other industrialized countries, offices, as a basic unit for work buildings, are intensively distributed in big cities and urban areas. As climate change becomes a very important global issue, the UK government has, for example, set a target of cutting CO_2 emission by 34% of 1990 levels by 2020 [7, 8]. In the UK, the energy consumed in the service sector took up 14% of overall energy consumption of the whole country in 2001. Most of the energy for the service sector is used in various kinds of offices for heating, lighting, computing, catering and hot water. Thus, energy consumption in office buildings is one of the research areas which have significant importance for meeting the UK government’s 2020 target [7, 8]. Similarly, commercial buildings consume approximately 19% of all energy and account for 18% of all CO_2 emissions in the U.S. [9]. By 2035, commercial building floor space is expected to increase by 28% in the U.S. compared to the total floor space area in 2009, reaching 103 billion sq. ft. [9]. This makes commercial buildings a significant target for achieving sustainability.

Different models have different data input requirements, so building energy consumption models may differ in their calculation and predictions. In general, according to the principle of energy consumption modeling technology, energy consumption models can be mainly divided into macro models and micro models [5]. However, in [6], only the macro model approach is discussed.

Two general classes of techniques are available to model residential power demand: top-down and bottom-up models [10]. Top-down models use estimates of total residential sector energy consumption, together with other pertinent macro variables, to attribute energy consumption to characteristics of the housing sector. This class of models can be compared to econometric models, which require little detail of the actual consumption process. These models treat the residential sector as an energy sink and regress or apply factors that affect consumption to determine trends [10, 11, 12, 13]. Depending on availability, the input data required to develop these models can include the structural characteristics of the dwellings, occupants and their behavior, appliances’ characteristics, historical energy consumption, weather conditions, and macroeconomic indicators. Stochastic predictors, based on time-series approach, such as auto regressive moving average methods, are also used to forecast home energy consumption [14, 15, 16].

Bottom-up models, on the other hand, identify the contribution of each end-use towards the aggregate energy consumption of the residential sector [17, 18, 19, 20]. Bottom-up approaches refine the modeling of energy consumption, allowing the simulation of the effects of technology improvements and policy decisions. These models calculate the energy consumption of an individual or group of households and extrapolate the results to a region or nation. This aggregate result is generally accomplished by using a weight for each modeled house or group of houses based on its representation of the sector [9]. Moreover, the bottom-up approach has the capability of determining total energy consumption of the residential sector without relying on historical data. Common input data to bottom-up models include dwelling characteristics (*e.g*., size and layout, building materials, and appliances’ characteristics), weather conditions, household occupant behavior and related use of appliances, lighting use, and characteristics of heating, ventilation, and air conditioning (HVAC) systems. This high level of detail represents the strength of bottom-up models, providing the ability to model the impact of different technology options and allowing the implementation of energy optimization techniques. On the other hand, the use of such detailed information, in particular regarding household members’ behavior, introduces great model complexity. The input data requirements are typically greater than that of top-down models.

2.0 Formulation of the Research Problem

The main problems attempted to be solved in this work can be formulated as follows. Due to the nature of the metering system in Nigeria as provided by the Power Holding Company of Nigeria (PHCN), only power consumption from the newly introduced pre-paid meters are available. The pre-paid metering system does not provide power consumed by individual items in homes, offices, industries nor industries. For this reason, the data obtained and used for this research work is classified as power consumed by home appliances including lighting.

In view of the above, the problems considered here is in terms of the power consumption by residential buildings which is further sub-divided into three categories according to the following six criteria: *1).* The type of housing units: *(i)* single-family homes (single flat housing unit, SFHU) and *(ii)* multi-family homes (multiple family housing units, MFHU); *2).* The type of climate during electricity consumption: *(i)* single-family homes (single flat housing units by climate, SFHUC) and *(ii)* multi-family homes (multiple family housing units by climate, MFHUC); *3).* The square footage of the housing units: *(i)* single-family homes (single-flat housing units by square footage, SFHUSF), and *(ii)* multi-family homes (multiple family housing units by square footage, MFHUSF). Please note that multi-family homes are apartments with more than 2 blocks of flat units. The climate referred to here is the data between November and January usually called the harmattan period in Nigeria.

The main objective of this work is on mathematical modeling for prediction of power consumption in residential buildings in Akure using efficient modeling algorithms. This is achieved through: *1).* Data collection and data pre-processing to remove outliers; *2).* The development of an ARX and ARMAX model predictors for the modeling and prediction of electricity consumption in residential buildings; and *3).* The validation of the efficiency of the ARX and ARMAX model predictors for the prediction of future power consumption in buildings.

3.0 Experimental Data Acquisition

The 3,000 data each for the six criteria making 18,000 data used in this work was collected from the distribution unit of the Power Holding Company of Nigeria, Akure which is now Benin Electricity Distribution Company (BEDC), Akure. However, various ideas have been proposed for modelling and prediction of power consumption but there are equally various challenges in their accuracy which is due to the complexity in the rate at which different building consumes power at a particular point in time throughout the year. Certain physical attributes such as climate, the number of occupants in the building, the square footage of the building etc plays a vital role in the rate at which power is consumed in residential building thus playing vital role in the prediction process. As mentioned earlier, the aim of this paper is on the model identification for the prediction of electricity consumption in residential building ARX and ARMAX model predictors.

4.0 Problem Formulation for the Power Consumption in Residential Buildings

Series of work have been carried out in order to model and predict to a high level of accuracy the power consumption in residential building but the required accuracy has been influenced by certain environmental and human factors. To carry out this work, the rate of power consumption in KWhr per number of houses needs to be estimated.

Below are the graphs plotted to show the rate of power consumption in residential building against the number of houses sampled. Fig. 1 shows the graph of Power Consumption in KWhr for a single flat housing unit and multiple flat housing units against the number of sampled houses. Also, Fig. 2 and Fig. 3 shows that of single flat housing unit by climate and single flat housing unit by square-feet and their corresponding multiple flat housing units respectively.

5.0 Power Consumption by Housing Units

Fig. 1(a) and Fig. 1(b) show the graph of power consumption by housing units taken over sampled houses. Fig. 1(a) represents single family flats by housing units while Fig. 1(b) represents multiple family flats by housing units in KWhr.

When Fig. 1(a) is compared with Fig. 1(b), there is a considerable difference in the power consumption of single family flats to that of a multiple family flats. In single family flats, the power consumption is above 100KWhr whereas in the multiple family flats, the power consumption based on the sampled houses is less than 80KWhr.

Based on the data collected and the graphs plotted, it can be deduce that more power is consumed in the single family flats compared to a multiple family flats. Various factors could be attributed to this such as:-

(i) Number of occupants in the building: This is one of the factors that determine the rate at which power is consumed in buildings. If there are many occupants in a building, there will be reasons to consume more power. For instance there will be more use of air conditioner during hot weather condition while during the cold weather, there will be much more use of water heater.

In relation to Fig. 1(a) and Fig. 1(b), the rate of power in single family flat which is considerably high compared to the multiple family flats could be as a result of more occupants in the single family flats under study.

(ii) The nature of their work : Depending on the nature of the occupants of a building, some people work from home, while some return home early after work whereas some spent most part of their of their day at work. This attributes also contribute to the rate of power consumption in residential buildings.People that spent most part of their day at work are likely to consume less power than those that work from home or return home early from work. This probably have an effect on why single family flats consumed more power than multiple family flats.

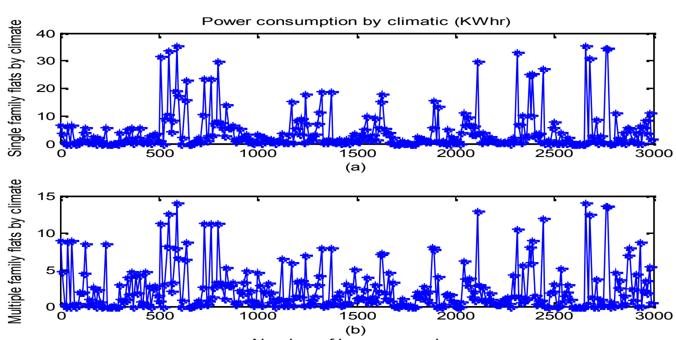

Fig. 1: Graph of power consumption in kWHr for: (a) single flat housing units (SFHU) and (b) multiple flat housing units MFHU).

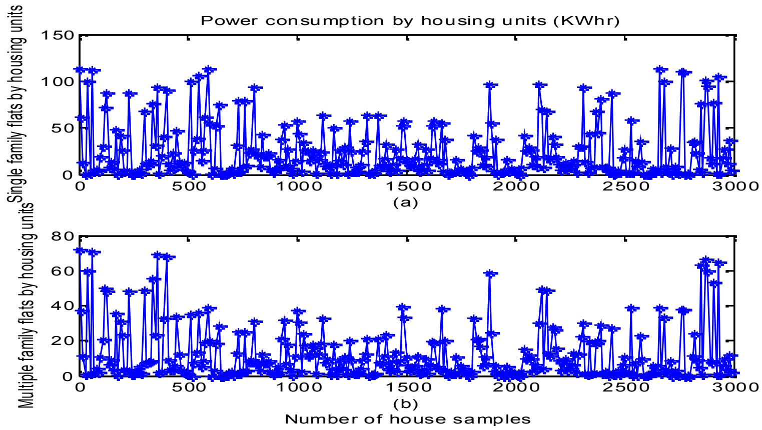

Fig. 2: Graph of power consumption in KWhr for: (a) single flat housing units by climate (SFHUC) and (b) multiple flat housing units by climate (MFHUC).

(iii) Their awareness about power consumption: This is another factor that influences how power is consumed in residential buildings. Some people, especially in Nigeria leaves their home without switching off their bulbs, electronic appliances like air conditioner, electric fan etc. This leaves these appliances to consume unnecessary power. This invariably determines the rate of power consumption in buildings.

(iv) Types of Electrical Appliances: The consuming rate of the electrical appliances is another factor that influence the rate of power consumption in residential building. Some homes have a lot of appliances that consume much energy whereas some homes have less electrical appliances.

6.0 Power Consumption by Climate

Climate change is expected to lead to changes in a range of climatic variables such as in the temperature levels, and as electricity demand is closely influenced by these, there is likely to be an impact on power demand patterns. The magnitude of the potential future climate changes on electricity will depend on the power use patterns in the absence of climate change as well as long term socio-economic trends.

The rate of power consumption during rainy or harmattan season is relatively low when compared to other period of the year. Fig. 2(a) and Fig. 2(b) show the power consumption of single family flats and multiple family flats respectively during a climatic weather conditions. The maximum value of power consumption during harmattan for single family flats is less than 40KW while that of multiple family flats is less than 15KW. When Fig. 2(a) and Fig. 2(b) are compared to the Fig. 1(a) and Fig. 1(b), it shows that less power is consumed in residential building during climatic condition than other time of the year. This explains why power is more stable during rainy or harmattan season. Various reasons could be attributed to this, such as:

(i) Reduction in the use of air conditioner: During cold weather, people tend to avoid the use of air conditioner simply because the weather is cold enough.

(ii) Change in living habits of occupants: During this period, there is a considerable change in the habits of people in residential areas. They tend to change their way of life to suit the environmental condition at this particular time.

(iii) Reduction in the rate of consumption of cold drinks etc.

7.0 Power Consumption By Sizes in Square Footage

Fig. 3(a) shows the graph of single family flats by sizes against the number of houses samples while Fig. 3(b) shows the graph of multiple family flats by sizes against the number of house samples.

It can be deduce from the graph that the power consumption of single family flat is much more than the power consumption of a multiple family flats. According to Fig. 3(a), the power consumption is about 20KW over the period when the data was taken whereas that of Fig. 3(b) shows that the power consumption is about 7KW under the same condition.

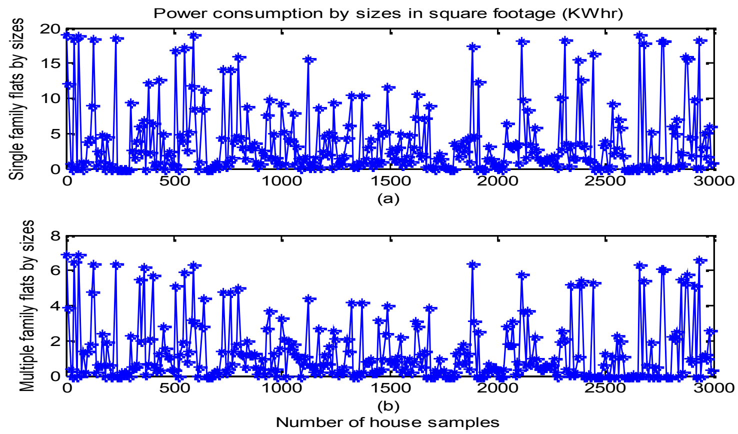

Fig. 3: Graph of power consumption in KWhr for: (a) single flat housing units by square-feet (SFHUSF) and (b) multiple flat housing unit by square-feet (MFHUSF).

Building designs and construction techniques can maximize the use of natural light and ventilation which minimizes the need for artificial light and HVAC equipments. Using building shading techniques, installing window to minimize solar intake, (depending on region) and properly insulating against unwanted airflow between indoor and outdoor spaces improve energy use.

8.0 Mathematical Models of Dynamic Systems and Their Model Predictors

The method of representing the behaviour of dynamical systems by vector difference or differential mathematical relationships is well established in system and control theories [21, 22, 23, 24, 25, 26]. These relationships constitute the so-called mathematical model of the system.

One very common method of modeling the behaviour of a *p*-input *q*-output multivariable plant in the discrete time space is by the family of the following general mathematical relationship [21, 24]:

$$A(z^{-1})Y(k) = z^{-d}\frac{B(z^{-1})}{F(z^{-1})}U(k) + \frac{C(z^{-1})}{D(z^{-1})}e(k) \tag{1}$$

where $Y(k)$ is the vector of order n of the *q* outputs at the timing instant k responding to the vector input $U(k)$; $e(k)$ is the noise disturbance vector; and $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$, $D(z^{-1})$ and $F(z^{-1})$ are polynomial matrices given by

$$\begin{aligned} A(z^{-1}) &= I + A_1 z^{-1} + \cdots + A_{n_a} z^{-n_a} \\ B(z^{-1}) &= B_0 + B_1 z^{-1} + \cdots + B_{n_b} z^{-n_b} \\ C(z^{-1}) &= I + C_1 z^{-1} + \cdots + C_{n_c} z^{-n_c} \\ D(z^{-1}) &= I + D_1 z^{-1} + \cdots + D_{n_d} z^{-n_d} \\ F(z^{-1}) &= I + F_1 z^{-1} + \cdots + F_{n_f} z^{-n_f} \end{aligned} \tag{2}$$

d is the system delay, *A*, *C*, *D* and *F* are *monic* polynomial matrices in the backward shift operator z^{-1} Their dimensions are $n_a \times n_a$ $q \times n_a$, $q \times n_d$, $q \times n_d$ and $q \times n_f$ and their degree n, m, c, l, respectively; B is a $n_b \times p$ stable polynomial matrix (i.e. all its zeros are all inside the unit circle) of degrees of degree r. The term *monic* implies that the leading coefficients of *A*, *C*, *D* and *F* are identity matrices of appropriate dimension to avoid division by zeros and also because the magnitude of $e(k)$ can be adjusted to compensate for this if necessary. In this discussion, it is assumed that: *1)* the time delay d of the system is known, i.e. $d = 1$; *2)* the coefficients of the polynomials matrices $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$, $D(z^{-1})$ and $F(z^{-1})$ are unknown; *3)* the polynomials matrices $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$, $D(z^{-1})$ and $F(z^{-1})$ are relatively prime; and *4)* that the upper bound on the order or each polynomial matrix is known or can be specified exactly.

Since the noise term $e(k)$ enters the general model equation (1) as a direct error term, the model of (1) is often called an equation error model [21, 24]. Depending on how the five parameters A, B, C, D and F are combined, several model structures can be obtained from (1)

The choice of the models that will represent the noise disturbances is as important as the choice of the system model.

Depending on the different assumptions made about the spectral density of the noise, $e(k)$ and how the noise is assumed to enter the system given by (1); 32 different model structures can be derived from (1) based on the combination of the five parameters *A*, *B*, *C*, *D* and *F* [24]. However, the model structures considered in the present work is limited to the structures derived from the combination of the four parameters *A*, *B*, *C* and *F*, that is ignoring the *D* parameter in (1). The reason for choosing these four parameters is because, as literature shows, they were adequate for the modeling needs of the model predictive control (MPC) for a wide range of dynamical systems [21, 27]. The combination of *A* and *B* results in an Auto Regressive with eXogenous inputs (ARX) model, the combination of *A*, *B* and *C* gives an Auto Regressive Moving Average with eXogenous inputs (ARMAX) model, and the combination of *B* and *F* corresponds to the output error (OE) model. The output error (OE) model is a form of equation error model [21, 23, 24] and can also take the form based on *A*, *B*, *C* and *D* which is widely used in MPC literature [21, 27, 28]. Rather than using *A*, *B*, *C* and *D* to describe the output error (OE) model, the choice of using *B* and *F* is adopted in this work for the output error (OE) model [21, 24].

Let $\theta(k)$ be a parameter vector which encapsulates the model parameters given in (2) and defined as:

$$\theta(k)=\left[-A_1,\ldots,-A_{n_a},\ B_1,\ldots,B_{n_b},\ C_1,\ldots,C_{n_c},\ -D_1,\ldots,-D_{n_d},\ -F_1,\ldots,-F_{n_f},\ \right]^T \tag{3}$$

Since the exact value of the parameter vector $\theta(k)$ in (3) is unknown, a parameterized set of model structures Θ can be defined as a set of candidate models given as:

$$\Theta:\ \theta(k)\in \Omega\subset\Re^{\eta}\rightarrow\hat{\theta}(k) \tag{4}$$

where Ω is some subset of $\Re^{\eta}$ inside which the search for a model is carried out; η is the dimension of $\theta(k)$; $\hat{\theta}(k)$ is the desired model associated with the parameter vector $\theta(k)$ and contained in the set of models $\Theta=\{\theta(k)_1,\theta_2(k),\ldots,\theta_\tau(k)\}$; $\theta(k)_1,\theta_2(k),\ldots,\theta_\tau(k)$ } Each member of this set is a distinct value of $\theta(k)$; and $\tau=1,2,\ldots,\max_{iter}$ is the number of iterations required to determine the $\hat{\theta}(k)$ from Θ.

Thus, the minimum variance (one-step) ahead predictor of (1) at time k based on the system information up to the time $k-1$ can be expressed as

$$\hat{Y}(k\,|\,k-1,\theta(k-1))=z^{-d}\,\frac{B(z^{-1})D(z^{-1})}{F(z^{-1})C(z^{-1})}U(k)+\left[1-A(z^{-1})\frac{D(z^{-1})}{C(z^{-1})}\right]Y(k) \tag{5}$$

Note the inclusion of $\theta(k)$ as an argument to indicate that the model structure represents a set of models. For notational convenience, the $k-1$ will be omitted henceforth. The prediction error $\varepsilon(k,\theta)$ can be computed directly from (1) and (5) as follows:

$$\begin{aligned}\varepsilon(k,\theta(k))&=Y(k)-\hat{Y}(k,\theta(k))\\&=\frac{D(z^{-1})}{C(z^{-1})}\left[A(z^{-1})Y(k)-z^{-d}\,\frac{B(z^{-1})}{F(z^{-1})}U(k)\right]\end{aligned} \tag{6}$$

By introducing $\tilde{d}(k,\theta(k))=z^{-d}\,\dfrac{B(z^{-1})}{F(z^{-1})}U(k)$ (7)

and $\tilde{v}(k,\theta(k))=A(z^{-1})Y(k)-\tilde{d}(k,\theta(k))$ (8)

and using (7) and (8), equation (6) can be expressed as

$$\varepsilon(k,\theta(k))=Y(k)-\hat{Y}(k,\theta(k))=\frac{D(z^{-1})}{C(z^{-1})}\tilde{v}(k,\theta(k)) \tag{9}$$

Let the regression vector (the so-called state vector) derived from the difference equation form of (1) be:

$$\begin{aligned}\varphi(k,\theta(k))=\big[&Y(k-1),\ldots,Y(k-n_a),\ \ U(k-d),\ldots,U(k-d-n_b),\ \ \varepsilon(k-1,\theta(k)),\ldots,\varepsilon(k-n_c,\theta(k))\\&\tilde{d}(k-1,\theta(k)),\ldots,\tilde{d}(k-n_f,\theta(k)),\ \ \tilde{v}(k-n_d,\theta(k)),\ldots,\tilde{v}(k-n_{d,\theta(k)})\big]\end{aligned} \tag{10}$$

Using the parameter vector given in (3) and the regression vector in (10) above, equations (7) and (9) can be expressed respectively as:

$$\tilde{d}(k,\theta(k))=B_1U(k-d)+\cdots+B_{n_b}U(k-d-n_b)+F_1\tilde{d}(k-1,\theta(k))+\cdots+F_{n_f}\tilde{d}(k-n_f,\theta(k)) \tag{11}$$

$$\varepsilon(k,\theta(k))=C_1\varepsilon(k-1,\theta(k))+\cdots+C_{n_c}\varepsilon(k-n_c,\theta(k))+\tilde{v}(k,\theta)+D_1\tilde{v}(k-d)+\cdots+D_{n_d}\tilde{v}(k-d-n_d) \tag{12}$$

Inserting $\tilde{v}(k,¯)$ from (12) and substituting $\tilde{d}(k,¯)$ from (11) into (6) gives

$$£(k,¯(k)) = Y(k) - ¨(k,¯(k))¯(k) \tag{13}$$

Thus, the one-step ahead predictor can then be expressed as:

$$\hat{Y}(k,¯(k)) = ¨(k,¯(k))¯(k) \tag{14}$$

9.0 Remarks on the Disturbance Model

The disturbance model, i.e. the second term in (1), plays significant role in modeling the overall system behaviour. Let the disturbance model be defined as

$$D_M = \frac{C(z^{-1})}{D(z^{-1})} e(k) \tag{15}$$

In MPC literature, the model of (15) is usually called CARIMA (controlled auto-regressive and integrated moving average) model [21, 27, 28]. In practice, $e(k)$ cannot be measured but it can be estimated as deterministic or stochastic noise [21, 23, 24, 25].

10.0 Autoregressive with Exogenous Input (ARX) Model

The deterministic case is simply achieved by setting $C(z^{-1}) = D(z^{-1}) = 1$ with the assumption that $e(k)$ is a zero-mean white noise with finite variance while its first few terms are made non-zero. Additional assumption on $e(k)$ is that it is independent of past inputs and that it can be characterized by some probability function [21, 24]. With these assumptions on (15) and setting $F(z^{-1}) = 1$ in (1), equation (1) essentially reduces to an autoregressive with exogenous input (ARX) model structure, which is stable for wide range of operations. From (5), the ARX model predictor becomes:

$$\hat{Y}(k \mid k-1,¯(k-1)) = z^{-d} B(z^{-1}) U(k) + \left[1 - A(z^{-1})\right] Y(k) \tag{16}$$

Note that the ARX model predictor uses the system outputs $Y(k)$ and the model parameters $¯(k)$ to predict future outputs $\hat{Y}(k)$.

11.0 Autoregressive Moving Average with Exogenous Input (ARMAX) Model

The stochastic case is somewhat more involved. Consider the case of modeling a stationary, zero-mean white noise process, namely $E\{e(k)^2\} = ±^2$, $E\{e(k)e(k-ª)\} = 0$ for all $ª \neq 0$, the probability distribution of $e(k)$ being the same for all (k), and each $e(k)$ being independent of $e(ª)$ if $ª \neq k$; where the term $E\{\ \}$ implies the expectation or mean value of its arguments. Then, if $C(z^{-1}) / D(z^{-1})$ is an asymptotically stable transfer function, the (15) will be a stationary process with spectral density given by

$$\Phi_{SD}(\mu) = ±^2 \frac{\left|C(e^{-j\mu T})\right|^2}{\left|D(e^{-j\mu T})\right|^2} \tag{17}$$

where $±$ is the spectral density. Note that since $\left|C(e^{-j\mu T})\right|^2 = C(e^{-j\mu T}) \cdot C(e^{-j\mu T})$, it is always possible to choose $C(z^{-1})$ such that all its roots lie inside the unit disc, i.e. without restricting the spectral densities which can be modeled in this way. Also for the same reason, the factors of $C(z^{-1})$ do not affect the spectral density. This property shows and guarantees a useful way of selecting $C(z^{-1})$ to lie inside the unit circle for models with moving average such as ARMAX and output error (OE) models.

As in the previous sub-section 4.1, with the assumptions on (15) and setting $D(z^{-1}) = F(z^{-1}) = 1$ in (1), equation (1) essentially reduces to an autoregressive moving average with exogenous input (ARMAX) model structure, which is usually unstable but find applications in wide range of systems with coloured noise. From (5), the ARMAX model predictor becomes:

$$\hat{Y}(k \mid k-1,¯(k-1)) = z^{-d} \frac{B(z^{-1})}{C(z^{-1})} U(k) + \left[1 - A(z^{-1}) \frac{1}{C(z^{-1})}\right] Y(k) \tag{18}$$

Note that the moving average filter $C(z^{-1})$ must be estimated at each time step and must equally lie within the left-hand plane of the unit circle for stability [21].

12.0 The ARX and ARMAX Model Validation Algorithms

Network validations are performed to assess to what extend the trained model predictors have approximated and captured the behaviour of the underlying dynamics of a system and as measure of how well the model being investigated will perform when deployed for the actual system modeling and future predictions [22, 24, 29].
The first test involves the comparison of one-step output predictions of the true training data using the model predictors given in (16) and (18) and the evaluation of their respective corresponding prediction errors using (13).
The second test involves the comparison of one-step output predictions of the true validation data that was not used during the model predictor development and the evaluation of their respective corresponding prediction errors using (13).
The third method is the *K*-step ahead predictions [21, 24] where the outputs of the trained network are compared to the unscaled output training data. The K-step ahead predictor follows directly from (8) and for $\ddot{}(k) = \hat{\ddot{}}(k+K)$ and $\bar{}(k) = \hat{\bar{}}(k)$, takes the following form:

$$\hat{Y}((k+K)\,|\,k, \hat{\bar{}}) = \hat{J}(Z^N, \hat{\ddot{}}(k+K), \hat{\bar{}}(k)) \tag{19}$$

where $\hat{\ddot{}}(k+K) = [U((k+K-1)\,|\,\hat{\bar{}}), \ldots, U((k+K-m)\,|\,\hat{\bar{}}),$

$\hat{Y}((k+K-1)\,|\,\hat{\bar{}}), \ldots, \hat{Y}((k+K+1-\min(k,n))\,|\,\hat{\bar{}}),\ Y((k+K-1)\,|\,\hat{\bar{}}), \ldots, Y((k+K-\max(n-k,0)\,|\,\hat{\bar{}})]^T$

The mean value of the K-step ahead prediction error (MVPE) between the predicted output and the actual training data set is computed as follows:

$$MVPE = mean\left(\sum_{k=m+K}^{N} \frac{Y(k) - \hat{Y}((k+K)\,|\,k, \hat{\bar{}})}{Y(k)}\right) \times 100\% \tag{20}$$

where $Y(k)$ corresponds to the actual output training data and $\hat{Y}((k+K)\,|\,k, \hat{\bar{}})$ the *K*-step ahead predictor output.

13.0 Simulation Studies and Discussion of Results

The ARX and ARMAX mathematical model with their respective predictors developed in Section 4 as well as the model validation algorithms proposed in Section 5 are applied for the modeling and prediction of power consumption in residential building based on the available data as discussed in Section 3.

14.0 Estimating the ARX and ARMAX Models

The input vector to the ARX and the ARMAX model predictors are the past values of the inputs (n_b) and outputs (n_a) as well as the order of moving average filter (n_c) which constitute the regression vector each defined from (10) for ARX and ARMAX as $\ddot{}_{ARX} = [\mathrm{n}_a, \mathrm{n}_b]$ and $\ddot{}_{ARMAX} = [\mathrm{n}_a, \mathrm{n}_b, \mathrm{n}_c]$ respectively. The inputs are the past values contained in the regression vector while the outputs are the predicted values of $\hat{Y}(k)$ given by (16) and (18) for ARX and ARMAX model predictors respectively while the optimal value of the adjustable parameters of the models $\bar{}(k)$ defined in (3) becomes $\hat{\bar{}}(k)$.

For assessing the model prediction performances, the model predictors was trained for ^{2} = 500 epochs (number of iterations) with the following selected parameters: $p = 6$, $q = 6$, $n_a = 4$, $n_b = 4$, $n_c = 4$, $n_{\ddot{}_{ARX}} = 48$ (ARX), $n_{\ddot{}_{ARMAX}} = 72$ (ARMAX). The details of these parameters are discussed in Section 4; where p and q are the number of inputs and outputs of the system, n_a, n_b and n_c are the orders of the regressors in terms of the past values, $n_{\ddot{}}$ is the total number of regressors (that is, the total number of inputs to the model predictor).

The 3,000 data each collected for the six criteria considered for the present case study is divided into two parts: 2400 (80%) to form the training data used for estimating the two models while the remaining 600 (20%) is reserved for the two estimated model validation.

15.0 Validation of the Estimated ARX and ARMAX Model

According to the discussion on model validation algorithms in Section 5, an estimated model can be used to model a process once it is validated and accepted, that is, the model demonstrates its ability to predict correctly both the data that were used for its development and other data that were not used during the model development.

The results shown in Fig. 4, Fig. 5 and Fig. 6 each having (a) to (l) corresponds to the one-step ahead output predictions of the training data, one-step ahead output predictions of the validation data and five-step ahead output predictions or the training data respectively for the six criteria considered, namely: SFHU, MFHU, SFHUC, MFHUC, SFHUSF and MFHUSF. Their respective corresponding prediction errors are given in Table 1 for the one-step ahead training, one-step ahead validation and 5-step ahead prediction errors.

Table 1: One-step ahead training, one-step ahead validation and 5-step ahead prediction errors.

	Training Errors		Validation Errors		5-Step Errors	
	ARX	ARMAX	ARX	ARMAX	ARX	ARMAX
SFHU	NaN	23.0593	NaN	25.7131	NaN	23.7404
MFHU	8.3377e+193	2.9089	5.8161e+47	3.3121	12.4594	11.7963
SFHUC	3.0735e+85	0.1833	2.0903e+22	0.2688	74.6538	4.3244
MFHUC	NaN	0.0422	NaN	2.2336	NaN	2.3666
SFHUSF	NaN	5.3131	NaN	13.7211	NaN	3.8967
MFHUSF	NaN	0.1104	NaN	0.1863	NaN	1.3115

16.0 Validation by the One-Step Ahead Predictions Simulations

In the one-step ahead output prediction method, the errors obtained from one-step ahead output predictions of the estimated model are assessed. In Fig. 4(a)–(l) the graphs for the one-step ahead predictions of the scaled training data (blue -) against the trained network output predictions (red --*) using the estimated ARX and ARMAX models respectively are shown for 500 epochs.

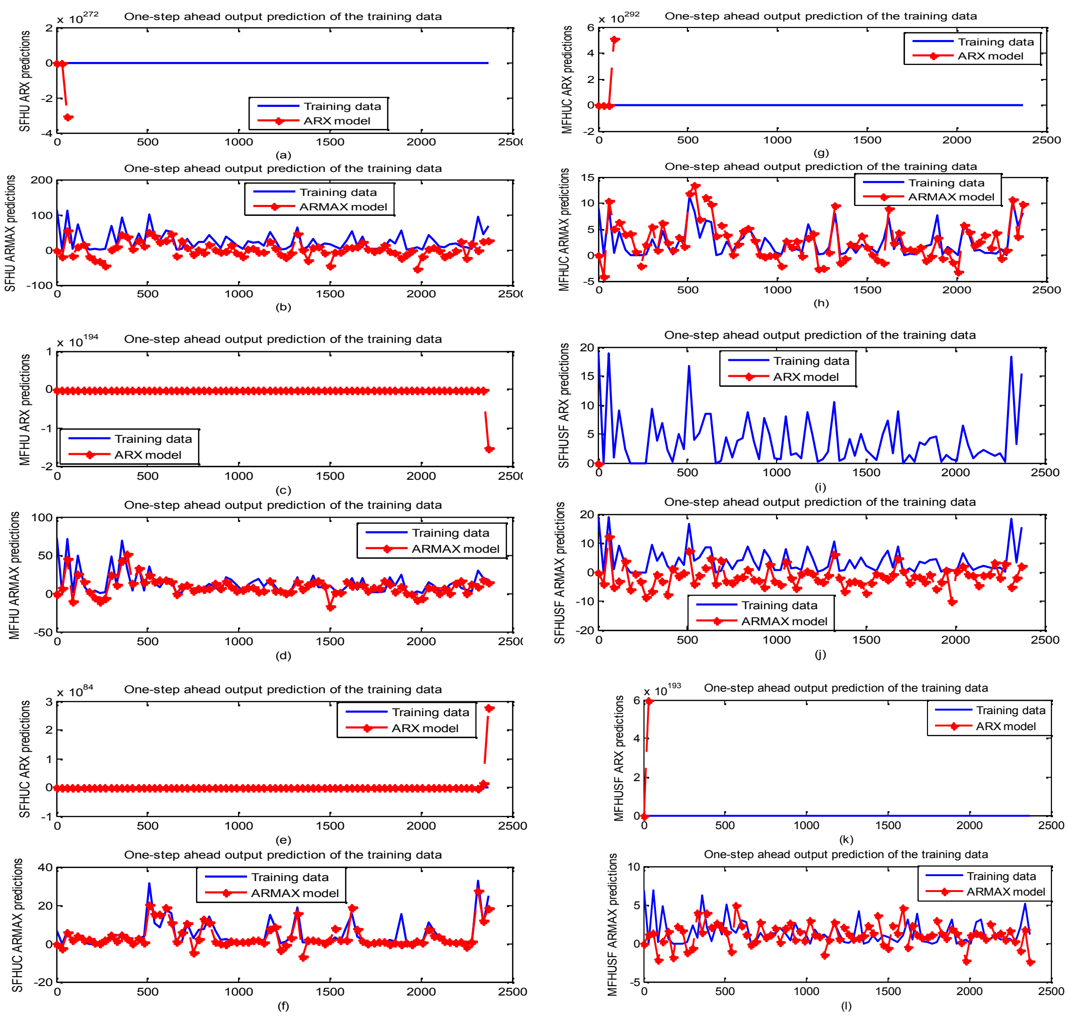

Fig. 4: Comparison of the one-step ahead output predictions of the original training data: (a) SFHU by ARX, (b) SFHU by ARMAX, (c) MFHU by ARX, (d) MFHU by ARMAX, (e) SFHUC by ARX, (f) SFHUC by ARMAX, (g) MFHUC by ARX, (h) MFHUC by ARMAX, (i) SFHUSF by ARX , (j) SFHUSF by ARMAX, (k) MFHUSF by ARX, (l) MFHUSF by ARMAX.

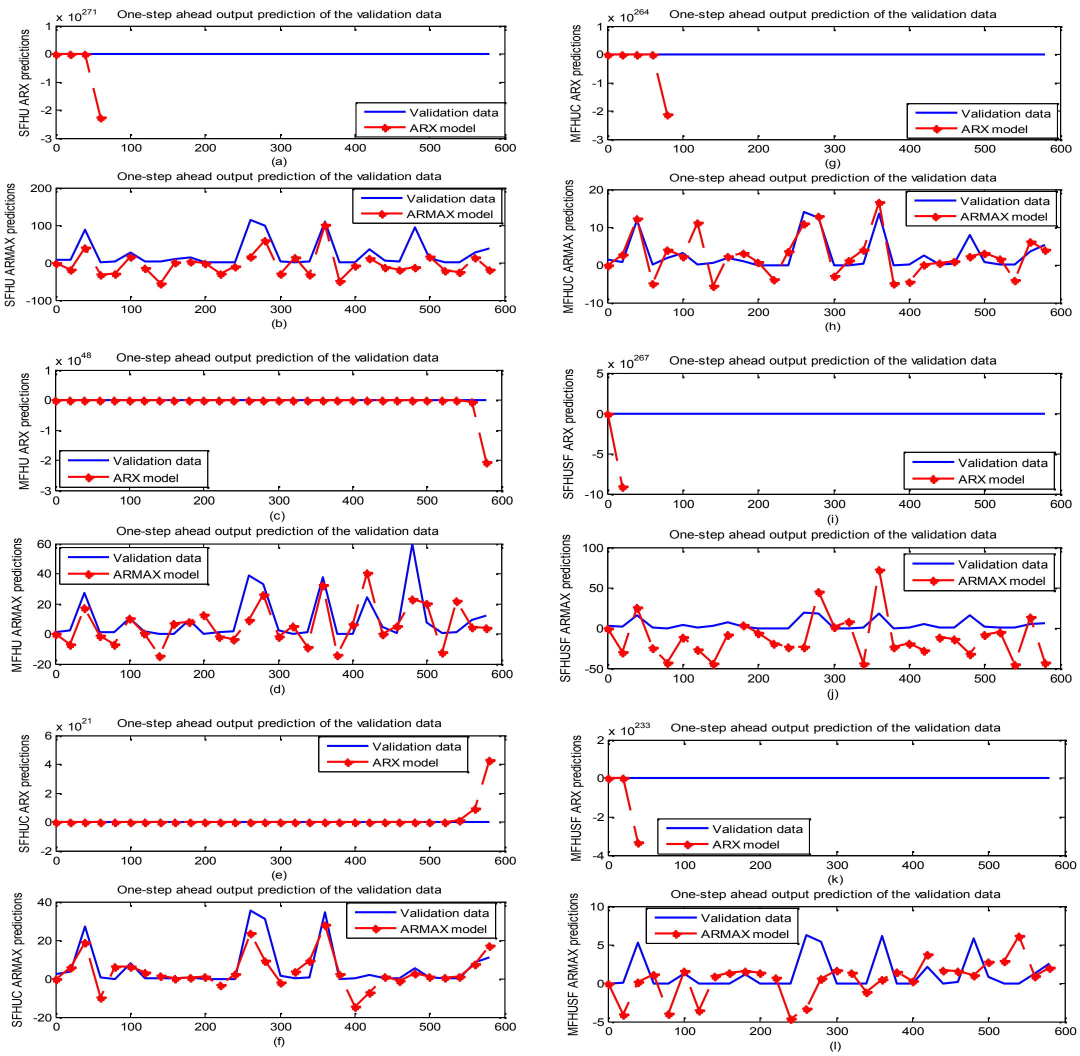

Fig. 5: Comparison of the one-step ahead output predictions of the validation data: (a) SFHU by ARX, (b) SFHU by ARMAX, (c) MFHU by ARX, (d) MFHU by ARMAX, (e) SFHUC by ARX, (f) SFHUC by ARMAX, (g) MFHUC by ARX, (h) MFHUC by ARMAX, (i) SFHUSF by ARX , (j) SFHUSF by ARMAX, (k) MFHUSF by ARX, (l) MFHUSF by ARMAX.

The mean value of the one-step ahead prediction errors are given in the second column of Table 1 for ARX and ARMAX respectively. It can be seen in these figures that the output predictions of the training data based on the ARMAX model closely match the original training data while the predictions based on the ARX model is completely out of phase with the true training data with much larger errors when compared to the errors obtained based on the ARMAX model as shown in the second column of Table 1. These small one-step ahead prediction errors are indications that the ARMAX model captures and approximate the dynamics inherent in the data to an appreciable degree of accuracy and that the ARMAX model predictor can be used for the future power consumption predictions.

Furthermore, the suitability of the ARX and the proposed ARMAX model predictors for the model identification and prediction for use in real power consumption predictions is investigated by validating the estimated models with the 600 validation data that was not used during model development. Graphs of the estimated ARX and ARMAX model one-step ahead output predictions (red --*) of the validation data with the actual validation data (blue -) using the ARX and the ARMAX models are shown in Fig. 5(a)–(l) for the six criteria considered. Again, as one can observe in these figures that the output predictions of the validation data based on the ARMAX model closely match the validation data while the predictions based on the ARX model is completely out of phase with the true validation data with much larger errors when compared to the errors obtained based on the ARMAX model as shown in the third column of Table 1. These small one-step ahead prediction errors of the validation data are indications that the ARMAX model can again be used for power consumption predictions in real scenarios.

17.0 *K*–Step Ahead Prediction Simulations

The results of the *K*-step ahead output predictions (red --*) using the *K*-step ahead prediction validation method discussed in Section 5 for 5-step ahead output predictions (K = 5) compared with the original training data (blue -) are shown in Fig. 6(a) to Fig. 6(l) for the estimated ARX and ARMAX models. Again, the value K = 5 is chosen since it is a typical value used in most model predictive control (MPC) applications to investigate the capabilities of trained model for future distant predictions. The comparison of the 5-step ahead output predictions performance by the ARX and the proposed ARMAX model indicate the superiority of the proposed ARMAX model predictor over the so-called ARX model predictor for distant predictions. This is further justified by the small output prediction errors produced by ARMAX model when compared to relatively large and infinite (NaN) error produced by ARX model as shown in the fourth column of Table 1.

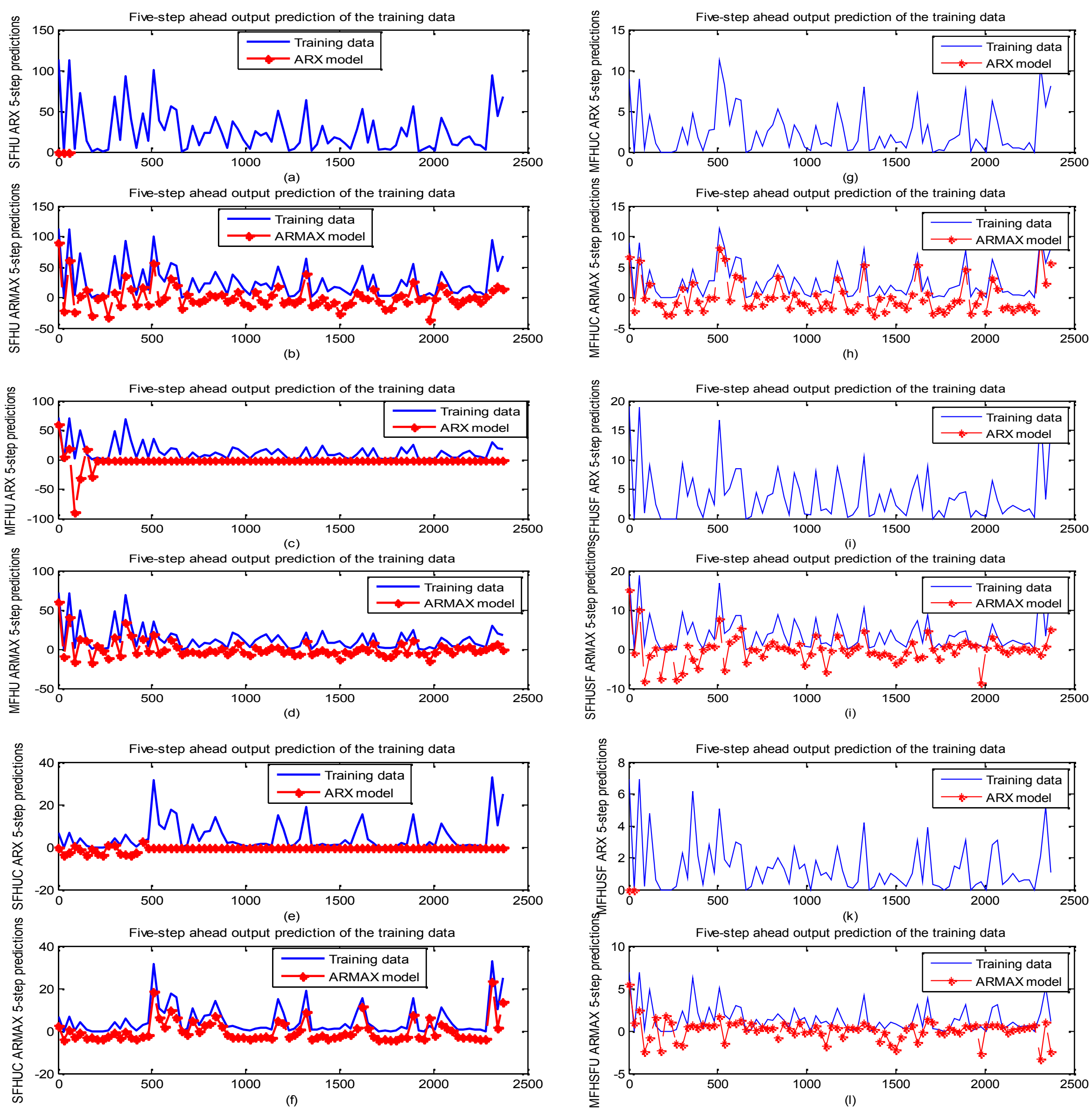

Fig. 6: Comparison of the five-step ahead output predictions of the original training data: (a) SFHU by ARX, (b) SFHU by ARMAX, (c) MFHU by ARX, (d) MFHU by ARMAX, (e) SFHUC by ARX, (f) SFHUC by ARMAX, (g) MFHUC by ARX, (h) MFHUC by ARMAX, (i) SFHUSF by ARX , (j) SFHUSF by ARMAX, (k) MFHUSF by ARX, (l) MFHUSF by ARMAX.

18.0 Conclusion

This paper presents the formulation of an ARX and ARMAX model predictors for model identification and prediction of power consumption in residential buildings. The 18,000 data used in the study has been obtained from the Power Holding Company of Nigeria, Akure. The results obtained from the application of these two model predictors for the modeling and prediction of power consumption in residential building as well as the validation results show that the ARMAX model outperforms the ARX model with much smaller predictions error and good prediction abilities with appreciable degree of accuracy and that the proposed ARMAX model predictor can be used for power consumption predictions in real scenarios.
Although the ARX model is stable while the ARMAX model is unstable due to the moving average filter coupled with the nonlinear nature of the power consumption data, the next aspect of the work is on the dynamic modeling and nonlinear model identification of the multivariable nonlinear systems using nonlinear neural network autoregressive moving average with exogenous input (NNARMAX) model which may give much better predictions and good tracking of the data for much more reliable power consumption predictions.

19.0 References

[1] Chowdhury,A. A., Rasul, M. G. and Khan, M. M. K. (2007): "Modelling and simulation of building energy consumption: a case study on an institutional building in central Queensland, Australia", *Proceedings: Building Simulation*, pp. 1916 – 1923.

[2] Hao, H., Kowli, A., Middelkoop, T., Barooah, P. and Meyn, S. (2013): "Using flexible HVAC power consumption of commercial buildings for regulation service to the grid", *Technical Report*, Available [Online]: http://plaza.ufl.edu/hehao/papers/ACC13_building.pdf.

[3] US Energy (2010): "U.S. Energy Information Administration", *Annual Energy Review 2009*, DOE/EIA-0384, August 2009 Edition.

[4] Li, X., Bowers, C. P. and Schnier, T. (2010): "Classification of Energy Consumption in Buildings with Outlier Detection", *IEEE Trans. Ind. Elect.*, vol. 57, no. 11, pp. 1 – 6.

[5] Seem, J. E. (2007): "Using intelligent data analysis to detect abnormal energy consumption in buildings", *Energy and Buildings*, vol. 39, pp. 52 – 58.

[6] Yu, W., Li, B., Lei, Y. and Liu, M. (2011): "Analysis of a Residential Building Energy Consumption Demand Model", *Energies*, vol. 4, pp. 475 – 487, doi: 10.3390/en4030475.

[7] Yao, R and Steemers, K. (2005): "A method of formulating energy load profile for domestic buildings in the UK", *Energy and Buildings*, vol. 37, no. 6, pp. 663 – 671.

[8] Zhang, T., Siebers, P. and Aickelin, U. (2011): "Modelling electricity consumption in office buildings: an agent based approach", *Energy & Building*, vol. 43, no. 10, pp. 2882 – 2892.

[9] Li, N., Kwak, J., Becerik-Geber, B. and Tambe, M. (2013): "Predicting HVAC energy consumption in commercial buildings using multi-agent systems", pp. 1 – 8, Available [Online]: http://teamcore.usc.edu/papers/2013/Camera%20Ready_Predicting%20HVAC%20Energy%20Consumption%20in%20Commercial%20Buildings%20Using%20Multiagent%20Systems.pdf.

[10] Swan, L. G. and Ugursal, V. I. (2009): "Modeling of end-use energy consumption in the residential sector: A review of modeling techniques", *Renewable and Sustainable Energy Reviews*, vol. 13, pp: 1819 – 1835.

[11] Hirst, E. (1978): "A model of residential energy use", *Simulation*, vol. 30, pp: 69–74.

[12] O'Neal, D. L. and Hirst, E. (1980): "An Energy Use Model of the Residential Sector", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 10, pp: 749–755.

[13] Wilson, J. W. (1971): "Residential demand for electricity", *Quarterly Review of Economics and Business*, vol. 11, pp: 7 – 22.

[14] Bunn, D. W. (2000): "Forecasting loads and prices in competitive power markets", *Proceedings of the IEEE 88*, pp: 163–169.

[15] Feinberg, E. A. and Genethliou, D. (2005): "Load Forecasting, in: J. H. Chow, F. F. Wu, J. Momoh (Eds.)", *Applied Mathematics for Restructured Electric Power Systems, Springer*, pp: 269–285.

[16] Arghira, N., Hawarah, L., Ploix, S. and Jacomino, M. (2012): "Prediction of appliances energy use in smart homes", *Energy*, vol. 48, pp: 128–134.

[17] Capasso, A., Grattier, W., Lamedica, R. and Prudenzi, A. (1994): "A bottom-up approach to residential load modeling", *IEEE Transactions on Power Systems*, vol. 9, pp. 957–964.

[18] Richardson, I., Thomson, M. and Infield, D. (2008): "A high-resolution domestic building occupancy model for energy demand simulations", *Energy and Buildings*, vol. 40, pp: 1560–1566.

[19] Widén, J. and Wäckelgård, E. (2010): "A high-resolution stochastic model of domestic activity patterns and electricity demand", *Applied Energy*, vol. 87, pp: 1880–1892.

[20] Muratori, M., Roberts, M. C., Sioshansi, R., Marano, V. and Rizzoni, G. (2013): "A highly resolved modeling technique to simulate residential power demand", *Applied Energy, March Edition*, pp. 1 – 15.

[21] Akpan, V. A. (Jul., 2011): "*Development of new model adaptive predictive control algorithms and their implementation on real-time embedded systems*", Department of Electrical and Computer Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece, Ph.D. Dissertation, 517 pages. [Online] Available: http://invenio.lib.auth.gr/record/127274/files/GRI-2011-7292.pdf.

[22] Akpan, V. A. and Hassapis, G. D. (2011) "Nonlinear model identification and adaptive model predictive control using neural networks", *ISA Transactions*, vol. 50, no. 2, pp. 177 – 194.

[23] Goodwin, G. C. and Sin, K. S. (1984): "Adaptive Filtering, Prediction and Control". *New Jersey, USA: Prentice-Hall.*

[24] Lung, L. (1999): "*System Identification: Theory for the User*". 2nd ed. Upper Saddle River, NJ: Prentice-Hall.

[25] Narendra, K. S. and Annaswamy, A. M. (1989): "Stable Adaptive Systems". *Englewood Cliff, NJ: Prentice-Hall.*

[26] Wang, L. (2009): *Model Predictive Control System Design and Implementation Using MATLAB*. London: Springer-Verlag.

[27] Maciejowski, J. M. (2002): Predictive Control with Constraints, England: Pearson Education Limited.

[28] Normey-Rico, J. E. and Camacho, E. F. (2007): *Control of Dead-Time Process,* 2nd ed., London: Springer-Verlag.

[29] Sjöberg, J. and Ljung, L. (1995): Overtraining, regularization, and searching for minimum in neural networks, *International Journal of Control,* vol. 62: 1391-1408.