Using Kalman Filter and Dynamic linear models for modeling and forecasting electricity load in Erbil city

(PP 347 - 373)

ID No.1993

https://doi.org/10.21271/zjhs.22.4.21

Kurdistan Ibrahim Mawlood
Salahaddin University – Erbil-College of Administration and Economics/ Statistics Department
kurdistan.mawlood@su.edu.krd

Rebaz Othman Yahya
M.Sc. in Statistics-Ministry of Higher Education and scientific Research
rebazstat@gmail.com

Received: 16 /01/2017
Accepted: 25/03/2018
Published: 10/09/2018

Abstract

In this paper we constructed four models for hourly (4*24 models for the 24 hours are 96 models of the day) and daily (4*7 models for the 7 days are 7 models of the week) electricity load using dynamic linear models (DLM) with various parameters. The Bayesian method and Kalman filter where used to estimate the parameters of four load models (Simple DLM, Trend DLM, Trend seasonal DLM and Regression DLM) and one step ahead forecasting.
In order to select the best and most efficient model for estimating and forecasting the electricity load in Erbil City, the four models were compared using (mean absolute error, mean absolute percentage error and root mean square error. The results presented in this paper based on real measured data. R-programming language and Microsoft Excel were used for data analyses. The result shows that Regression-DLM has best estimation and forecasting results compared to other models using the accuracy criteria, and the Kalman filter algorithm is a well-established technique and is suitable for estimating the parameters of load models that are used in this work as dynamic linear equations that include load signal with uncorrelated Gaussian white noise.

1. Introduction

Dynamic linear models provide a very rich class of models for the analysis and forecasting of signal data. They are used in a large number of applied areas in statistics, such as econometrics, signal processing, genetics, and population dynamics. Dynamic linear models (DLMs) are a particular class of state space models that allow many of the relevant inferences to be carried out exactly using the Kalman filter especially in the case of a completely specified model. At the same time they are edible enough to capture the main features of a wide array of diferent data, estimating unknown parameters in a DLM requires numerical techniques.

The system applied by DLMs is actually a sequence of models, which are updated at each time step, justifying the word 'dynamic'. The characteristic of interest or unknown parameters in the signal is modeled as posterior, possibly a vector of parameters, and the analysis of the development of posterior. Forecasting is also a common goal, but this also depends on how posterior is behaving over time. A model is fitting to the parameter(s) of interest at time t-1 and using this model a value is forecasted for time t. The model is then updated given this new information as well as any newly obtained and relevant outside information, the Bayesian analysis is the most natural method to allow for this updating of information at each time t. (Durbin and S. J. Koopman 2012).

Successful application of quantitative investment strategies rests on the availability of reliable forecasts of asset returns, volatilities and co-volatilities as they vary over time. Dynamic linear models (DLMs) are widely used due to flexibility in model specification, ability to changing location dynamics and predictor-outcome relationships, ability to incorporate external/intervention information, and flexibility in automatically incorporating series dependence characteristics among multiple series. So that models are updated dynamically, responding to the latest locations events, while being open to user-intervention at all times (Prado and West, 2010, West and Harrison, 1997).

The main aims for this research are:
1) Using dynamic linear models and its techniques for estimation and forecasting signal with Gaussian noise.

2) Using the dynamic linear models with the theoretical background estimation and forecasting electricity load for 24 hours and 7 days by using Kalman filtering based on dataset (2009-2013) which includes simple, trend, trend-seasonal components and regression models.

3) Selecting the best model(s) using Accuracy (or Error) Measures, and hence using the selected model for forecasting

2. Related Work

Stochastic filtering theory was first established in the early 1940s by Norbert Wiener and Andrey N. Kolmogorov, and it culminated in 1960 for the publication of classic Kalman filter (KF) (Kalman 1960).
Kalman & Bucy (1961) designed an algorithm to estimate dynamic state at continuous time which is called Kalman-Bucy Filter as a special case of nonlinear filter (Kalman & Bucy 1961).
Bucy (1970) developed the idea of filter theory special for stochastic process, taken the way series for difference between linear and nonlinear (Bucy 1970).
Harrison & Stevens (1976) defined the class of dynamic linear models (DLM's) and developed the Bayesian approach to dynamic modeling and forecasting (Harrison & Stevens 1976).
Harvey & Koopman (1993) suggested a structural univariate time series model for hourly electricity loads involving time-varying splines to model the intra weekly pattern (Harvey & Koopman 1993).
Taylor (2003) applied Double seasonal exponential smoothing models to make univariate electricity demand forecasting for lead times from a half hour ahead to a day ahead (Taylor 2003).
Kurdistan, M. (2008) used Kalman filter to estimate the delay time of calls for real data in Aria Phone system in Erbil City and to estimate the position and velocity of moving object of a simulated data of a radar system, She applied the Kalman filtering to signal estimation (Kurdistan 2008).
Hinman, J. & Hickey, E. (2009) used an autoregressive moving average model with exogenous weather variables (ARMAX) to forecast short-term electricity load using hourly load data from Commonwealth Edison Company. This model treats each hour’s load separately as an individual daily time series. Avoids modeling the complicated intraday pattern (load profile) displayed by the load, which varies through the week as well as through the seasons (Hinman & Hickey 2009).
Chaturvedi, D., Premdayalh, S. & Chandiok, A. (2010) worked on the Wavelet Transform Short Term Load Forecasting Soft Computing Techniques using Generalized Neural Network (GNN) (Chaturvedi, Premdayalh & Chandiok 2010).
Shankar, R., Chatterjee, K. & Chatterjee, T. (2012) developed a technique to establish the relationship between the economic load dispatch and load forecasting mechanism to the classical of the load frequency control (LFC) they used the Kalman filter prediction recursive algorithms and a bank of hourly predicted load data was obtained and then the concepts of 5-minute ahead forecasting (Shankar, Chatterjee & Chatterjee 2012).
Yoo, J. & Hur, K. (2013) conducted forecasting based on a selected model among ARMAX (autoregressive moving average with exogenous variable) models with the processed input data. They applied Kalman filtering to estimate model parameters (Yoo & Hur 2013).
Ramchandani, V. & Chowdhury, S. (2014) used load consumption and temperature of previous few days as parameters for estimation and forecasting data. Estimation was implemented

separately for (Weekdays & Weekends) Public Holidays, And then pplied Kalman filter algorithm to estimate parameters for hourly loads (Ramchandani & Chowdhury 2014).

Potočnik, P., Strmčnik, E. & Govekar E. (2015) applied stepwise regression, autoregressive model and neural network for short-term heat load forecasting (Potočnik, Strmčnik& Govekar 2015).

3. Background Information:

This section reviews signal and noise, the basic principles of modeling ideas and theoretical Bayesian forecasting and the structure of important special classes of dynamic models and their analyses. It is devoted to dealing with dynamic linear models for simple, trend, seasonal, trend-seasonal and regression models.

3.1: *Signals and Noise*

A signal is a function of independent variable such as time, distance, position and pressure. Signals represent air pressure as function of time at point in space and signal processing on area of applied mathematics that deals with operations (Smith 1999).

A signal carries information, and the objective of signal processing is to extract useful information carried by the signal. The method of information extraction depends on the type of signal and the nature of the information being carried by the signal (Prandoni & Vettenli 2008).

Noise can be defined as an unwanted signal that interferes with the communication or measurement of another signal. A noise itself is a signal that conveys information regarding the source of the noise.

Modeling and removal of the effects of noise and distortion have been at the core of the theory and practice of communications and signal processing. Noise reduction and distortion removal are important problems in applications such as cellular mobile communication, speech recognition, image processing, medical signal processing, radar, sonar, and in any application where the signals cannot be isolated from noise and distortion (Benesty, Chen, Huang & Cohen 2009).

3.2: Dynamic Linear Models

Dynamic linear models provide a very flexible yet fairly simple tool for analyzing dynamic phenomena and evolving systems, and have significantly contributed to extend the classical domains of application of statistical time series analysis to non-stationary for continuous and discrete data. The Kalman filter can use to estimation and forecasting in a dynamic linear model with known parameters (Prado & West 2010).

The main area in which DLMs are used is in modeling observations collected over time for the purposes of forecasting or detection of model shifts. The characteristic of interest or unknown parameters in the signal is modeled as θ. A model fits to the parameter(s) of interest at time t-1 and using this model a value is forecasted for time t, a dynamic linear model is specified by its parameter quadruple $\{F,G,V,W\}_t$ (West & Harrison, 1997).

The dynamic linear model is specified by the following observation equation and system equation:

Observation equation: $Y_t = F_t\theta_t + \upsilon_t \qquad \upsilon_t \sim N[0,V_t]$ …. (5.1)

System equation: $\theta_t = G_t\theta_{t-1} + \omega_t \qquad \omega_t \sim N[0,W_t]$ …. (5.2)

t = 1, 2, …,n

Y_t : Is the vector of observation at time t, of dimension mx1.

F_t : Is the known matrix of non-random variable at time t, of dimension mxp.

θ_t : Is the vector of unknown state parameters at time t, of dimension px1.

υ_t : Is the vector of observation error (stochastic error term) at time t, of dimension mx1.

G_t : Is the matrix of evolution or system (known coefficients), of dimension pxp.

ω_t : Is the vector of evolution error (stochastic error term) at time t of dimension px1.

The two stochastic terms (V_t) and (W_t) are assumed to be independent.

3.3 *Types of Dynamic Linear Models*:

3.3.1 *The Constant Dynamic Linear Models*

Assume that we know that the level $\mu_t = \mu_0$ is a constant does not change with time it is one of the simplest linear models (Petris, Petrone & Campagnoli 2009).

Observation equation: $Y_t = \mu_t + \varepsilon_t \quad e_t \overset{iid}{\sim} N(0, s_e^2)$ …. (3.1)

t = 1 , 2 , …,n

μ_t represents unknown parameters model θ_t but level is constant don't change over time

3.3.2 *The Simple Dynamic linear Models*

The simplest model for a univariate process is so-called random walk plus noise model. it is appropriate for time series which shows no clear trend or seasonal variation, which will be expected to change significantly over time. This change is due to the jump.

The observations (Y_t) are modeled as noisy observations of a level μ_t it is randomly changing over

This model can be considered as:

Observation equation $Y_t = \mu_t + \varepsilon_t \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$ …. (3.2)

System equation $\mu_t = \mu_{t-1} + \sigma_{\mu_t} \quad \sigma_{\mu_t} \overset{iid}{\sim} N(0, \sigma_\mu^2)$ …. (3.3)

t = 1, 2, … ,n

Where:

$W_t = W = \sigma_\mu^2$, $V_t = V = \sigma_\varepsilon^2$, $G_t = G = 1$, $F_t = F = 1$, $\theta_t = \mu_t$

It is characterized by the quadruple {1, 1, V, W} like this in the standard DLM $\{1, 1, \sigma_\varepsilon^2, \sigma_\omega^2\}$.

3.3.3 *The Trend (Slope) Dynamic linear Models*

Trend DLMs assumes that the current level μ_t changes linearly through time and that the slop rate may also evolve (Petris 2010).

This model equally spaced interval can be considered as:

Observation equation: $Y_t = \mu_t + \varepsilon_t \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$ …. (3.4)

System equation: $m_t = m_{t-1} + b_t + s_{m_t} \quad s_{m_t} \overset{iid}{\sim} N(0, s_m^2)$ …. (3.5)

$$\beta_t = \beta_{t-1} + \sigma_{\beta_t} \qquad \sigma_{\beta_t} \overset{iid}{\sim} N(0, \sigma_\beta^2) \qquad \text{…. (3.6)}$$

t = 1, 2, …, n

3.3.4 ***The Seasonal Dynamic Linear Models*** SEDLM

Cyclical or periodic behavior is evident in many time series associated with economic, commercial, physical and biological systems. The term seasonality is used as in many applications; this period represents a seasonal cycle, quarter, month, or year.

We have two ways of modeling a time series which shows a cyclical behavior or seasonality:

1. The seasonal factor model.
2. The Fourier-form seasonal model.

In this paper, we use the second form for modeling the load data

3.3.4.1 *The Fourier – Form Seasonal Models*

We using Fourier analysis for estimating periodic parameters of hydrologic time series shows that for small time interval series, such as hourly, daily series only the first few harmonics are necessary for a good Fourier series fit in the periodic estimate of model parameters.
We introduce the following trigonometric representation of seasonal components based on Fourier series (West & Harrison 1997, Pole, West, & Harrison 1994).
The SEDLM the representation of seasonality components as:
The S_t^i is evolution of time t is:

$$Y_t = S_t^i + \varepsilon_t \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$$

$$S_t^i = \sum_{j=1}^{n_i} S_{j,t}^i$$

$$S_{j,t}^i = S_{j,t-1}^i \cos(w_j) + S_{j,t-1}^{i*} \sin(w_j)$$

$$S_{j,t}^{i*} = -S_{j,t-1}^i \sin(w_j) + S_{j,t-1}^{i*} \cos(w_j)$$

Where the Fourier frequencies

$$w_j = \frac{2\pi j}{C} \qquad j = 0, 1, 2, 3, \dots, \frac{C}{2}$$

We describe the stochastic level of the i^th seasonal component by $S_{j,t}^i$, and the i^th seasonal component that is needed to describe the change in the seasonal component over time by $S_{j,t}^{*i}$ is denoted n_i., the number of harmonics required for the i^th seasonal component.

3.3.5 The Trend – Seasonal Dynamic Linear Models

The seasonal pattern to change over time, a simple way to achieve this is to add an error term ε_t to this relation giving the model (West & Harrison 1997).

Observation equation: $Y_t = \mu_t + S_t^i + \varepsilon_t \quad , \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$ …. (3.7)

Where

$$S_t^i = \sum_{j=1}^{n_i} S_{j,t}^i \qquad \text{.... (3.8)}$$

Then we can write the observation equation as

$$Y_t = \mu_t + S_{j,t}^i + S_{j,t}^{i*} + \varepsilon_t \qquad \text{.... (3.9)}$$

System equation:

$$\mu_t = \mu_{t-1} + \beta_t + \sigma_{\mu_t} \qquad \sigma_{\mu_t} \overset{iid}{\sim} N(0, \sigma_\mu^2) \qquad \text{.... (3.10)}$$

$$\beta_t = \beta_{t-1} + \sigma_{\beta_t} \qquad \sigma_{\beta_t} \overset{iid}{\sim} N(0, \sigma_\beta^2) \qquad \text{.... (3.11)}$$

$$S_{j,t}^i = S_{j,t-1}^i \cos(w_j) + S_{j,t-1}^{i*} \sin(w_j) + \zeta_{S_{j,t}} \qquad \text{.... (3.12)}$$

$$\zeta_{S_{j,t}} \overset{iid}{\sim} N(0, \sigma_{S_{j,t}}^2)$$

$$S_{j,t}^{i*} = -S_{j,t-1}^i \sin(w_j) + S_{j,t-1}^{i*} \cos(w_j) + \zeta_{S^*_{j,t}} \qquad \text{.... (3.13)}$$

$$\zeta_{S^*_{j,t}} \overset{iid}{\sim} N(0, \sigma_{S^*_{j,t}}^2)$$

When:
j^{th} harmonic J=1, 2, 3, 4,…., $\frac{c}{2}$, If j = 1, and w = $\frac{2\pi}{c}$.

3.3.6 The Regression Dynamic Linear Models

The effects of other variables on the development of a particular dynamic linear model, the explanatory, regression variables can be added to the measurement equation of the model (Gentleman, Hornik, & Parmigiani 2009).
If regression variables are added to the local level model, then the measurement equation becomes.

$$y_t = \beta_0 + \sum_{i=1}^{k} \beta_{it} x_{it} + \varepsilon_t \qquad \text{.... (3.14)}$$

The assumption of i.i.d. errors is often not realistic if the observations are taken over time. This is in fact quite interesting in many problems, is to think that the relationship between y and x evolves over time.
That is to consider a dynamic linear regression model in this paper of the form:
Observation Equation:

$$y_t = \beta_{0,t} + \beta_{1,t} x_{1t} + \beta_{2,t} x_{2t} + \beta_{3,t} x_{3t} + \beta_{4,t} x_{4t} + \varepsilon_t \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \quad \text{.... (3.15)}$$

The parameter of the model are ($\beta_{0,t}$, $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$ و $\beta_{4,t}$).
System Equation:

$$\beta_{i,t} = \beta_{i,t-1} + \omega_{i,t} \qquad \text{.... (3.16)}$$

i=0, 1, 2, 3, 4

$\omega_{0,t}$, $\omega_{1,t}$, $\omega_{2,t}$, $\omega_{3,t}$ and $\omega_{4,t}$ are independent. Where $F_t = x_{1,t}, x_{2,t}, x_{3,t}, x_{4,t}$ are the values of the four explanatory variables at time t.

3.4 Kalman Filtering

The Kalman filter is the Bayesian optimum solution to the problem of sequentially estimating the states of a dynamical system in which the state evolution and measurement processes are both linear and Gaussian. Kalman filter is an optimal linear estimator which provides the estimation of signals in noise. Kalman used the state transition models for dynamic system. The Kalman filter allows estimating the state of dynamic systems with certain types of random behavior by using such statistical information (Grewal & Andrews 2008).

The optimal state vector (variables) estimation process applied to a dynamic system that involves random disturbance. The Kalman filter gives a linear, unbiased, and minimum error variance recursive algorithm to optimally estimate the unknown state of a dynamic system from noisy data taken at discrete real-time (Eubank 2006).

The solution is recursive in that each updated estimate of the state is computed from the previous estimate and the new input data, therefore only the previous estimate requires storage.

The Kalman filter is computationally more efficient than computing the estimate directly from the entire past observed data at each step of the filtering process (Schutter, Geeter, Lefebvre & Bruyninckx 1999).

3.5 Bayesian Representation of Dynamic Linear Models

For each time t, the general dynamic linear model is defined by the following two equations are:

Observation equation: $Y_t = F_t\theta_t + \upsilon_t \qquad \upsilon_t \sim N[0, V_t]$

System equation: $\theta_t = G_t\theta_{t-1} + \omega_t \qquad \omega_t \sim N[0, W_t]$

The initial information about parameter θ_0 at time zero is:

$\theta_0 \sim N(M_0, C_0)$

By induction and using Bayes' theorem we have:

$(\theta_{t-1} \mid D_{t-1}) \sim N(M_{t-1}, C_{t-1})$

Where

$D_{t-1} = \{Y^{t-1}, \sigma_\varepsilon^2, \sigma_\mu^2, \sigma_\beta^2)$

and

Posterior $\propto$ Observed likelihood × Prior

$P(\theta_t \mid D_t) \alpha P(Y_t \mid \theta_t) * P(\theta_t \mid D_{t-1})$

Where

$D_t = \{D_{t-1}, Y_t\}$

to find mean and variance for each of the likelihood distribution $P(Y_t \mid \theta_t)$ and the prior distribution $P(\theta_t \mid D_{t-1})$ respectively:

$(Y_t \mid \theta_t) \sim N(F\theta_t, \sigma_\upsilon^2)$

and

$(\theta_t \mid D_{t-1}) \sim N(GM_{t-1}, Q_t)$

Where:

$= GC_{t-1}G^T + W = Q_t$

Substituting for each of $(Y_t \mid \mu_t)$ and $P(\mu_t \mid D_{t-1})$ we get:

$$P(\theta_t \mid D_t) \alpha e^{-\frac{1}{2}[(Y_t - F\theta_t)^T (\sigma_\upsilon^2)^{-1}(Y_t - F\theta_t) + (\theta_t - GM_{t-1})^T (Q_t)^{-1}(\theta_t - GM_{t-1})]}$$

$$P(\theta_t \mid D_t) \alpha e^{-\frac{1}{2}[(\theta_t - M_t)^T (C_t)^{-1}(\theta_t - M_t)]}$$

We can obtain that:

$C_t = [Q_t^{-1} + F^T(\sigma_\upsilon^2)F]^{-1}$

Or it can be written as:

$C_t = Q_t + K_t F Q_t$

Where

$K_t = Q_t F^T [F Q_t F^T + \sigma_\upsilon^2]^{-1}$

and

$M_t = GM_{t-1} + K_t(Y_t - FGM_{t-1})$

or

$(\theta_t \mid D_t) \sim N(M_t, C_t)$

Then we can write the Kalman filter equation as:

$$M_t = GM_{t-1} + K_t(Y_t - FGM_{t-1}) \quad \ldots.\ (3.17)$$

$$C_t = Q_t + K_t F Q_t \quad \ldots.\ (3.18)$$

$$K_t = Q_t F^T [F Q_t F^T + \sigma_\varepsilon^2]^{-1} \quad \ldots.\ (3.19)$$

From the relations (3.17) and (3.18) we obtain:

$C_t = \begin{pmatrix} c_{1,t} & c_{2,t} \\ c_{2,t} & c_{3,t} \end{pmatrix}$ $Q_t = \begin{pmatrix} r_{1,t} & r_{2,t} \\ r_{2,t} & r_{3,t} \end{pmatrix}$ $K_t = \begin{pmatrix} k_{1,t} \\ k_{2,t} \end{pmatrix}$

Where:

$M_t = E(\theta_t \mid D_t) = E = \left(\begin{matrix} \theta_t \\ \beta_t \end{matrix} \middle| D_t\right) = \begin{pmatrix} a_t \\ f_t \end{pmatrix}$

Where

$$\left.\begin{matrix} a_t = (a_{t-1} + f_{t-1}) + k_{1,t}(a_{t-1} - f_{t-1}) \\ \\ \\ \end{matrix}\right\}$$

$$f_t = f_{t-1} + k_{2,t}[Y_t - (a_{t-1} - f_{t-1})] \quad \text{…. (3.20)}$$

$$\left.\begin{aligned} k_{1,t} &= \frac{r_{1,t}}{r_{1,t} + \sigma_\varepsilon^2} \\ k_{2,t} &= \frac{r_{2,t}}{r_{2,t} + \sigma_\varepsilon^2} \end{aligned}\right\} \quad \text{…. (3.21)}$$

$$r_{1,t} = c_{1,t} + 2c_{2,t-1} + c_{3,t-1} + \sigma_\beta^2$$
$$r_{2,t} = c_{2,t-1} + c_{3,t-1} + \sigma_\beta^2$$
$$r_{2,t} = c_{3,t-1} + \sigma_\beta^2$$

$$\left.\begin{aligned} c_{1,t} &= k_{1,t} + \sigma_\beta^2 \\ c_{2,t} &= k_{2,t} + \sigma_\beta^2 \\ c_{1,t} &= r_{3,t} - k_{1,t} + \sigma_\beta^2 \end{aligned}\right\} \quad \text{… (3.22)}$$

Then the equation (3.20) is filter for parameter from model (μ_t, β_t) , equation (3.21) is component of Kalman gain filter gain , equitation (3.22) is component gain for the variation estimate θ_t (West & Harrison 1997, Gentleman, Hornik & Parmigiani 2009).

3.6 Forecasting Dynamic Linear Models

We mean by forecasting finding the *h*-step ahead for each distribution given the data $y_1, y_2, ..., y_t$. The DLM based on the normal distribution of initial parameters.

With $y_{1:t}$ at hand, one can be interested in forecasting future values of the observations, Y_{t+h} , or of the state vectors.

For a dynamic linear model, the *h*-step-ahead forecasting distributions, for states and observations, are obtained as a product of the Kalman filter (West & Harrison 1997, Meng 2011).

Then the general dynamic linear model for h step is:

Observation equation: $Y_{t+h} = F_{t+h}\theta_{t+h} + \upsilon_{t+h} \qquad \upsilon_t \sim N[0, V_t]$ …. (3.23)

System equation: $\theta_{t+h} = G_{t+h}\theta_{t+h-1} + \omega_{t+h} \qquad \omega_t \sim N[0, W_t]$ …. (3.24)

Let $a_0 = m_t, R_0 = C_t$, then for $h \geq 1$, for forecasting at time t, the forecaster requires the *h*-step ahead marginal distributions, $(Y_{t+h} \mid D_t)$ and $(\theta_{t+h} \mid D_t)$. Predictions start from the posterior estimate of the state, obtained from the
Kalman filter, on the time (t) on which the forecast is made.
678We can make *h*-step ahead forecast in the follows are:
The mean and variance for observation equation is:

$$\hat{Y}_{t+h} = E(Y_{t+h} \mid D_t)$$
$$= E(F_{t+1}\theta_{t+1} + \upsilon_{t+1} \mid D_t)$$

$$= FG^{k} m_t \qquad \ldots (3.25)$$

$$Q_t(t) = Var(Y_{t+h} \mid D_t)$$
$$= FG^{h} C_t (G^{k})^{T} F^{T} + \sum_{i=0}^{h-1} FG^{i} W (G^{i})^{T} F^{T} + V \qquad \ldots (3.26)$$

3.7 Load Forecasting

Accurate models for electricity power load forecasting are essential to the operation and planning of a utility company. Load forecasting helps an electric utility to make important decisions including decisions on purchasing and generating electricity power, load switching, and infrastructure development. Load forecasting is a way of estimating what future electricity load will be for a given forecast horizon based on the available information about the state of the system. Short-term forecasting is important because the national grid requires a balance between the electricity produced and consumed at any moment in ~~the~~ hours (Hong 2010).

Load forecasts can be divided into four categories: very short term load forecasting, forecasting horizon ranging from a few minutes ahead to a few hours ahead, short-term forecasts which are usually from one hour to one week, medium forecasts which are usually from a week to a year, and long-term forecasts which are longer than a year. There are two types of forecasting models. First, Qualitative is based on experience, judgments, and knowledge. Second, Quantitative is based on statistics data (Rothe &Wadhwani 2009).

Electricity demand forecasting is of great importance for the management of power systems. Long-term forecasts of the peak electricity demand are needed for capacity planning and maintenance scheduling. Medium-term demand forecasts are required for power system operation and planning. Short-term load forecasts are required for the control and scheduling of power systems. Short-term forecasts usually correspond to a forecasting horizon from the hours to the next hours up. Short-term forecasting models are therefore based on high frequency data, usually hourly and daily measured (Willis 2002).

Electric load has an obvious correlation to weather; the most important variables responsible in load changes are dry and wet bulb temperature, dew point humidity, wind speed and wind direction sky.

3.8 Evaluating Accuracy (or Error) Measures

Evaluating measures are indispensable for helping us to (a) choose an appropriate model among the many available, (b) select the best method for our particular estimation and forecasting situation, (c) measure the most likely size of estimation and forecasting errors and (d) quantify the extent of uncertainty surrounding the estimates forecasts (Rabinovich 2010).

In judging the value of models or methods and the size of their estimation and forecasting errors uncertainty, it is critical to distinguish between models fitting and post-sample measures. In this paper we used thre evalution accuracy measures which are (mean absolute error *(MAE)*, mean absolute percentage error *(MAPE)* and root mean square error *(RMSE))*

4: Application Part

4.1 *Introduction*

This section presents an analysis of the electricity load data. The data set analyzed in this section is the hourly electricity load covering the period starting on (Jun 1, 2009) and ending on (Dec 31, 2013) obtained from (General Directorate of Erbil-Dispatch Control Center of Erbil and General Directorate of Agriculture-Erbil).

Load profile for each hour in the day and for each day in the week possesses different patterns according to treating them separately it produces better result than combing all (hours and days) as they have the same profile patterns.

For this reason the paper generated 24 profiles or small data sets each set consists of (1461) values for hourly estimation and 7 profile or small data sets each set consists of (209) values for analyzing weekdays load.

Due to increasing the accuracy for load estimate and forecast, four models (Simple-DLM, Trend-DLM, Trend and Seasonal-DLM, and Regression-DLM) have been proposed for each data set for the short term-load forecasting problem, where the first two models are non-weather sensitive and the rest two models are weather sensitive. The model coefficients are estimated, using the Kalman filter. The three criteria used for model selection are (MAE, MAPE and RMSE), one model is better than the other if it has smaller accuracy criteria.

4.2 Characteristics of the Power System Electricity Load

The system electricity load is the sum of the entire consumers, electricity load at the same time. The objective of system terms electricity load forecasting is to forecast the future system electricity load. A good understanding of the system characteristics helps to design reasonable forecasting models and select appropriate models operating in different situations. Various factors that influence the system load behavior can be classified into the following major categories:

1. Weather

The weather variable**s** are shown to have greatest correlation with season load and influence the electricity load, the factors that can contribute significantly to weather-dependent electricity load are:

a. Temperature
b. Rate of Humidity
c. Wind speed

2. Time

Time factors influencing the electricity load at time point of the hourly, daily and season property. From the observation of the load curves it can be seen that there are certain rules of the electricity load variation with the time point of the hour or day.

4.3 Estimation Hourly and Daily Electricity Load For Real Data by Kalman Filter

The main point of estimation and forecasting using the Kalman filter is initializing the parameter values, mean and variance $\hat{\theta}_0$ and $\hat{P}_0$ of data respectively. when the Kalman gain and covariance converges, this point is the best estimate, and this point is selected since it is the best estimate.

Y_t is the hourly electricity load and the data used covers four years consisting of (35064) observations . The initial state θ_0 is Gaussian with mean $\hat{\theta}_0$ and the covariance matrix $\hat{P}_0$.

The table 3.1 describes (35064) hourly electricity load observations with three independent weather variables (temperature, rate of humidity and wind speed) , from January 1, 2009 to December 31, 2013).

Table (3.1) Descriptive of the observation for variables

N. Observation	Hours	Electricity Load (Y)	T.C. (X1)	R.H. (X2)	W.S. (x3)
1	1	453.5	8.62	86.4	0.255
2	2	461.5	7.467	91.5	0
3	3	454.5	7.331	93.6	0.003
4	4	437.5	6.783	93.8	0.029
5	5	434.4	6.545	94.7	0.093

N. Observation	Hours	Electricity Load (Y)	T.C. (X1)	R.H. (X2)	W.S. (x3)
6	6	442.9	6.205	94.6	0.364
7	7	442.2	5.626	94.2	0.173
8	8	449.4	6.757	92	0.289
9	9	453.9	7.768	85.8	0.467
10	10	447.2	6.296	95.5	0.721
11	11	460.1	9.67	84.3	0.673
12	12	448.1	8.64	90.4	0.656
13	13	446	8.61	90.5	0.51
14	14	450	8.98	87.4	0.448
15	15	430.5	9.7	84.5	0.474
16	16	423.8	9.5	84.8	0.263
17	17	441.5	8.58	88.4	0.123
18	18	422.9	8.29	90.8	0
19	19	420.2	7.947	92.8	0.025
20	20	415.8	7.604	93	0
21	21	420.5	7.741	93.2	0.067
22	22	431	7.194	94.2	0.043
23	23	439	6.442	92.8	0
24	24	421.8	8.22	89.6	0.255
25	1	414.6	6.32	93.7	0.276
26	2	428.7	6.315	90.3	0.507
27	3	316.4	5.015	95.3	0.228
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
35062	22	1141.3	9.85	67.12	1.165
35063	23	1213.7	9.16	75.47	1.234
35064	24	1205.6	8.52	81.3	0.517

4.3.1. Hourly Electricity Load Estimation

The estimation process of parameters is made using the initial values of electricity load mean and variances are being of an estimate of it is values in each hour.

Case One: SDLM of equations (3.2) and (3.3) are used to estimate parameters, Kalman Gain and Error Covariance for univariate model for each twenty four hours. The electricity load of each hour of the day is used to estimate the electricity load parameters of the same hour. Table 3.5 shows all convergences points for twenty-four hours. Including estimated electricity load parameter and actual values with Kalman Gain, Error Covariance and Error Estimate of the time which is convergence.

Table (3.2) Estimated load parameters using SDLM

Hours	Convergence Time	Actual Values	Estimate	Kalman Gain Load	Error Covariance	Error Estimate
1	16	439.1	439.0119	0.4442	6885.0422	0.0881
2	16	449.66	445.5144	0.4442	6885.0422	4.1456
3	16	436.2	440.6219	0.4442	6885.0422	-4.4219
4	16	450.45	429.2103	0.4442	6885.0422	21.2397
5	16	427.58	432.5572	0.4442	6885.0422	-4.9772
6	16	445.24	441.9727	0.4442	6885.0422	3.2673
7	16	456.24	453.771	0.4442	6885.0422	2.469
8	16	451.25	443.4807	0.4442	6885.0422	7.7693
9	16	452.07	434.9009	0.4442	6885.0422	17.1691
10	16	429.38	429.9956	0.4442	6885.0422	-0.6156
11	16	447.94	422.19	0.4442	6885.0422	25.75
12	16	418.83	415.6895	0.4442	6885.0422	3.1405
13	16	427.45	417.678	0.4442	6885.0422	9.772

Hours	Convergence Time	Actual Values	Estimate	Kalman Gain Load	Error Covariance	Error Estimate
14	16	421.25	418.0306	0.4442	6885.0422	3.2194
15	16	435.7	418.5257	0.4442	6885.0422	17.1743
16	16	427.8	410.0586	0.4442	6885.0422	17.7414
17	16	431.7	411.0182	0.4442	6885.0422	20.6818
18	16	414.9	433.5536	0.4442	6885.0422	-18.6536
19	16	421.4	433.6256	0.4442	6885.0422	-12.2256
20	16	402.67	416.2096	0.4442	6885.0422	-13.5396
21	16	401.17	410.6931	0.4442	6885.0422	-9.5231
22	16	408.37	409.6359	0.4442	6885.0422	-1.2659
23	16	397.21	421.0353	0.4442	6885.0422	-23.8253
24	16	410.66	427.5384	0.4442	6885.0422	-16.8784

Table 3.2 shows that:

In the first hour (hour one) of the day, the Error Covariance is convergence at time point (t=16) with the value (6885.0422). Kalman Gain is convergence at time point (t=9) with the value (0.4442). There are two convergences time and the last time convergence is the best estimate for the same time.

In the second hour (hour two) of the day, the Error Covariance is convergence at time point (t=16) with the value (6885.0422). Kalman Gain is convergence at time point (t=9) with the value (0.4442). There are two time convergences and the last time convergence is the best estimate for the same time.

In the last hour (twenty-four) of the day, the Error Covariance is convergence at time point (t=16) with the value (6885.0422). Kalman Gain is convergence at time point (t=9) with the value (0.4442). There are two time convergences and the last time convergence is the best estimate for the same time.

A compression between electricity load resulting from the estimate and actual electricity load values for hour (one) in SDLM is shown in Figure 3.1. The results show how closely the estimate model matches the actual electricity load. The blue line is actual values and the red line is estimate value

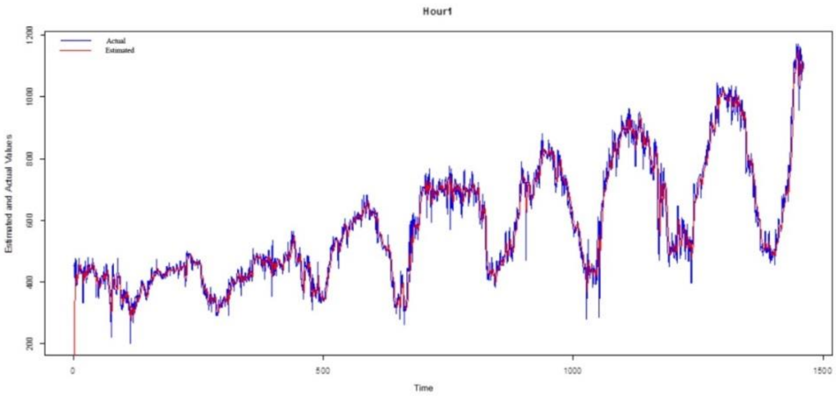

Figure (3.1) Difference between actual values and estimate values for hour one

Case Two: TDLM of equations (3.4), (3.5) and (3.6) are used to estimate parameters, Kalman Gain and Error Covariance for univariate model for each twenty four hours. The electricity load of the each hour of the day is used to estimate the electricity load parameters.

Table 3.9 shows all convergences points for twenty-four hours. Including estimated electricity load parameter and actual values with Kalman Gain, Error Covariance and Error Estimate of the time which is convergence.

Table (3.3) Estimated load parameters using TDLM

Hours	Convergence Time	Actual Values	Estimate Load	Estimate Trend	Kalman Gain Load	Kalman Gain Trend	Error Estimate

Hours	Convergence Time	Actual Values	Estimate Load	Estimate Trend	Kalman Gain Load	Kalman Gain Trend	Error Estimate
1	35	446.5	441.22854	0.23475	0.69638	0.18133	5.31146
2	34	443.3	443.53888	-0.94385	0.70038	0.18594	-0.21888
3	33	453.3	448.58223	0.09051	0.70421	0.19038	4.71777
4	32	426.4	430.70287	-5.09363	0.7098	0.19692	-4.35287
5	33	446.9	441.77011	0.6353	0.71599	0.20421	5.10989
6	32	445.7	451.35178	-2.49702	0.71483	0.20283	-5.70178
7	33	454.3	448.7475	-2.77035	0.70699	0.19362	5.5725
8	37	452	438.14193	4.79563	0.70433	0.19052	13.81807
9	35	397.4	404.81265	-5.34553	0.69581	0.18068	-7.41265
10	37	447.6	437.15028	5.55225	0.69105	0.17524	10.39972
11	37	372	388.69309	-10.22054	0.68837	0.1722	-16.66309
12	39	421	427.74937	1.10524	0.68923	0.17317	-6.71937
13	36	421.6	422.18252	3.27239	0.6913	0.17553	-0.63252
14	37	445	442.36976	8.3545	0.69233	0.17669	2.63024
15	35	415.4	425.54231	0.98268	0.69451	0.17918	-10.13231
16	36	404.1	409.98391	-3.42292	0.69482	0.17955	-5.93391
17	35	396.2	403.60154	-4.14609	0.69322	0.17771	-7.44154
18	36	424.6	424.06869	-0.78729	0.69203	0.17635	0.51131
19	38	395.9	417.04614	-5.00936	0.6927	0.17712	-21.10614
20	36	455.5	442.71127	4.53651	0.69124	0.17545	12.78873
21	36	438.8	434.50797	3.66707	0.69336	0.17787	4.29203
22	36	435.1	433.67433	4.87317	0.69497	0.17971	1.41567
23	36	368.2	395.94253	-11.09241	0.6966	0.18158	-27.78253
24	35	455.4	453.17555	3.45517	0.69512	0.17989	2.20445

Table (3.4) Error Covariance parameters using TDLM

Hours	Error Covariance Load	Error Covariance Trend
1	25415.6962	15179.07687
2	23989.95237	14887.89055
3	22715.95246	14620.36413
4	20996.05996	14247.17342
5	19272.41005	13857.76795
6	19584.1937	13929.43301
7	21840.20414	14432.1481
8	22676.90775	14612.04964
9	25626.58907	15221.4542
10	27480.09504	15586.69631
11	28593.05823	15800.14056
12	28231.53112	15731.26789
13	27376.67348	15566.6454
14	26966.47069	15486.74362
15	26118.99298	15319.73039
16	25998.28285	15295.7242
17	26616.01752	15418.00032
18	27086.41091	15510.16845
19	26818.39378	15457.75202
20	27403.57842	15571.86523
21	26560.33891	15407.03734
22	25943.57396	15284.82577
23	25335.11044	15162.83774
24	25885.60096	15273.26465

Tables (3.3) and (3.4) are show that:

In the first hour (one) Error Covariance to electricity load parameter is convergence at time point (t=35) with the value (25415.6962) and the Error Covariance for trend parameter is convergence at time point (t=35) with the value (15179.07687). The Kalman Gain to electricity load parameter is convergence at time point (t= 17) with the value (0.69638) and the Kalman Gain to trend parameter is convergence at time point (t= 18) with the value (0.18133). There are four time convergences and the last time convergence is the best estimate for the same time.

A compression between electricity load resulting from the estimate and actual electricity load values for hour (one) in TDLM is shows in Figure 3.2. The results show how closely the estimate model matches the actual electricity load. The blue line is actual values and the red line is estimate values.

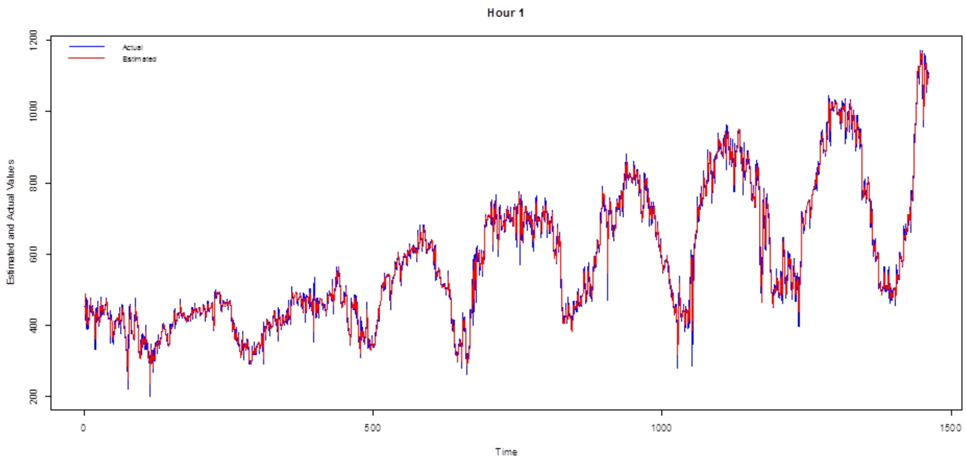

Figure (3.2) Difference between actual values and estimate values for hour one

In the second hour (two) of the day, the Error Covariance to electricity load parameter is convergence at time point (t=34) with the value (23989.95237) and the Error Covariance for trend parameter is convergence at time point (t=34) with the value (14887.89055).

The Kalman Gain to electricity load parameter is convergence at time point (t=17) with the value (0.70038) and the Kalman Gain to trend parameter is convergence at time point (t=17) with the value (0.18594). There are four time convergences and the last time convergence is the best estimate for the same time.

In the last hour (twenty-four) of the day, the Error Covariance to electricity load parameter is convergence at time point (t=35) with the value (25885.60096) and the Error Covariance for trend parameter is convergence at time point (t=35) with the value (15273.26465).

The Kalman Gain to electricity load parameter is convergence at time point (t=18) with the value (0.69512) and the Kalman Gain to trend parameter is convergence at time point (t=17) with the value (0.17989). There are four time convergences and the last time convergence is the best estimate for the same time.

Case Three: In the preliminary analysis we saw the presence of seasonal effects in the data. TSEDLM which contains the seasonal factor of equations (3.7) to (3.13) are used to estimate parameters, Kalman Gain and Error Covariance for univariate model for each twenty-four hours. In this analysis seasonal component (C=4) due to the quartile cycle is used. The electricity load of each hour of the day is used to estimate the electricity load parameters include trend and seasonal.

Table 3.17 shows all convergences points for twenty-four hours. Including estimated electricity load parameter and actual values with Kalman Gain, Error Covariance and Error Estimate of the time which is convergence.

Table (3.5) Estimated load parameters using T-SEDLM

Hours	Covariance Time	Actual	Estimate Load	Estimate Trend	Estimate S	Estimate S*	Error Estimate
1	105	351.8	352.6926297	0.166029049	-4.527108865	1.004859043	-0.892629659
2	111	300.3	304.1820616	-5.27795208	5.789328762	-5.472258677	-3.882061614
3	95	287.5	307.2193889	-7.443953535	-7.567995688	4.57002198	-19.71938888
4	97	285.2	283.1524247	-8.579210568	3.919503958	-3.992815502	2.047575311
5	99	332	320.4926577	-0.926289592	7.804878534	-3.34790436	11.50734232
6	98	312.1	313.7464755	-6.407101993	-4.241459777	-1.979653036	-1.646475501
7	91	434.3	427.9771019	7.295172046	-12.06590867	1.81503659	6.322898107
8	97	375.6	375.540603	2.684366031	-9.124081258	8.367417002	0.059396997
9	88	445.9	438.4509597	2.041382065	-9.038164776	7.109298724	7.449040277
10	98	386.2	386.9761382	-0.915213769	6.137721032	-7.887235265	-0.776138189
11	116	388.3	380.6929441	3.998426017	5.069833166	6.502481958	7.607055888
12	105	339.9	346.4326476	-1.659137798	-11.66830884	2.871220276	-6.532647591
13	98	364.5	358.0905311	-4.245098358	0.354329498	1.226918135	6.409468936
14	95	345.4	347.9968415	-9.580104438	11.53618307	5.711467698	-2.596841454
15	96	344.4	352.5293499	-9.066275025	-1.977931097	-8.570676377	-8.129349933
16	103	375.2	369.4960399	-1.079379474	0.035545939	-2.072745523	5.703960129
17	102	356.1	360.2343075	-3.908203197	9.522266565	-7.802810019	-4.134307491
18	99	352.59	347.1555002	-6.390124327	7.609335081	-5.426734362	5.434499809
19	101	401	407.4385669	-0.288938238	1.985267234	-7.736426616	-6.43856693
20	104	387.9	396.586376	-0.592622092	5.017091992	-6.897634651	-8.686376031
21	90	446.8	447.9929025	8.900914193	-27.94367871	26.28518595	-1.192902526
22	105	406.4	405.0228454	-1.45050985	3.617349315	-3.449408803	1.377154568
23	116	393.5	402.3117792	-0.041155486	-5.076022025	2.380155435	-8.811779239
24	84	402	409.8617955	-2.72822742	-0.341160426	4.830944795	-7.861795478

Table (3.6) Error Covariance parameters using T-SEDLM

Hours	Convergence Time	Error Covariance Load	Error Covariance Trend	Error Covariance S	Error Covariance S*
1	105	17835.92559	2682.166492	4335.103805	4084.776054
2	111	17076.46007	2658.03398	4215.279336	3960.426242
3	95	16391.25246	2635.789021	4105.866762	3846.692848
4	97	15485.85634	2605.984537	4038.8496	3756.479878
5	99	14505.22191	2572.03312	3797.842155	3525.424039
6	98	14815.23259	2584.180353	4175.851696	3841.552176
7	91	16106.63122	2628.328635	4553.355629	4196.000868
8	97	16407.41181	2636.68783	4204.687481	3929.99231
9	88	18179.86304	2695.1866	5055.526595	4688.825959
10	98	19106.11473	2723.147363	5083.463191	4748.406996
11	116	19729.95467	2742.206216	5283.017504	4961.232552
12	105	19444.02497	2732.812281	4974.06204	4673.453855
13	98	19097.48613	2723.212442	5172.083648	4845.973751
14	95	18883.1114	2716.700737	5134.241796	4806.621085
15	96	18438.42497	2703.070493	5055.346168	4724.515299
16	103	18412.12135	2702.596068	5148.98008	4803.773677
17	102	18714.20229	2711.717418	5160.789539	4813.487926
18	99	18960.4976	2719.228879	5204.808139	4859.290007
19	101	18805.58786	2714.336538	5120.526624	4792.353763
20	104	19148.45485	2725.08598	5282.013071	4942.30016
21	90	18707.42007	2711.670115	5202.512799	4859.545058

Hours	Convergence Time	Error Covariance Load	Error Covariance Trend	Error Covariance S	Error Covariance S*
22	105	18150.05326	2692.326289	4480.826974	4219.054122
23	116	17828.15201	2682.25698	4429.338491	4165.638163
24	84	18496.51297	2706.478993	5538.917153	5137.652397

Tables (3.5) and (3.6) are show that:

The Error Covariance to electricity load parameter is convergence at time point (t=87) with the value (17835.92559), the Error Covariance for trend parameter is convergence at time point (t=83) with the value (2682.166492), The Error Covariance to S parameter is convergence at time point (t=105) with the value (4335.103805) and the Error Covariance for S* parameter is convergence at time point (t=102) with the value (4084.776054). There are four time convergences and the last time convergence is the best estimate for the same time.

A compression between electricity load resulting from the estimate and actual electricity load values for hour (one) in T-SEDLM is shown in Figure 3.3. The results show how closely the estimate model matches the actual electricity load. The blue line is actual values and the red line is estimate values.

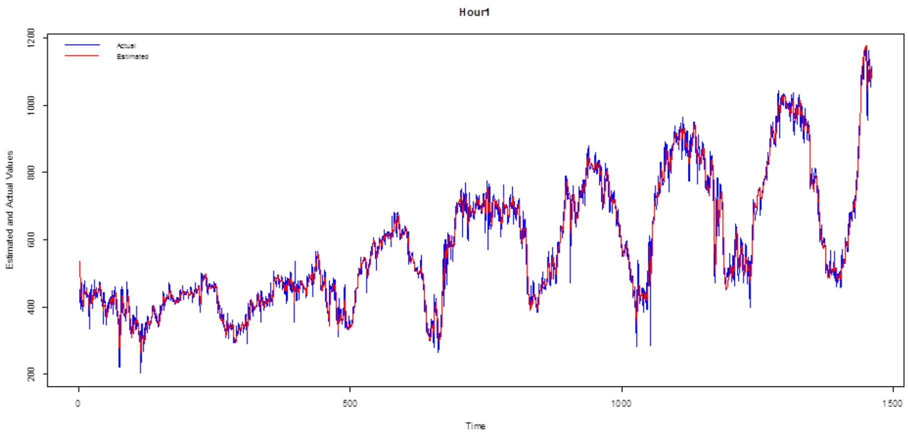

Figure (3.3) Difference between actual values and estimate values for hour one

In the second hour (two) of the day the Error Covariance to electricity load parameter is convergence at time point (t=91) with the value (17076.46007), the Error Covariance for trend parameter is convergence at time point (t=79) with the value (2658.03398), The Error Covariance to S parameter is convergence at time point (t=102) with the value (4215.279336) and the Error Covariance for S* parameter is convergence at time point (t=111) with the value (3960.426242). There are four time convergences and the last time convergence is the best estimate for the same time.

In the last hour (twenty-four) of the day the Error Covariance to electricity load parameter is convergence at time point (t=68) with the value (18496.51297), the Error Covariance for trend parameter is convergence at time point (t=82) with the value (2706.478993), The Error Covariance to S parameter is convergence at time point (t=84) with the value (5538.917153) and the Error Covariance for S* parameter is convergence at time point (t=83) with the value (5137.652397). There are four time convergences and the last time convergence is the best estimate for the same time.

Case Four: The electricity load at instant of a time is expressed as a linear combination of previous load and weather data with time varying coefficients. RDLM of equations (3.15) and (3.16) which contains the weather factors are used to estimate parameters, Kalman Gain and Error Covariance for univariate model for each twenty four hours.

Where: X_1 is the previous load values.

X_2 is the temperature values.

X_3 is the rate of humidity values.

X_4 is the wind speed values.

Table 3.19 shows estimated electricity load parameter and actual values with estimated coefficient weather variables for twenty-four hours. The results show that the estimated electricity load is very close to the actual electricity load values, which means that the model assumed with weather factors is accurate enough to present the electricity load.

Table (3.7) Estimate load parameters using RDLM

Hours	Bo	B1	B2	B3	B4	Actual Value	Estimate	Error Estimate
1	68.65852	0.6496772	1.854102	0.8122782	-17.55946	360	363.0809	-3.0809
2	39.97512	0.7186626	1.720259	0.7804937	-9.586071	352.4	340.4137	11.9863
3	45.70766	0.7006409	1.816982	0.8048018	-20.04099	323.6	323.8995	-0.2995
4	43.29189	0.7099722	1.729292	0.7644335	-13.57656	323.5	324.2118	-0.7118
5	40.61952	0.7050792	1.844437	0.8071401	-19.78089	325.2	325.0181	0.1819
6	89.71536	0.5639716	1.692802	0.8769712	-13.23974	352.1	330.6393	21.4607
7	106.5541	0.5212312	1.971742	1.043915	-18.49214	377.7	369.0777	8.6223
8	138.82485	0.4174388	2.469237	1.2091362	-26.68152	368.7	368.4026	0.2974
9	117.22045	0.4424556	2.38424	1.3684174	-19.17112	368	368.9145	-0.9145
10	96.61076	0.4975305	2.614981	1.4688176	-20.45762	360.7	363.814	-3.114
11	112.63918	0.462994	2.679074	1.58474	-27.31971	356.2	353.39	2.81
12	131.54182	0.4697398	1.9747437	1.216986	-18.6053	352.7	352.1601	0.5399
13	140.34896	0.3994521	2.64432	1.2622865	-18.51475	356.7	352.0817	4.6183
14	157.4172	0.3910942	2.41827	1.1921009	-24.47464	349.8	356.8326	-7.0326
15	140.21752	0.4457632	2.17998	1.0853226	-18.75885	365.4	365.1069	0.2931
16	114.16425	0.4997458	2.125466	1.0367658	-11.62889	363	362.7329	0.2671
17	76.93243	0.550724	2.825553	1.1505954	-14.95535	356.1	357.3592	-1.2592
18	90.7044	0.5637701	2.259752	0.9454142	-17.61734	349.4	353.993	-4.593
19	92.28469	0.5762689	2.120332	0.8816061	-16.4601	419.4	419.7626	-0.3626
20	81.58675	0.6407005	1.684693	0.6724413	-18.55384	403.3	405.9698	-2.6698
21	83.94648	0.6573737	1.4153369	0.5752139	-15.74711	407.6	411.6166	-4.0166
22	71.49914	0.7074771	1.3102815	0.5139203	-17.26171	397	399.2465	-2.2465
23	108.03059	0.5953917	1.6992314	0.6240192	-17.66982	396.4	399.6216	-3.2216
24	93.58771	0.6428407	1.4285121	0.5137729	-3.976381	372.2	377.5938	-5.3938

3.3.2. Daily Electricity Load Estimation

Case One: SDLM is used to estimate electricity load parameters, Kalman Gain and Error Covariance.

Table (3.8) presents estimated electricity load parameters for all seven weekdays which consist of estimated electricity load, Kalman gain and error estimate.

Table (3.8) Estimated daily load parameters using SDLM

Days	Convergence Time	Actual Values	Estimate Load	Kalman Gain Load	Error Covariance	Error Estimate
Saturday	15	361.9720	373.986	0.4629	4675.492	-12.0139167
Sunday	15	376.0083	382.135	0.4629	4675.492	-6.1266667
Monday	15	364.8791	374.4596	0.4629	4675.492	-9.5804333
Tuesday	15	376.75	375.6127	0.4629	4675.492	1.1373
Wednesday	15	359.6208	376.2525	0.4629	4675.492	-16.6316667
Thursday	15	343.9125	371.1295	0.4629	4675.492	-27.217
Friday	15	371.9366	371.9588	0.4629	4675.492	-0.0221333

Case Two: TDLM is used to estimate electricity load, trend parameters, Kalman Gain and Error Covariance.
Table (3.9) presents estimated electricity load and trend parameters for all seven weekdays which consist of estimated load, Kalman gain and error estimate.

Table (3.9) Estimated daily load parameters using TDLM

Days	Convergence Time	Actual Value	Estimate load	Estimate Trend	Kalman Gain Load	Kalman Gain Trend	Error Estimate
Saturday	31	416.2208	416.90937	4.63962	0.70459	0.19136	-0.688536667
Sunday	31	404.5625	408.50139	2.35826	0.70329	0.18983	-3.93889
Monday	32	418.4875	420.27426	4.45107	0.70102	0.18719	-1.78676
Tuesday	31	416.9041	415.50969	4.50752	0.70119	0.18739	1.394476667
Wednesday	30	402.2916	404.7106	1.02208	0.70239	0.18878	-2.418933333
Thursday	32	394.3916	398.30331	-0.7523	0.70294	0.18943	-3.911643333
Friday	30	405.7916	407.41502	3.34771	0.70589	0.19288	-1.623353333

Table (3.10) presents Error Covariance for all seven weekdays using TDLM

Days	Convergence Time	Error Covariance Load	Error Covariance Trend
Saturday	31	22906.39527	14838.34848
Sunday	31	23336.1559	14930.0023
Monday	32	24106.08487	15092.1397
Tuesday	31	24044.46925	15079.25945
Wednesday	30	23636.64695	14993.59232
Thursday	32	23451.74487	14954.51093
Friday	30	22489.70008	14748.66639

Case Three: T-SEDLM is used to estimate electricity load, trend, seasonal parameters and Error estimate. In this analysis seasonal component (C=365) due to the daily cycle is used.
Table (3.11) presents estimated electricity load, trend and seasonal parameters and error estimate for all seven weekdays.

Table (3.11) Estimated daily load parameters using T-SEDLM

Days	Convergence Time	Actual Value	Estimate Load	Estimate Trend	Estimate S	Estimate S*	Error Estimate
Saturday	107	695.0625	696.4842474	9.614224817	3.761326402	-6.731204326	-1.421747436
Sunday	104	680.65	679.337293	12.23029252	7.101454824	-0.05178806	1.312706955
Monday	108	694.8	696.3575025	6.502598434	0.100730376	2.505799315	-1.557502479

Days	Convergence Time	Actual Value	Estimate Load	Estimate Trend	Estimate S	Estimate S*	Error Estimate
Tuesday	110	685.375	685.2813412	3.192004745	0.845168037	0.646916414	0.093658835
Wednesday	100	669.3833	667.5007387	20.94932294	-6.002774578	5.544137698	1.882594683
Thursday	99	661.6041	661.2344844	21.04111568	-8.284144861	-8.260281435	0.369682278
Friday	106	688.3333	687.0652458	8.658956762	5.389224107	8.670408696	1.268087554

Table (3.12) presents Error Covariance for all seven weekdays using T-SEDLM

Days	Convergence Time	Error Covariance Load	Error Covariance Trend	Error Covariance S	Error Covariance S*
Saturday	107	17127.48951	2504.968807	4159.546419	3872.207666
Sunday	104	17360.62035	2510.820943	4195.604265	3909.789944
Monday	108	17709.57238	2519.021831	4043.045108	3814.981536
Tuesday	110	17712.87563	2519.363945	4150.89393	3897.201977
Wednesday	100	17622.85023	2518.021031	4500.019685	4162.724123
Thursday	99	17482.02508	2514.310472	4404.710151	4060.062857
Friday	106	16870.03045	2498.226274	4023.433521	3764.454399

Case four: RDLM is used with the same parameters of hourly electricity load estimation to estimate electricity load, weather factor coefficients and Error estimate.

Table (3.13) presents estimated regression and electricity load parameters, for all seven weekdays which consist of error estimate.

Table (3.13) Estimated daily load parameters using RDLM

Days	B0	B1	B2	B3	B4	Actual Value	Estimate	Error Estimate
Saturday	75.60497	0.07412743	7.666765	3.746165	-57.95118	440.6041667	441.371	-0.766833333
Sunday	148.82699	0.2560306	3.6494509	1.811544	-22.260152	441.4333333	443.304	-1.870666667
Monday	38.90046	0.3004703	5.9539222	2.8050402	-20.55628	414.4916667	414.6143	-0.122633333
Tuesday	158.74359	0.1952518	4.1911701	2.1416931	-42.67936	429.775	430.8917	-1.1167
Wednesday	178.986112	0.2974422	2.9040504	1.6873835	-59.16235	432.3375	431.9379	0.3996
Thursday	184.364189	0.2694014	3.2016155	1.6462279	-60.63332	432.5791667	433.1055	-0.526333333
Friday	155.46553	0.223161	3.8318566	2.1767382	-52.90777	406.5416667	406.5939	-0.052233333

4.4 *Evaluating Accuracy (or Error) Measures for Estimation*

In order to select the best model for estimation, three three accuracy criteria which are (Mean Absolute Error, Mean Absolute Personage Error and Root Mean Square Error) are used. They are based on the error estimate which is the difference between the estimated and actual load electrical values. The results are shown in table (3.14).

Table (3.14) using accuracy criteria for each model for 24 hours

Models	MAE	MAPE	RMSE
Simple	10.8147375	2.531374741	13.4268498
Trend	7.61955958	1.849629988	10.0891725
Seasonal	5.6341522	1.568636473	7.05164161
Regression	3.749725	1.034449	6.017486

Table (3.15) using accuracy criteria for each model for 7days

Models	MAE	MAPE	RMSE
Simple	10.38987	2.917227	13.58742
Trend	2.251799	0.556234	2.533024
Seasonal	1.129426	0.165227	1.279957
Regression	0.693571429	0.159168821	0.909389137

The RDLM has the smallest accuracy criteria values compared to other models, due to the seasonal behavior of the electricity load data, the T- SEDLM in each case (hourly and daily) provided accurate estimations and its accuracy criteria is very close to the accuracy criteria of RDLM.

. The results shown in tables (3.14) and (3.15) are illustrating the effectiveness of trend and seasonal factors in increasing the accuracy of the model.

4.5 Forecasting Hourly and Daily Electricity Load

The one hour ahead forecasting is repeated for 24 hours, using the 24 sets of estimated parameters for each hour model, and one day ahead forecasting is repeated for 7 days using 7 sets estimated parameters for each day model.

4.5.1 Forecasting Hourly Electricity Load

Case One: based on estimation in table (3.5) one hour ahead forecasting values of electricity load using SDLM are shown in table (3.16).

Table (3.16) the forecasting and actual values for 24 hours

Hours	Actual Values	Forecasting	Error Covariance	Error Forecasting
1	427.26	439.0119	7135.042	-11.7519
2	457.63	445.5144	7135.042	12.1156
3	423.3	440.6219	7135.042	-17.3219
4	435	429.2103	7135.042	5.7897
5	430.3464	432.5572	7135.042	-2.2108
6	443.424	441.9727	7135.042	1.4513
7	454.8677	453.771	7135.042	1.0967
8	424.91	443.4807	7135.042	-18.5707
9	442.5274	434.9009	7135.042	7.6265
10	429.7221	429.9956	7135.042	-0.2735
11	433.628	422.19	7135.042	11.438
12	417.0845	415.6895	7135.042	1.395
13	422.0187	417.678	7135.042	4.3407
14	317.83	418.0306	7135.042	-100.2006
15	333.7	418.5257	7135.042	-84.8257
16	373.2	410.0586	7135.042	-36.8586
17	411.38	411.0182	7135.042	0.3618
18	361.06	433.5536	7135.042	-72.4936
19	360	433.6256	7135.042	-73.6256
20	353.49	416.2096	7135.042	-62.7196
21	368.47	410.6931	7135.042	-42.2231
22	445.97	409.6359	7135.042	36.3341
23	432.95	421.0353	7135.042	11.9147

Hours	Actual Values	Forecasting	Error Covariance	Error Forecasting
24	419.58	427.5384	7135.042	-7.9584

Case Two: based on estimation in table (3.9) one hour ahead forecasting values of electricity load using TDLM are shown in table (3.17).

Table (3.17) the forecasting and actual values for 24 hours

Hours	Actual Values	Forecasting	Error Covariance	Error Forecasting
1	444.21	441.5337	68555.21	2.6763
2	449.9	442.3119	65989.77	7.5881
3	451.71	448.6999	63671.17	3.0101
4	432.32	424.0812	60498.26	8.2388
5	426.08	442.596	57263.87	-16.516
6	461.48	448.1057	57853.26	13.3743
7	436.65	445.146	62061.99	-8.496
8	445.5	444.3762	63599.7	1.1238
9	442.2	397.8635	68932.19	44.3365
10	344.68	444.3682	72219.58	-99.6882
11	438.5	375.4064	74172.38	63.0936
12	458.7	429.1862	73539.71	29.5138
13	418.1	426.4366	72037.33	-8.3366
14	407.46	453.2306	71313.14	-45.7706
15	423.34	426.8198	69810	-3.4798
16	412.2	405.5341	69595.12	6.6659
17	427.37	398.2116	70692.71	29.1584
18	455.02	423.0452	71525.11	31.9748
19	442.4	410.534	71051.19	31.866
20	450.08	448.6087	72084.75	1.4713
21	449.1	439.2752	70593.99	9.8248
22	445.2	440.0095	69497.66	5.1905
23	428.95	381.5224	68410.99	47.4276
24	434.08	457.6673	69394.35	-23.5873

Case Three: based on estimation in table (3.17) one hour a-head forecasting values of electricity load using T-SEDLM are shown in table (3.18).

Table (3.18) the forecasting and actual values for 24 hours

Hours	Actual Values	Forecasting	Error Covariance	Error Forecasting
1	401.1	356.2315	45165.35	44.8685
2	269.1	303.3372	43748.96	-34.2372
3	286.8	309.473	42461.41	-22.673
4	337.4	282.3678	40999.5	55.0322
5	312.3	315.9431	38865.08	-3.6431
6	353.1	319.3269	40519.4	33.7731

Hours	Actual Values	Forecasting	Error Covariance	Error Forecasting
7	393.5	438.9575	43506.59	-45.4575
8	390.7	376.5657	42800.9	14.1343
9	398.6	440.584	47918.04	-41.984
10	390.3	388.6341	49249.64	1.6659
11	369.1	369.5205	50733.58	-0.4205
12	213.9	355.0638	49360.16	-141.1638
13	384.34	356.0848	49537.69	28.2552
14	351.4	329.7912	49130.71	21.6088
15	401.2	362.1713	48283.78	39.0287
16	414.5	371.4253	48549.27	43.0747
17	407.2	358.124	48982.38	49.076
18	382	344.3339	49451.75	37.6661
19	419.4	413.1608	48983.33	6.2392
20	330	398.4077	49957.19	-68.4077
21	457.4	450.5415	49115.13	6.8585
22	375.3	404.7099	46055.89	-29.4099
23	382.3	405.0035	45455.66	-22.7035
24	321.7	405.0992	49916.42	-83.3992

Case Four: based on estimation in table (3.19) one hour a-head forecasting values of electricity load using RDLM are shown in table (3.19).

Table (3.19) the forecasting and actual values for 24 hours

Hours	B0	B1	B2	B3	B4	Actual Value	Forecasting	Error Forecasting
1	68.64568	0.649662	1.853726	0.812329	-17.5609	371.6	369.6726	1.9274
2	40.02004	0.718712	1.721543	0.780366	-9.58295	361.3	363.4596	-2.1596
3	45.70641	0.70064	1.816946	0.804805	-20.0411	321.6	346.5373	-24.9373
4	43.28937	0.70997	1.729216	0.76444	-13.5767	310.5	347.8896	-37.3896
5	40.62023	0.70508	1.844459	0.807138	-19.7808	310.9	347.6026	-36.7026
6	89.68777	0.563949	1.691948	0.877108	-13.2437	369.9	353.0437	16.8563
7	106.6822	0.52138	1.975797	1.04304	-18.44	389.7	395.5597	-5.8597
8	138.9717	0.417578	2.474666	1.207816	-26.6135	414.3	381.2255	33.0745
9	117.3394	0.442574	2.388283	1.367362	-19.1239	403.4	381.5842	21.8158
10	96.75026	0.49763	2.620214	1.467484	-20.3729	403.2	387.0869	16.1131
11	112.9543	0.462941	2.688759	1.582893	-27.1515	339.1	383.9589	-44.8589
12	131.6158	0.469787	1.977302	1.216504	-18.5462	375.2	361.0347	14.1653
13	140.4178	0.399468	2.646448	1.26208	-18.4731	374.2	361.3475	12.8525
14	157.3918	0.391074	2.417382	1.192144	-24.485	357.2	361.4776	-4.2776
15	140.2185	0.445764	2.180011	1.085323	-18.7583	378.2	368.4797	9.7203
16	114.0912	0.499652	2.123375	1.036661	-11.6354	375.2	367.6363	7.5637
17	76.79787	0.550482	2.820754	1.150221	-14.9552	407.2	365.5855	41.6145
18	90.60634	0.563656	2.256688	0.94535	-17.6244	382.5	359.3858	23.1142
19	92.35795	0.576395	2.122619	0.881537	-16.4561	428.5	408.4099	20.0901
20	81.55577	0.640629	1.683778	0.672516	-18.5548	394	400.2481	-6.2481

Hours	B0	B1	B2	B3	B4	Actual Value	Forecasting	Error Forecasting
21	84.00007	0.657469	1.416701	0.575129	-15.7461	387	388.5676	-1.5676
22	71.3968	0.707251	1.307779	0.513923	-17.2572	408	391.8394	16.1606
23	107.9476	0.595217	1.696548	0.624098	-17.6773	393.7	399.3912	-5.6912
24	93.50912	0.642719	1.426044	0.514146	-3.98771	385.6	383.1022	2.4978

4.5.2 Forecasting Daily Electricity Load

Case One: One step (or one day ahead) forecasting electricity load for 7 weekdays is obtained and is based on the estimate values in table (3.20) of SDLM as presented in table (3.20).

Table (3.20) the forecasting and actual values for 7 days

Days	Actual Values	Forecasting	Error Covariance	Error Forecasting
Saturday	358.6416	373.986	4925.492	-15.3443333
Sunday	371.8458	382.135	4925.492	-10.2891667
Monday	365.6541	374.4596	4925.492	-8.8054333
Tuesday	351.2541	375.6127	4925.492	-24.3585333
Wednesday	355.525	376.2525	4925.492	-20.7275
Thursday	356.2166	371.1295	4925.492	-14.9128333
Friday	308.0541	371.9588	4925.492	-63.9046333

Case Two: One step (or one day ahead) forecasting of electricity load for 7 weekdays is obtained and is based on the estimate values in table (3.21) of TDLM as presented in table (3.21).

Table (3.21) one-ahead forecasting and actual values for 7 days

Days	Actual Values	Forecasting	Error Covariance	Error Forecasting
Saturday	406.7333	420.3899	3259.597	-13.65656667
Sunday	419.3041	408.6475	3308.851	10.65666667
Monday	433.5291	422.6902	3403.108	10.83896667
Tuesday	385.475	421.0593	3395.796	-35.5843
Wednesday	410.9916	406.3388	3343.51	4.652866667
Thursday	432.5791	398.3747	3319.59	34.20446667
Friday	410.4291	409.8658	3208.273	0.563366667

Case Three: One step (or one day ahead) forecasting of electricity load for 7 weekdays is obtained and is based on the estimate values in table (3.23) of T-SEDLM as presented in table (3.22).

Table (3.22) the forecasting and actual values for 7 days

Days	Actual Values	Forecasting	Error Covariance	Error Forecasting
Saturday	722.21667	698.4902	42162.1	23.72646667
Sunday	683.625	688.0834	42585.96	-4.4584
Monday	674.55	707.2159	42584.25	-32.6659
Tuesday	662.175	689.2327	42922.21	-27.0577
Wednesday	708.04167	706.2818	43889.46	1.759866667
Thursday	637.3125	688.6118	43389.26	-51.2993
Friday	690.3375	701.6031	41397.73	-11.2656

Case Four: One step (or one day ahead) forecasting electricity load for 7 weekdays is obtained and is based on the estimate values of RDLM as presented, shown in table (3.23).

Table (3.23) the forecasting and actual values for 7 days

Days	B0	B1	B2	B3	B4	Actual Value	Forecasting	Error Forecasting
Saturday	75.45585	0.07411609	7.726745	3.738953	-58.41364	429.3	420.7229	8.5771
Sunday	148.2336	0.2576236	3.7323884	1.7936956	-22.47484	417.4166667	411.0466	6.370066667
Monday	38.84362	0.3007919	5.9982375	2.7946566	-20.76184	407.5979167	414.6143	-7.016383333
Tuesday	158.29728	0.1947682	4.2659921	2.137767	-43.09304	421.3166667	421.8543	-0.537633333
Wednesday	182.384302	0.2862304	3.077299	1.6642943	-60.10725	402.025	408.0742	-6.0492
Thursday	184.562277	0.27047	3.2667803	1.6331713	-61.7196	411.35	401.0022	10.3478
Friday	155.22253	0.2234513	3.8721989	2.1694384	-52.86432	432.9041667	429.311	3.593166667

4.6 *Evaluating Accuracy (or Error) Measures for Forecasting*

The comparison of the models of forecasting electricity load is based on three criteria (MAE, MAPE and RMSE) with respect to the error forecasting table (3.24) and (3.25) which show the values of accuracy criteria for the four relevant models.

Table (3.24) the accuracy of hourly forecasting criteria for four models

Models	MAE	MAPE	RMSE
Simple	26.0374	7.145814	39.63042
Trend	22.6004	5.437278	32.57204
Seasonal	36.4492	11.25187	47.17583
Regression	16.9691	4.641016	21.48871

Table (3.25) the accuracy of daily forecasting criteria for four models

Models	MAE	MAPE	RMSE
Simple	22.6203	6.735649	28.65562
Trend	15.7367	3.82958	20.26857
Seasonal	21.7476	3.256568	27.10728
Regression	6.070193	1.460467	6.763808

In table (3.24) and (3.25) three measurements for each model are used and the results give the best model is regression for one hour and one day ahead forecasting because it has the less error value for each measurements compared to other models.

4.1. Conclusions

In the light of the analysis carried in the previous sections, the following conclusions can be drowning:

1- The Kalman filter algorithm is a well-established technique for estimating parameters of load models that are used in this work as dynamic linear equations which include load signal with uncorrelated Gaussian white noise.

2- The Bayesian method and Kalman algorithm used to estimate the parameters of four load models (SDLM, TDLM, T-SEDLM and RDLM), the result shows that Regression-DLM has best estimation and forecasting results compared to other models using the accuracy criteria. the results showed that the Regression-DLM estimation has (MAE=3.749725, MAPE=1.034449 and RMSE=6.017486), the forecasting has (MAE=16.9691, MAPE=4.641016 and RMSE=21.48871) for hourly electricity load. The Regression-DLM estimation has (MAE=0.693571429, MAPE=0.159168821 and RMSE=0.909389137), the forecasting has (MAE=6.070193, MAPE=1.460467 and RMSE=6.763808) for daily electricity load.

3- One of the most important tasks of estimation DLM parameters is initializing of parameter values, because setting an appropriate initial has a great impact on estimation and forecasting performance.

4- We saw the presence of strong seasonal effects in the load data analyzed in this paper and because the T-SEDLM contains the seasonal factor is produces acceptable estimation and forecasting results.

5- MAE, MAPE and RMSE can be regarded as good accuracy measures for comparing DLM's to parameter estimation and load forecast.

6- Short-term load forecasting is an important issue for ministry of electricity. Long-term and mid-term forecasting (few weeks to several years) has an essential role in load planning but they are inaccurate due to the uncertain nature of forecasting process.

References

1. Anderson & Moore, B.G. (1979) *Optimal Filtering.* Prentice-Hall Publishing, University of Newcastle, New South Wales, Australia.

2. Brockwell & Davis, P., R. (2002) *Introduction to Time Series and Forecasting*, Second Edition, Springer-Verlag New York, United States of America.

3. Brown, & Hwang, R.P. (2012) *Introduction to Random Signals and Applied Kalman Filtering With MATLAB Exercises*. Fourth Edition, John Wiley & Sons, United States of America.

4. Candy, J. (2009) *Bayesian Signal Processing Classical, Modern, and Particle Filtering Methods*. One Edition, John Wiley & Sons, United States of America.

5. Durbin, & Koopman, J.S. (2012) *Time Series Analysis by State Space Methods*. Second Edition Oxford University publishing, UK.

6. Eubank, R. (2006) *A Kalman Filter Primer*. CRC Press Publishing, United States of America.

7. Fahad, & Arbab, M. N. (2014) *Factor Affecting Short Term Load Forecasting*. Journal of Clean Energy Technologies, Vol. 2, No. 4, October 2014.

8. Gentleman, R., Hornik, K., & Parmigiani, G. (2009) *Dynamic Linear Models with R*. Springer Science &Business Media publishing, New York.

9. Grewal, M. & Andrews, A. (2008**)** *Kalman Filtering Theory and Practice Using MATLAB*. Third Edition, John Wiley & Sons Publishing, United States of America.

10. Hong, T. (2010) *Short Term Electric Load Forecasting*. Raleigh, Operations Research PhD, North Carolina united of American.

11. Mawlood, Kurdistan I. (2008) Bayesian Detection and Kalman Filter estimation of signals in White Gaussian noise Process, PhD theses, university of Salahddin-Erbil.

12. Meng, M., Niu, D., & Sun, W. (2011) *Forecasting Monthly Electric Energy Consumption Using Feature Extraction*. Journal Energies 2011, 4, 1495-1507; doi: 10.3390/en4101495.

13. Pole, A., West, M. & Harrison, J. (1994) *Applied Bayesian Forecasting and Time Series Analysis*, Springer Science & Business Media.

14. Prado, R. & West, M. (2010) *Time Series Modeling, Computation, and Inference*. Chapman & Hall, United States of America United States of America.

15. Prandoni, P. & Vetterli, M. (2008) *Signal processing for communications*. CRC Publishing, Italy.

16. Rabinovich, Semyon, G. (2010) *Evaluating Measurement Accuracy a Practical Approach*. Springer Science & Business Media, New York.

17. Rothe, J., Wadhwani, A. & Wadhwani, S. (2009) *Short Term Load Forecasting Using Multi Parameter Regression*. International Journal of Computer Science and Information Security, Vol. 6, No.2.

18. Schutter, J., Geeter, J., Lefobvre, T. & Bruyninckx, H. (1999**)** *Kalman filters a tutorial*. Journal, Vol. 40(4), 538-546.

19. Smith, S. (1999) *Digital Signal Processing*. Second Edition, California Technical Publishing San Diego, California.

20. Taylor, J., & Buizza, R. (2003) *Using Weather Ensemble Predictions in Electricity Demand Forecasting*, International Journal of Forecasting, 2003, Vol. 19, pp. 57-70.

21. Taylor, J., Menezes, L., & McSharry, P. (2006) *Comparison of univariate methods for forecasting electricity demand up to a day ahead*. International Journal of Forecasting 22(1), 1–16.

22. Wang, S. (2006) *Exponential Smoothing for Forecasting and Bayesian Validation of Computer Models*. A Thesis Presented to the Academic Faculty, Industrial and Systems Engineering Georgia Institute of Technology.

23. West, M., & Harrison, J. (1997) *Bayesian Forecasting and Dynamic Models*. Second Edition, Springer-Verlag New York.

24. Willis, H. (2002) *Spatial Electric Load Forecasting*. Second Edition, Marcel Dekker Company, United States of American.

پوختە

مۆدێل (Model) پێکدێت لە سیستەمێکی هاوکێشەیی کە زانیاری پێویستی لە خۆوە دەگرێت بۆ نواندن و شیکارکردنی دیاردەکان (phenomenon) دوای بەدەستهێنانی مۆدێلێکی گونجاو بۆ هەر دیاردەیەک دەتوانرێت ئەو مۆدێلە بەکاربهێنرێت بۆ پرۆسەی پێشبینی کردن لە داهاتوو .

ئامانج لە پێشبینی کردن بۆ داهاتوو بریتیە لە زانینی چۆنیەتی ڕوودانی دیاردەکە لە داهاتوو ، وهەروەها بڕیاردانی گونجاو و دروست و بەکەم کردنەوەی مەترسی لە بڕیار دان دا ،واتە بە مانایەکی تر بڕیاری گونجاو بەندە لەسەر بنیاد نانی مۆدڵیکی دروست وپێشبینی کردنێکی ڕاست و گونجاو بۆ دیاڕدەی ژێر لێکۆلینەوەکە .

توێژەران زۆر گرنگیان داوە بە پرۆسەی پێشبینی کردن لە داهاتوو لە بوارە جیا جیاکانی زانستیدا، بە جۆرەها شێوازی جیا. کاتێک داتای دیاردەی ژێر لێکۆلینەوە پێکهاتوە لە داتایەک کەپەیوەستە بە کات ، ئەو زانیاریانە کە دەست دەکەوێت لە داتاکە بە شێوازێکی فیزیایی(جووڵەکی) دەبێت. لەم حاڵەتانەدا بەکارهێنانی مۆدێلە جێگیرەکان (Fixed Models) گونجاو نیە بۆ وەسف کردنی دیاردەکە و تایبەتمەندیە فیزیاییەکانی سیستەمەکە وە لە هەمان کاتدا ناتوانرێت پشت ببەسترێت بەو پێشبینیەی بۆ داهاتوو کە بەندە بەو مۆدێلیە جێگیرانە. مۆدێلی ڕاستە هێڵی فیزیایی (Dynamic Linear Models) بە گرنگترین جۆرەکانی مۆدێل هەژماردەکرێت بۆ نواندنی ئەو داتایانەی کە پەیوەستن بە کات.

لەم لێکولینەوە زانستیەدا توانرا چوار مۆدێل بنیات بنرێت بە پارامیتەری جیا جیا بۆ هەڵگری وزەی کارەبا (بۆ هەموو بیست و چوار کاتژمێرەکانی ڕۆژ و شەو و حەوت ڕۆژەکانی هەفتە هەر یەکەو بەجیا) بەبەکارهێنانی مۆدێلەکانی ڕاستەهێلی فیزیایی (ڕێگای بیز و کالمن فلتر Bayesian (و Kalman) بۆ خەمڵاندنی پارامیتەری مۆدێلەکان وە پێشبینی کردنی یەک هەنگاو (Forecast) بە مەبەستی بەراوردکرنی مۆدێلەکان بە بەکارهینانی ئەم پێوەرانە(MAE ، MAPE and RMSE)بۆ دەست نیشانکردنی باشترین مۆدێل بۆ پێشبینی هەڵگری وزەی کارەبا (Electricity Load) لە شاری هەولێر .

ئەو دەرەنجامەی کە لەم لێکۆلینەوە زانستیەدا بەدەست کەوتووە بەندە لەسەر ئەو داتا ڕاستەقینەی کە وەرگیراوە لە(بەڕێوەبەرایەتی کۆنترۆڵ وگەیاندنی کارەبای هەولێر وە بەڕێوەبەرایەتی گشتی کشتوکاڵی هەولێر) بە بەکارهێنانی پرۆگرامی (R-Language) وە ئێکسل.

بە پێی ئەم شیکارە زانستیەی کە بۆ داتاکەمان کردووە گەیشتین بە ئەوەی کە باشترین مۆدێل بۆ خەمڵاندن و پێشبینی کردنی پارامیتەرەکانی بارکردنی وزەی کارەبا بۆ 24 کاتژمێر وە هەروەها بۆ 7 ڕۆژەکانی هەفتە مۆدێلی لاری (Regression) بە بەراورد لەگەڵ مۆدێلەکانی تر بەپێی ئەم پێوەرانە(MAE ، MAPE and RMSE).