Advanced Research in Electrical and Electronic Engineering
p-ISSN: 2349-5804; e-ISSN: 2349-5812 Volume 3, Issue 3, April-June, 2016, pp. 236-242

http://www.krishisanskriti.org/Publication.html

# Kalman Filters Coupled with Multi-resolution ANFIS Models for Real-Time Electric Power Load Forecasting

**Poulomi Paul[1] and Vijaya Laxmi[2]**

*[1,2]Department of Electrical and Electronics Engineering, Birla Institute of Technology, Mesra, India*
*E-mail: [1]polopaul10@gamil.com, [2]vlaxmi@bitmesra.ac.in*

**Abstract—***This study explores the application of multi- resolution modeling using wavelet analysis and Adaptive Network based Fuzzy Inference System (ANFIS) in conjunction with Kalman filters, called WANFIS-KF models, for real- time Electric Power Load Forecasting (EPLF) application. The proposed models are used for forecasting electric load for 1- day lead time. WANFIS-KF model utilizes Discrete Wavelet Transform (DWT) to obtain the sub- time series of the original load data at multiple time- frequency resolution, thus providing better representation of the variations in the electric load data. The Kalman filter algorithm mimic the real- time application scenario and hence, updates the model at each time step to provide a reliable forecast. The proposed models are compared with stand- alone ANFIS models to highlight the effectiveness of the proposed WANFIS-KF models in real- time EPLF using Root mean squared error (RMSE), Correlation coefficient (CC) and Mean Absolute Error (MAE) as statistical indices. The results indicate that WANFIS-KF models are better suited for forecasting load data in real- time and are able to capture the peak loads with higher accuracy compared to standalone ANFIS models.*

## 1. INTRODUCTION

The continuous monitoring of the electric load, which is supplied by the electric energy system, is a basic requirement for its reliable operation [1]. The most important planning resources for power system to forecast the future load demand in the present scenario and also one of the essential requirement for the efficient operation and planning of the power system [2]. Load forecasting, with lead times from a few minutes to several months, helps the system operator to schedule spinning reserve allocation efficiently and is crucial for power system security [3]. The EPLF can be classified in terms of the forecasting horizon. EPLF for up to 1 day/week ahead is termed as short term, whereas forecasting from 1 day/week to 1 year ahead is classified as medium-term forecasting, and ELPF with greater than 1 year lead in forecasting horizon is termed as long term forecasting.

Several approaches are used for the purpose of EPLF like regression analysis, autoregressive moving average (ARIMA) [4] and autoregressive distributed-lag models. Application of Artificial intelligence approaches have been relatively new and have been successfully implemented for the purpose of short term EPLF. Machine Learning and Soft Computing techniques have been proven to represent electric consumption uncertainities with very good detail [5]. Al-Hamadi and Soliman [6] used a time-varying weather and load model for the short-term EPLF. Jain and Satish [7] proposed a hybrid technique using Support Vector Machines (SVM) and Artificial Neural Networks (ANN) to forecast the electric load with 1-day lead time. Several other studies have emphasized on the application of ANN for EPLF [3, 8, 9].

However, these models do not address to the inherent non-stationarity in the dataset such as trends and seasonal variations which leads to poor predictability of electric load in real- time applications. Discrete Wavelet Transforms (DWT) provides an excellent tool to counter the stationarity issues in time series modeling by analyzing the time series data in both time and frequency domains, give considerable information about the physical structure of a signal [10]. Wavelet transform has been largely applied for analyzing variations, periodicities and trends in time series [11]. Also, due to inherent non- stationarity of the process, it is imperative to constantly update the model parameters in order to accommodate the short- term fluctuations in the dataset which are otherwise unaccounted in stochastic models. Kalman filters (KF) are very popularly used for recursive updating of the forecasting scheme and hence to obtain reliable forecast in the real- time [1, 2, 6, 12].

This study proposes a hybrid of DWT, ANFIS and KF in order to obtain 1- day lead time electrical power load forecasts. The proposed methodology is compared with Wavelet-ANFIS hybrid models (without recursive updating using KF), and stand- alone ANFIS models. The results establish the usefulness of the proposed approach over the conventional models. The WANFIS- KF models capture the transient shifts and peaks in the dataset with greater accuracy and hence are more suited for application in electrical load management.

## 2. BRIEF DATA DESCRIPTION

This study uses hourly load data from 1st Jan 2009 to 31st Dec 2015 aggregated to daily average values for 1-day lead time forecasting obtained from the Electric Reliability Council of Texas, U.S.A., (ERCOT) archives. The data is available at hourly resolution for eight weather zones. For the purpose of this study, the data from eight climate zones is added to obtain the combined electric load. Data for the years 2009- 2014 is taken as the training dataset for the models and 365 input data points from 1st January 2015 to 31st December 2015 are selected for the model validation. The data can be accessed through the following link:

http://www.ercot.com/gridinfo/load/load_hist/.

**Table 1: Provides the descriptive statistics of the training and the validation dataset.**

| Descriptive statistics of the load data in MW | | |
|---|---|---|
| | **Training data (1st Jan' 09- 31st Dec' 04)** | **Validation data (1st Jan' 15- 31st Dec' 14)** |
| Mean | 844.85 | 835.20 |
| Max | 1287.59 | 1154.55 |
| Min | 602.53 | 632.98 |
| Std. Dev. | 148.21 | 137.43 |

## 3. MATHEMATICAL TOOLS AND TECHNIQUES

### 3.1 Discrete Wavelet Transform

The Wavelet analysis is similar to Fourier analysis. In Fourier analysis, signal is broken into sinusoids of unlimited duration, whereas, in wavelet analysis, wavelets are used instead of the sinusoids. Wavelets have waveforms of limited duration with a mean value of zero. In wavelet analysis, the wavelet is shifted forward in steps along full signal. At each step, correlation of wavelet to the signal is measured. When the full series is covered, a set of wavelet coefficients is generated having same consistency in time as that of original signal. The process is repeated. Thus, sets of wavelet coefficients at different scales are generated and can be used to obtain the high frequency (D1, D2..etc) and low frequency (A1, A2.. etc) sub- time series of the original data corresponding to different time- frequency resolutions. The main advantage of using the wavelet method is its robustness since it does not include any potentially erroneous assumptions or parametric testing procedures. Another advantage of the wavelet method is that wavelet variance decomposition allows one to study different investing behavior in different time scales independently.

For a discrete time series, $x_i$, with integer time steps, DWT in the dyadic decomposition scheme is defined as

$$T_{m,n} = 2^{-m/2} \sum_{i=0}^{N-1} x_i \, \varphi(2^{-m} i - n) \qquad (1)$$

where $T_{m,n}$ is the discreet wavelet coefficient for scale a=$2^m$ and location b =$2^m$ *n, m* and *n* being positive integers; N is the data length of the time series which is an integer power of 2, i.e., N=$2^M$. This gives the ranges of m and n as $0 < n < 2^{M-m} -1$ and $1 < m < M$, respectively. This implies that only one wavelet is needed to cover the time interval producing only one coefficient at the largest scale (i.e., $2^m$ where m=M). At the next scale ($2^{m-1}$), two wavelets would cover the time interval producing two coefficients, and so on till m=1. Thus, the total number of coefficients generated by DWT for a discreet time series of length N = $2^M$ is 1+2+3+…+$2^{m-1}$ = N-1 [13].

The original time series may, then, be reconstructed employing inverse discrete transform, i.e;

$$x_i = \bar{T} + \sum_{m=1}^{M} \sum_{n=0}^{2^{M-m}-1} T_{m,n} 2^{-m/2} \varphi(2^{-m} i - n) \qquad (2)$$

or, in a simple format as:

$$x_i = \bar{T}(t) + \sum_{m=1}^{M} W_m(t) \qquad (3)$$

where $\bar{T}(t)$ is called approximation sub-time series (denoted by $A_m$ in this study) at level m and $W_m(t)$ are details sub-time series (denoted by $D_m$ in this study) at levels $m = 1, 2, \ldots M$.

For a detailed illustration on wavelet analysis, the readers are referred to Mallat [14].

### 3.2 Adaptive network based fuzzy inference system

ANFIS is a data-driven modeling technique which combines human knowledge and reasoning ability of fuzzy inference system (FIS) with the adapting capability of ANN. Systems which use the theory of employing fuzzy sets to classes of unclear, imprecise and incomplete information using linguistic labels stipulated by membership functions are referred to as FIS. Linguistic terms and structure of if-then rules make fuzzy an easily understandable technique, but lacks the ability to deal with changing external environments.

The degree of agreement of a fuzzy set is represented by a membership function, varying from 0 to 1. The final membership functions obtained after each rule's output differ in value yet, are similar in shape to initial membership functions. The artificial intelligence is achieved by specifying category of input variables, such as 'low', 'medium' and 'high'. Optimal number of categories is chosen through comparison, whichever provides the best result. Jang, Sun and Mizutani [15] provide a good illustration of the working of ANFIS. In this work, a five-layer ANFIS structure is implemented where the first layer executes a fuzzification process, the second layer executes the fuzzy AND of the antecedent part of the fuzzy rules, the third layer normalizes the membership functions, the fourth layer executes the consequent part of the fuzzy rules and the last layer computes the output of the system by summing up the outputs of the fourth layer (Fig. 1).

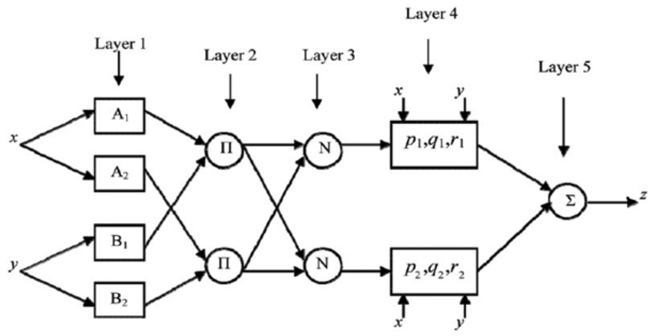


**Fig. 1: five-layer ANFIS structure**

### 3.3 Kalman Filter

In this section, the equation for the development of the basic recursive discrete Kalman filter have been addressed. Given the discrete state equations:

$$x(k+1) = A(k)x(k) + w(k) \tag{4}$$

$$z(k) = C(k)x(k) + v(k) \tag{5}$$

where x(k) is n x 1 system states, A(k) is n x n time-varying state transition matrix, z(k) is m x 1 measurement vector, C(k) is m x n time-varying output matrix, w(k) is n x 1 system error and v(k) is m x 1 measurement error.

The noise vectors w(k) and v(k) are uncorrected white noises that have:

Zero means:

$$E[w(k)] = E[v(k)] = 0 \tag{6}$$

No time correlation:

$$E[w(i)w^T(j)] = E[v(i)v^T(j)] = 0 \tag{7}$$

For i = j

Known covariance matrices (noise levels):

$$E[w(k)w^T(k)] = Q_1 \tag{8}$$

$$E[v(k)v^T(k)] = Q_2 \tag{9}$$

where $Q_1$ and $Q_2$ are positive semi-definite and positive definite matrices, respectively. The basic discrete-time Kalman filter algorithm is given by the following set of recursive equations. Given as priori estimates of the state vector x^(0)=$x_0$^ and its error covariance matrix, P(0)=$P_0$, set k=0 then recursively computer:

Kalman gain :

$$K(k) = [A(k)P(k)C^T(k)]/[C(k)P(k)C^T(k) + Q_2] \tag{10}$$

New State Estimate :

$$x\hat{}(k+1) = A(k)x\hat{}(k) + K(k)[z(k) - C(k)x\hat{}(k)] \tag{11}$$

Error covariance update:

$$P(k+1) = A(k) - K(k)C(k)]p(k)[A(k) - K(k)C(k)]^T + K(k)Q_2K^T(k) \tag{12}$$

An intelligent choice of the priori estimate of the state $x_0$^ and its covariance error $P_0$ enhances the convergence characteristics of the Kalman filter. Few samples of the output waveform z(k) can be used to get a weighted least-squares as an initial values for $x_0$^ and $P_0$:

$$x_0\hat{} = H^T Q_2{}^{-1} H]^{-1} H^T Q_2{}^{-1} z_0 \tag{13}$$

$$P_0 = [H^T Q_2{}^{-1} H]^{-1} \tag{14}$$

Where $z_0$ is $mm_1$ x 1 vector of $m_1$ measured samples and H is $mm_1$ x n matrix.

$$z_\circ = \begin{bmatrix} z(1) \\ z(2) \\ . \\ z(m_1) \end{bmatrix} \quad (15) \quad \text{And,} \quad H = \begin{bmatrix} C(1) \\ C(2) \\ . \\ C(m_1) \end{bmatrix} \tag{16}$$

## 4. MODEL DEVELOPMENT

Wavelet coupled bootstrap- Artificial Neural Network (WBNN) models are developed in this study to evaluate its effectiveness in EPLF. DWT is applied on the input time series (training and validation) of the load data to obtain the wavelet sub- time series of the input dataset. Selection of suitable level of decomposition of the input data is crucial in a wavelet based model.

### 4.1 Selection of decomposition level of decomposition for obtaining the wavelet sub- time series

Selection of the suitable level of decomposition to obtain the wavelet sub- time series of the input data for the models is crucial for explaining the overall variability in the data. Each wavelet sub-time series captures different realizations of the load data in time- frequency domain.

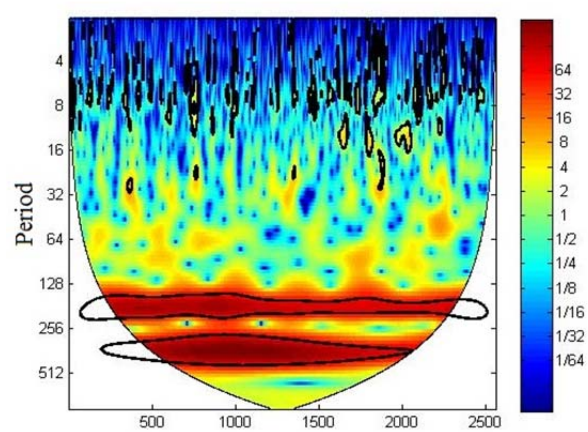


**Fig. 2: Wavelet Power Spectrum (WPS) of the observed total dataset (Training+ Validation period). Significant occurrences can be observed at 128- 512 period interval corresponding to the seasonal and annual variations in the dataset.**

Fig. 2 provides the wavelet power spectrum (WPS) [16] of the total model data (training+ validation). The WPS indicates significant occurrences in the 128 to 256 and 256- 512 day period owning to annual and seasonal variation in the dataset. Hence, the dataset is decomposed to 9$^{th}$ dyadic scale ,corresponding to 256- 512 day period, to obtain the high and low frequency wavelet sub- time series corresponding the to the 9$^{th}$ level of wavelet decomposition (A9, D9, D8, D7..D1) for training and validation period. Fig. 3 shows the wavelet decomposition of the total input dataset to nine level of decomposition.

### 4.2 Selection of suitable wavelet for decomposition of the data

Selection of a suitable mother wavelet has significant impact on the performance of a wavelet based model. Effective representation of the variation in the dataset with in a given constraint of the level of decomposition is affected by the mother wavelet selected for applying the DTW. A *Daubechies* class wavelet with vanishing moment of 45 (*db45*) is selected for this study to carry out the wavelet decomposition of the input dataset for the models. The selection of wavelet in this study is in line with Sehgal et. al. in [17] which highlights that the wavelets with a high vanishing moment are more suited for effective representation of a signal in the form of wavelet sub- time series for a wavelet based time series model.

### 4.3 WANFIS-KF, WANFIS and ANFIS models

Once the input dataset is decomposed into its wavelet sub-time series, the decomposed data is arranged in lagged form till seven antecedent lads in order to relate current load information with the load data for past seven days. The selection of suitable number of lags is carried out by observing the Sample Partial Autocorrelation plots of the observed total dataset as provided in Fig. 4. Seven lags are found to be sufficient to explain the variation in the data. This corresponds to the weekly cycles in electricity usage. Separate ANFIS models are developed for each wavelet sub- time series to obtain the wavelet sub- components of the predicted electric load at 1- day lead. These predicted components are later added to obtain the predicted load at original time- frequency resolution. These models are called WANFIS models in this study. Kalman filters are used to update the model parameters with each forecasting step and hence provide an updated forecast compared to the WANFIS model output. As these models are hybrid of DWT, ANFIS and KF, they are called WANFIS- KF models. For the purpose of comparison with these two models, stand- alone ANFIS models are also developed with original dataset (DWT is not applied). A schematic for the proposed modeling scheme is provided in Fig. 5.

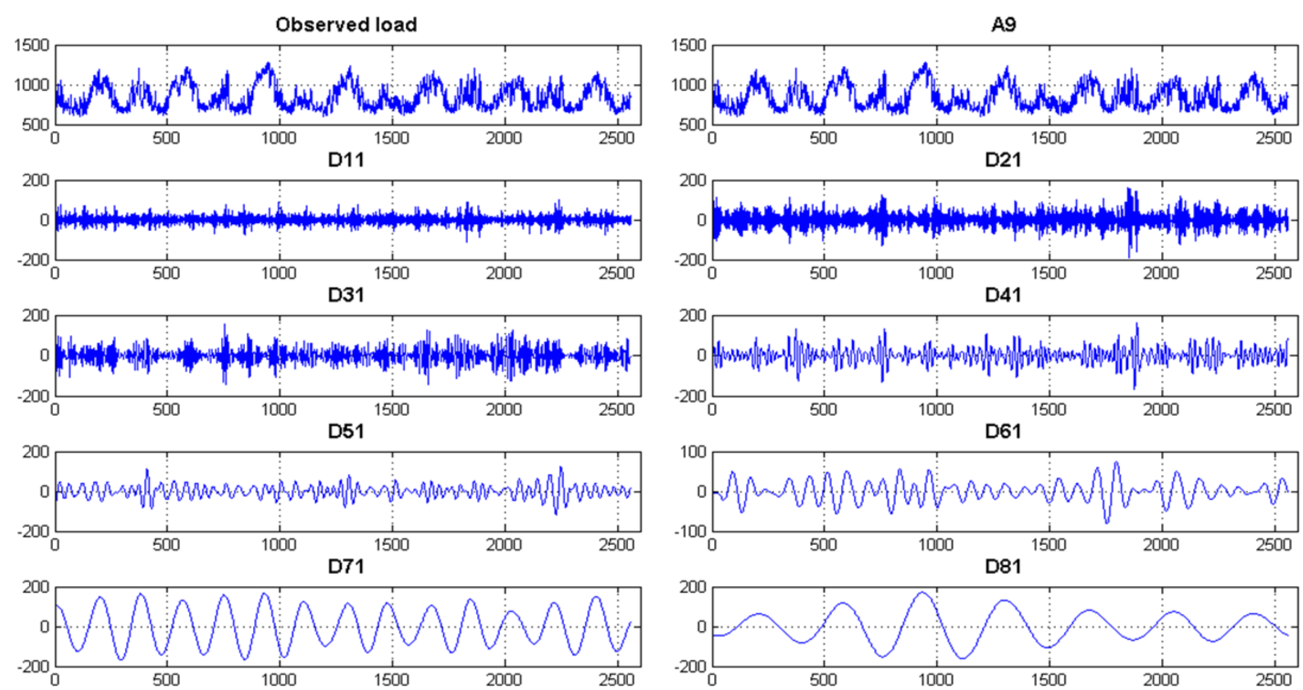


**Fig. 3: Wavelet decomposition of the total input dataset to nine level of decomposition.**

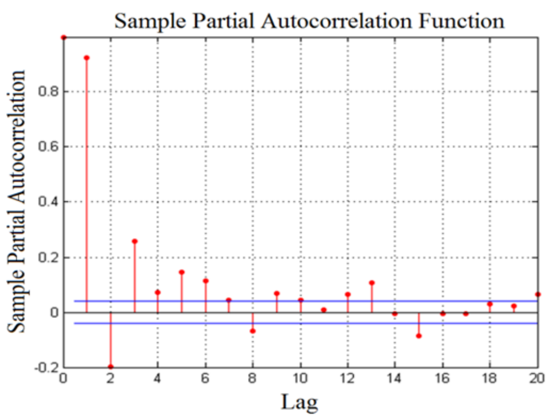


**Fig. 4: Sample Partial autocorrelation of the observed total dataset. Seven antecedent lags are found to be significant for the model development.**

## 5. MODEL PERFORMANCE INDICES

For the performance evaluation of the WANFIS-KF, WANFIS and ANFIS models in forecasting 1- day lead load data for the validation period, three statistical indices namely Root Mean Square Error (RMSE), Correlation Coefficient (CC) and Mean Absolute Error (MAE) which are defined as follows:

(i) ***Root Mean Square Error (NRMSE)*** is expressed as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2} \quad (17)$$

where $O_i$ and $Pi$ are the observed and estimated load and $n$ is the number of data points in the validation dataset.

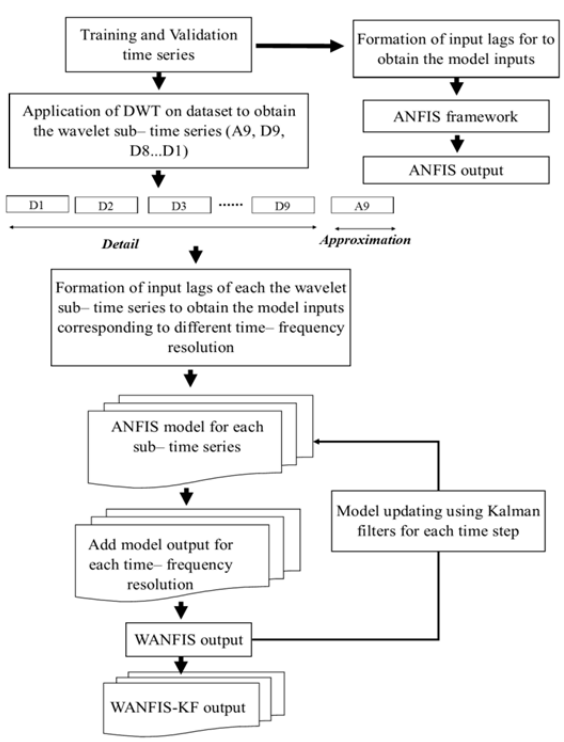


**Fig. 5**

(ii) ***Correlation coefficient (CC)*** is defined as:

$$CC = \left( \frac{\sum_{i=0}^{n}\left(\text{Oi} - \bar{O}\right)\left(\text{Pi} - \bar{P}\right)}{\sqrt{\sum_{i=0}^{n}\left(\text{Oi} - \bar{O}\right)^2}\sqrt{\sum_{i=0}^{n}\left(\text{Pi} - \bar{P}\right)^2}} \right) \quad (18)$$

where, $\overline{Pi}$ is the mean of the estimated load time series for the validation dataset.

(iii) ***Mean absolute error*** **(MAE)** is expressed as

$$MAE = 1/n\left(\sum_{i=1}^{n}|Oi - Pi|\right) \quad (19)$$

## 6. RESULTS AND DISCUSSION

Comparison between the performances of the three models for the validation period is summarized in Table 2. It can be observed that the wavelet based models namely WANFIS- KF and WANFIS perform better than stand-alone ANFIS models in terms of all three statistical indices. However, the WANFIS-KF models are better in capturing the transient fluctuations in the electric load and hence outperform WANFIS models in all three performance indices. The WANFIS- KF model gives an RMSE of 13.48 MW compared to 37.08 MW and 51.32 MW obtained from WANFIS and ANFIS models. The CC and MAE for the WANFIS- KF models is observed to be 0.99 and 10.26 KW respectively. CC and MAE observed from the WANFIS and ANFIS models respectively are 0.97 and 31 MW; and 0.93 and 37.42 MW. Fig. 6 provides a comparison between the observed and model outputs from WANFIS- KF, WANFIS and ANFIS models using line and scatter plots.

The wavelet based models provide information to the modeling system about the variations in the electric load data at multiple time- frequency resolutions. Hence model for each sub- time series is able to adapt to the required complexity and the combined approach of wavelet analysis with ANFIS (WANFIS) provides better accuracy compared to stand- alone ANFIS models. However, the load forecasting system should be able to adapt to the fluctuations in the power demand very quickly to accommodate unexpected transient fluctuations in electrical load demand. The proposed WANFIS- KF approach is able to rapidly adapt to the changing electrical demand scenarios and hence is able to provide reliable forecast for electrical load.

**Table 1: Performance comparison of WANFIS-KF, WANFIS and ANFIS models for the validation period (1st January 2015 to 31st December 2015)**

| Performance of models for validation period | | | |
|---|---|---|---|
| | WANFIS-KF | WANFIS | ANFIS |
| RMSE (MW) | 13.48 | 37.08 | 51.32 |
| CC | 0.99 | 0.97 | 0.93 |
| MAE (MW) | 10.26 | 31.00 | 37.42 |

## 7. CONCLUSION

This paper provides application of a real- time electrical power load forecasting methodology using Wavelet analysis, ANFIS and Kalman filters hybrid approach, called WANFIS- KF. The proposed approach is compared with wavelet- ANFIS hybrid models (without real- time updating using Kamlan filters) and stand- alone ANFIS models. The models are applied for 1- day lead electrical load forecasting for calendar days of the year 2015. The results from the three models are compared using three statistical indices RMSE, CC and MAE. From the results, it is evident that the WANFIS- KF models are better suited for real- time applications as the models capture the transient changes in the load accurately thus providing a reliable forecast compared to the other two models explored in this study.

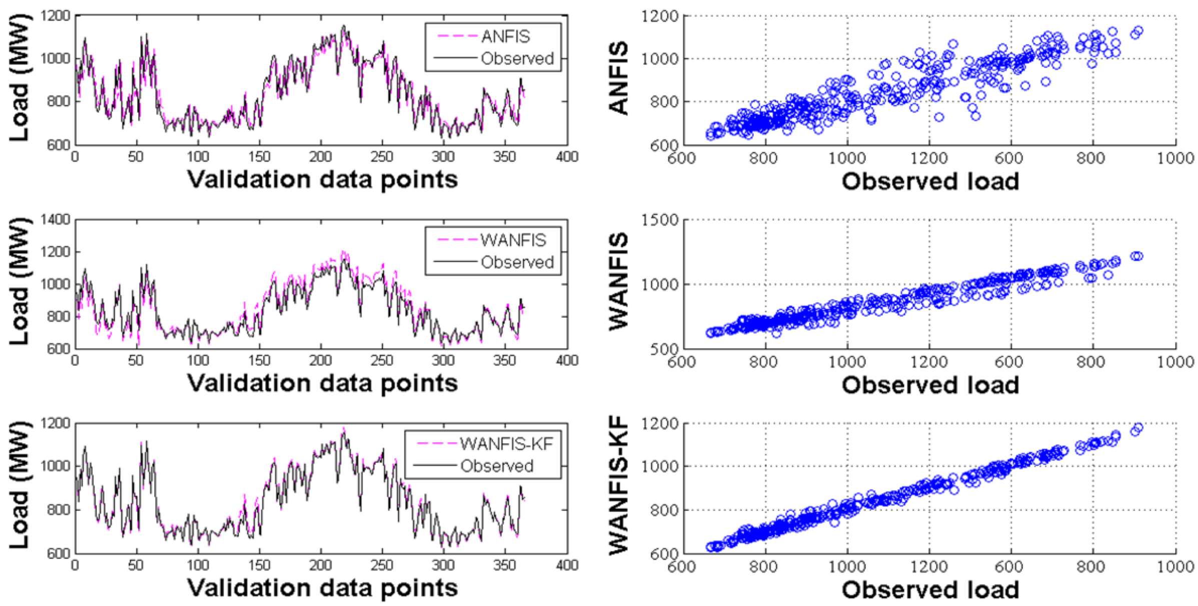


**Fig. 6: Line and Scatter plots for (a) WANFIS- KF (b) WANFIS (c) ANFIS models**

## REFERENCES


[1] Markoulakis, S., Stavrakakis, G., and Nikolaou, T.: ‘Short-term load forecasting based on the Kalman filter and the neural-fuzzy network (ANFIS)’, WSEAS TRANSACTIONS ON POWER SYSTEMS, 2006, 1, (5), pp. 924.

[2] Shankar, R., Chatterjee, K., and Chatterjee, T.: ‘A very short-term load forecasting using Kalman filter for load frequency control with economic load dispatch’, Journal of Engineering Science and Technology Review, 2012, 5, (1), pp. 97-103.

[3] Metaxiotis, K., Kagiannas, A., Askounis, D., and Psarras, J.: ‘Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher’, Energy Conversion and Management, 2003, 44, (9), pp. 1525-1534.

[4] Contreras, J., Espinola, R., Nogales, F.J., and Conejo, A.J.: ‘ARIMA models to predict next-day electricity prices’, Power Systems, IEEE Transactions on, 2003, 18, (3), pp. 1014-1020.

[5] Jurado, S., Peralta, J., Nebot, A., Mugica, F., and Cortez, P.: ‘Short-term electric load forecasting using computational intelligence methods’, in Editor (Ed.)^(Eds.): ‘Book Short-term electric load forecasting using computational intelligence methods’ (IEEE, 2013, edn.), pp. 1-8.

[6] Al-Hamadi, H., and Soliman, S.: ‘Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model’, Electric power systems research, 2004, 68, (1), pp. 47-59.

[7] Jain, A., and Satish, B.: ‘Integrated approach for short term load forecasting using SVM and ANN’, in Editor (Ed.)^(Eds.): ‘Book Integrated approach for short term load forecasting using SVM and ANN’ (IEEE, 2008, edn.), pp. 1-6.

[8] Živčák, D.: ‘Electricity load forecasting using ann’, in Editor (Ed.)^(Eds.): ‘Book Electricity load forecasting using ann’ (Citeseer, 2001, edn.), pp.

[9] Bhattacharyya, S.C., and Thanh, L.T.: ‘Short-term electric load forecasting using an artificial neural network: case of Northern Vietnam’, International journal of energy research, 2004, 28, (5), pp. 463-472.

[10] Daubechies, I.: ‘The wavelet transform, time-frequency localization and signal analysis’, Information Theory, IEEE Transactions on, 1990, 36, (5), pp. 961-1005.

[11] Nalley, D., Adamowski, J., and Khalil, B.: ‘Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008)’, Journal of Hydrology, 2012, 475, pp. 204-228.

[12] Guan, C., Luh, P.B., Michel, L.D., and Chi, Z.: ‘Hybrid Kalman filters for very short-term load forecasting and prediction interval estimation’, Power Systems, IEEE Transactions on, 2013, 28, (4), pp. 3806-3817.

[13] Addison, P.S.: ‘Wavelet transforms and the ECG: a review’, Physiological measurement, 2005, 26, (5), pp. R155.

[14] Mallat, S.: ‘A wavelet tour of signal processing’ (Academic press, 1999. 1999).

[15] Jang, J.-S.R., Sun, C.-T., and Mizutani, E.: 'Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence', 1997.

[16] Grinsted, A., Moore, J.C., and Jevrejeva, S.: 'Application of the cross wavelet transform and wavelet coherence to geophysical time series', Nonlinear processes in geophysics, 2004, 11, (5/6), pp. 561-566.

[17] Sehgal, V., Sahay, R.R., and Chatterjee, C.: 'Effect of utilization of discrete wavelet components on flood forecasting performance of wavelet based ANFIS models', Water resources management, 2014, 28, (6), pp. 1733-1749.