Short-term Electric Power Load forecasting using Wavelet based Bootstrap-Neural Network hybrid models

Poulomi Paul, Dr. Parshuram Thakura

Abstract— **Accurate Electric Power Load Forecasting (EPLF) is a key issue in planning and operation of the electricity industry. However, uncertainty related to the choice of the model for load forecasting and estimation of the model parameters significantly affects the load forecasts. Also, most of the time series forecasting approaches do not address to the non-stationarity of the process. In order to counter these two key issues, this study proposes a hybrid approach using Discrete Wavelet Transform (DWT) coupled with Bootstrap resampling and Artificial Neural Networks (ANN), called WBNN, for electric load forecasting. For the purpose of comparison with the proposed approach, stand- alone Artificial Neural Network (ANN) and Auto- regression (AR) models are also developed. The models are developed and evaluated for their forecasting ability of 1- day lead time electric load. The results indicate that the WBNN models outperform ANN and AR models and are better suited for real- time application in the industry for EPLF.**

***Index Terms*—ANN, Bootstrap resampling, Electric load forecasting, Wavelet analysis**

I. INTRODUCTION

Electric Power Load Forecasting (EPLF) is crucial to the utilities and power industry by providing the estimate for the required future expansion of generation, transmission and distribution of electric energy [1]. 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 [2]. 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) [3] and autoregressive distributed-lag models. Application of Artificial intelligence approaches have been relatively new and has 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 [1]. Al-Hamadi and Soliman [4] used a time-varying weather and load model for the short-term EPLF. Jain and Satish [5] 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 [2, 6, 7].

However, the point forecasts made using these data driven models are subjected to change by a change in the length of the training dataset, parameters and the structure of the selected model. Also, 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. Wavelet transforms provides an excellent tool to counter the stationarity issues in time series modeling by analyzing the time series data in multiple time and frequency domains, providing a better analysis of the information about the physical structure of a signal [8].Discrete Wavelet Transform (DWT) has been largely applied for analyzing variations, periodicities and trends in time series [9]. The bootstrap resampling method has been used in several studies to generate different realizations of the time series to develop ensemble forecasting models [10, 11].

This study demonstrates the application of a hybrid approach using DWT, bootstrap resampling and ANN for 1- day lead time EPLF. 200 realizations of the observed load time series are obtained using the bootstrapping approach. DWT is used to obtain the sub- time series of each resampled time series which is later fed to ANN framework to obtain forecasts for 1-day lead time. Forecasts obtained from models using the resampled data are used to obtain the 95% confidence band of the forecast, thus providing a reliable increment over standard point forecasts using stand-alone time series models. The results from the proposed WBNN models are compared with that obtained from ANN and AR models to demonstrate the effectiveness of the proposed approach.

II. BRIEF DATA DESCRIPTION

The data for this study is taken from the Electric Reliability Council of Texas, U.S.A., (ERCOT) archives and provides hourly load data for its eight weather zones. 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 for the combined electric load for the eight zones. 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.

Table 1: Descriptive statistics of the training and validation data used in this study

Manuscript received May 25, 2016.
Poulomi Paul, Department of Electrical and Electronics Engineering, Birla Institute of Technology, Mesra, India
Dr. Parshuram Thakura, Department of Electrical and Electronics Engineering, Birla Institute of Technology, Mesra, India

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

III. Mathematical tools and techniques

A. Discrete Wavelet Transform

Wavelet analysis is similar to the Fourier analysis. In Fourier analysis, a 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:

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

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 $1+2+3+\ldots+2^{m-1} = N-1$ [12].

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

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

or, in a simple format as:

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

where $\overline{T}(t)$ 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 [13].

B. Artificial Neural Networks

Artificial neural networks (ANNs) are information processing systems composed of simple processing elements (nodes) linked by weighted synaptic connections [14]. The multilayer feed-forward neural network is a popular ANN algorithm consisting of a set of sensory units that constitute the input layer, one or more hidden layers of computation nodes and an output layer of computation nodes. The input signal propagates through the network in a forward direction, layer by layer. These neural networks are commonly referred to as multilayer perceptrons. For a detailed explanation of ANNs, interested readers can refer to Haykin and Network [15] and Bishop [16]

C. Bootstrap resampling

Bootstrap is a data-driven simulation method that uses intensive resampling, with replacement, to reduce uncertainties [17]. For a data consisting of random sample D_n of size n drawn from a population of unknown probability distribution P. If $D_n = [(x_1, y_1), (x_2, y_2)\ldots (x_n, y_n)]$, then $d_i\ (x_i, y_i)$ is an independently and identically distributed (i.i.d) realization drawn from P. Let $\hat{P}$ be the empirical distribution function for D_n with mass 1/n on $d_1, d_2, \ldots, d_n$; and let D^* be a random sample of size n taken from i.i.d. with replacement from $\hat{P}$. The set of B bootstrap samples can be represented as $D^1, D^2, \ldots, D^b, \ldots, D^B$, in which B is the total number of bootstrap samples [18, 19]

IV. 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.

A. 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. Figure 1 provides the wavelet power spectrum (WPS) [19] 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 9th level of wavelet decomposition (A9, D9, D8, D7..D1) for training and validation period.

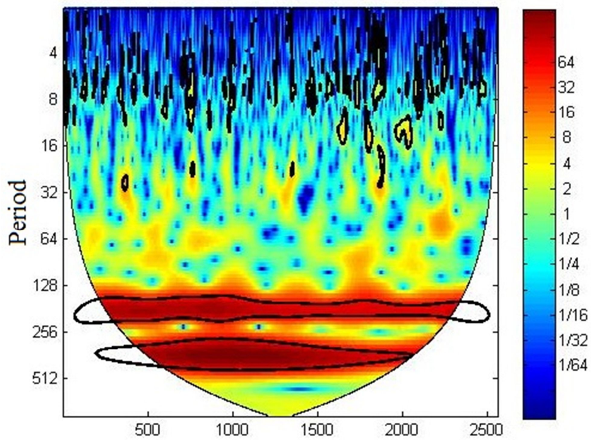

Figure 1: Wavelet Power Spectrum (WPS) of the observed total dataset (Training+ Validation period). Significant occurrences can be observed at 128- 512 period interval corresponding to the seasonal and annual variations in the dataset.

B. Selection of suitable wavelet for decomposition of the data

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 [21] 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.

C. Selection of suitable antecedent lags for model development

Once the input dataset is decomposed into its wavelet sub-time series, the decomposed data is arranged in lagged form till seven antecedent lags 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 Autocorrelation and Sample Partial Autocorrelation plots of the observed total dataset as provided in Figure 2. Seven lags are found to be sufficient to explain the variation in the data owning to the weekly usage cycle.

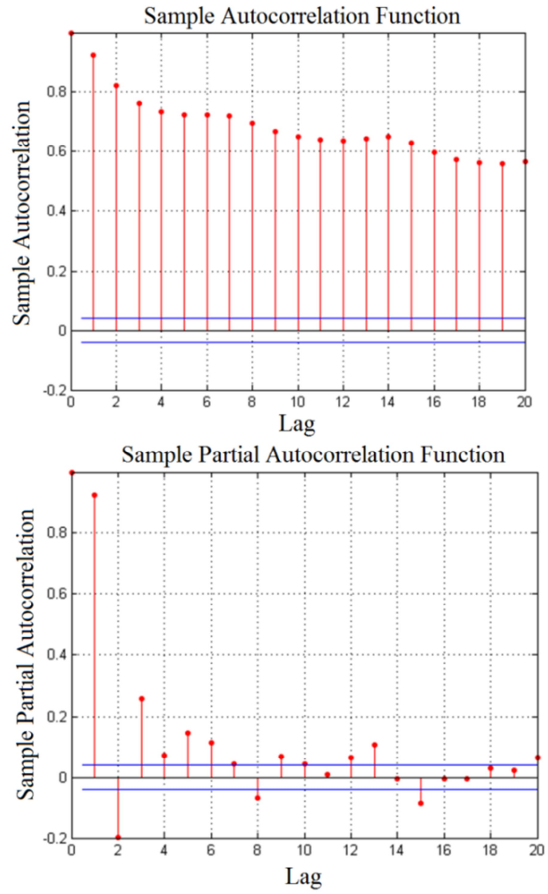

Figure 2: Sample Autocorrelation and Sample Partial autocorrelation of the observed total dataset. Seven antecedent lags are found to be significant for the model development. This corresponds to the weekly cycle of electricity usage.

D. Obtaining Bootstrap resamples and coupling with ANN

Once the decomposed and lagged version of the model input dataset is obtained from the processes explained above, bootstrap resampling is carried out to obtain 200 resamples of the input dataset. Each resample is fed to an ANN model separately and is validated against the observed dataset for the validation period. The output for each model based on the resamples is saved. The average of all model outputs provides the WBNN model output. Since for each data-point in the validation period we have a set of 200 forecasts, a 95% confidence interval (CI) is obtained for the distribution of the forecasts obtained from models based on the 200 resampled inputs. Hence, the proposed WBNN model provides a band forecast along with the point forecast, thus providing a more reliable model out for application in real- time scenarios.

For the purpose of comparison with the proposed approach, stand- alone ANN and AR models with seven antecedent lags of the input dataset are developed and validated. A schematic for the proposed modeling scheme is provided in Figure 3.

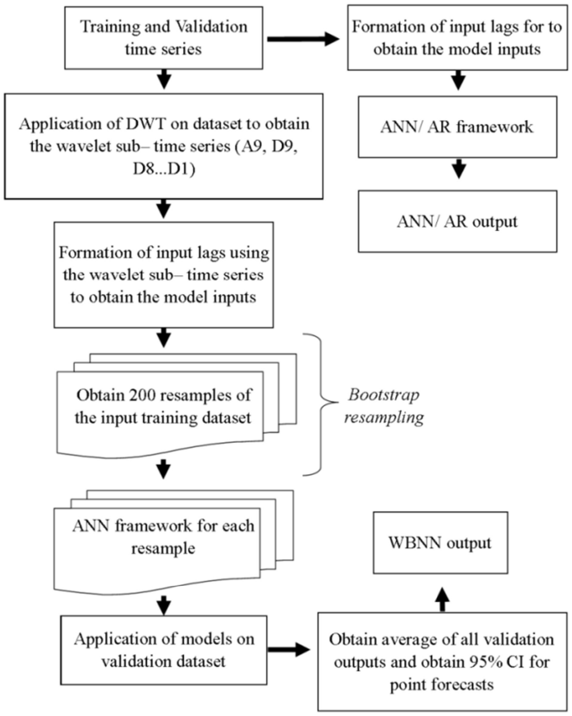

Figure 3: Schematic for the proposed WBNN, ANN and AR modeling scheme

V. MODEL PERFORMANCE INDICES

For the performance evaluation of the WBNN, ANN and AR 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) are used, which are defined as follows:

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

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2} \quad (4)$$

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

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

$$CC = \left(\frac{\sum_{i=0}^{n}\left(\mathrm{Oi} - \bar{O}\right)\left(\mathrm{Pi} - \bar{P}\right)}{\sqrt{\sum_{i=0}^{n}\left(\mathrm{Oi} - \bar{O}\right)^2}\sqrt{\sum_{i=0}^{n}\left(\mathrm{Pi} - \bar{P}\right)^2}} \right) \quad (5)$$

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

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

$$MAE = 1/n(\sum_{i=1}^{n}|Oi - Pi|) \quad (6)$$

VI. 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 outperform the ANN and AR models in terms of all three statistical indices The WBNN modes give and RMSE of 35.10 MW compared to 51.32 MW and 84.03 MW obtained from ANN and AR models. The CC and MAE for the WBNN models is observed to be 0.98 and 29.38 MW respectively. CC and MAE observed from the ANN and AR models respectively are 0.93 and 37.42 MW; and 0.80 and 62.03 MW. Figure 4 provides a comparison between the observed and model outputs from WBNN, ANN and AR models using line and scatter plots. Line plots for the WBNN model outputs contains the 95% CI for the estimation of the point forecast for 1- day lead time. It can be observed that the WBNN models are consistent with the observed validation dataset whereas ANN and AR models give significant deviation from the observations.

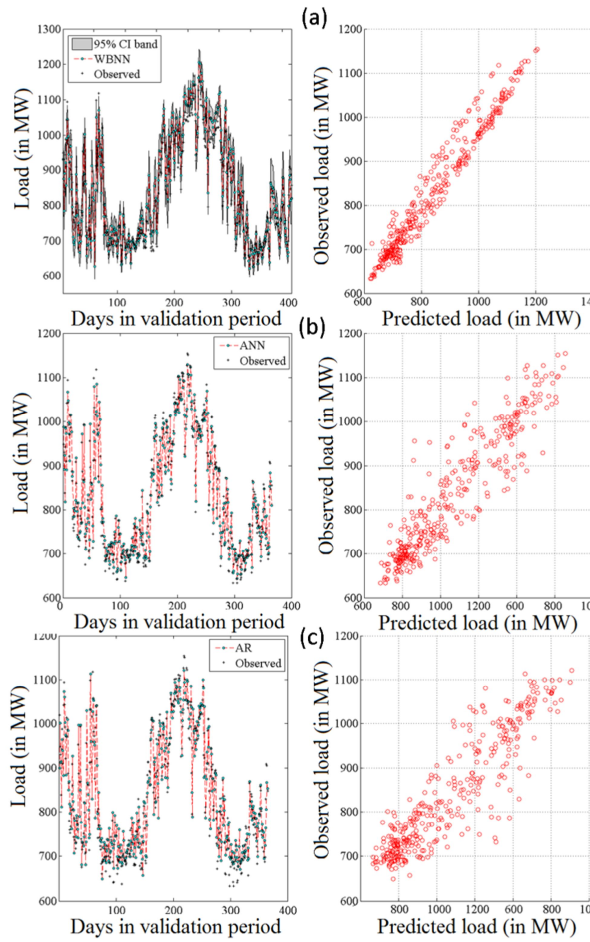

Figure 4: Line and Scatter plots for (a) WBNN with the 95% CI using the model outputs from based on the resampled dataset (b) ANN (c) AR models

The proposed WBNN models employ wavelet analysis for obtaining the wavelet sub- time series at multiple time-frequency resolutions of the input dataset. Thus, the ANN models are able to capture the long and short term variations in the dataset with greater accuracy compared to the stand alone ANN and AR models. The bootstrap resampling provides a useful tool to counter the uncertainty due to insufficient data and provides a band forecast which is considered to be useful in practical application for the decision makers.

Table 2: Performance comparison of WBNN, ANN and AR models for the validation period (1st January 2015 to 31st December 2015)

Performance of models for validation period			
	WBNN	ANN	AR
RMSE (MW)	35.10	51.32	84.03
CC	0.98	0.93	0.80
MAE (MW)	29.38	37.42	62.03

CONCLUSION

This paper provides a comparative analysis of wavelet coupled Bootstrap- Artificial Neural Network models (WBNN) and stand- alone ANN and AR models in 1- day lead time electric power load forecasting. The study reveals that the wavelet based models are able to capture the peaks and low values with greater accuracy thus giving an edge over the stand-alone ANN and AR models. The bootstrapping approach provides band forecast for the model outputs, and thus enhances the reliable of the model outputs. Overall, the combined approach using Discrete Wavelet Transform and Bootstrap resampling coupled approach can be a promising method for electric load forecasting application.

REFERENCES

[1] Jurado, S., Peralta, J., Nebot, A., Mugica, F., and Cortez, P.: 'Short-term electric load forecasting using computational intelligence methods', *Short-term electric load forecasting using computational intelligence methods* (IEEE, 2013, edn.), pp. 1-8.

[2] 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.

[3] 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.

[4] 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.

[5] Jain, A., and Satish, B.: 'Integrated approach for short term load forecasting using SVM and ANN', in *Integrated approach for short term load forecasting using SVM and ANN* (IEEE, 2008, edn.), pp. 1-6.

[6] Živčák, D.: 'Electricity load forecasting using ann', in *Electricity load forecasting using ann*' (Citeseer, 2001, edn.), pp.

[7] 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.

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

[9] 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.

[10] Fan, S., and Hyndman, R.J.: 'Short-term load forecasting based on a semi-parametric additive model', *Power Systems*, IEEE Transactions on, 2012, 27, (1), pp. 134-141.

[11] Bartkiewicz, W.: 'Prediction Intervals for Short-Term Load Forecasting Neuro-Fuzzy Models', *Przegląd Elektrotechniczny*, 2012, 88, (10b), pp. 284-287.

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

[13] Mallat, S.: 'A wavelet tour of signal processing' (*Academic press*, 1999).

[14] Müller, B., Reinhardt, J., and Strickland, M.T.: 'Neural networks: an introduction' (*Springer Science & Business Media*, 2012)

[15] Haykin, S., and Network, N.: 'A comprehensive foundation', *Neural Networks*, 2004, 2.

[16] Bishop, C.M.: 'Neural networks for pattern recognition' (*Oxford university press*, 1995).

[17] Efron, B., and Tibshirani, R.J.: 'An Introduction to the Bootstrap, Monographs on Statistics and Applied Probability, Vol. 57', New York and London: Chapman and Hall/CRC, 1993.

[18] Efron, B., and Tibshirani, R.J.: 'An introduction to the bootstrap' (CRC press, 1994. 1994).

[19] Kumar, S., Tiwari, M.K., Chatterjee, C., and Mishra, A.: 'Reservoir inflow forecasting using ensemble models based on neural networks, wavelet analysis and bootstrap method', *Water Resources Management*, 2015, 29, (13), pp. 4863-4883.

[20] 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.

[21] 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.