4

Short Term Forecasting of Solar Irradiance Using Hybrid of Wavelet Transform and Feedforward Neural Networks

In recent years, more emphasis has been given on developing hybrid/ensemble methods that can improve the forecast accuracy and incorporate the advantages of individual models. Ensemble forecasting methods can be divided into two categories; competitive ensemble and cooperative ensemble forecast [Opitz and Maclin, 1999]. Cooperative ensemble forecasting is further categorized into preprocessing and postprocessing ensemble forecasting. This chapter focuses on the preprocessing ensemble forecasting. Sometimes, the pre-processing divides the input data set into groups of sub-datasets and chooses suitable predictors for each or few sub-datasets based on the characteristics of the sub-series. Final forecast value then obtained by combining outputs of all predictors [Ren *et al.*, 2015]. Many preprocessing based ensemble frameworks are proposed in the literature that uses data processing techniques primarily for short-term solar radiation forecasting.

4.1 INTRODUCTION

In this chapter we implement a hybrid model comprising of Discrete Wavelet Transform (DWT) and Feedforward Neural Network (FFNN) and show that this approach of hybrid model performs better for the underlying solar data. In this process, we measure the difference in forecast error with and without applications of wavelet transform and highlight the improvement in forecast accuracy. We have used half-hourly Solar-GIS GHI data of Rajasthan from the year 1999 to 2001 for analysis purpose. All the data of same month throughout this period is kept together because data of the same month will have similar variability and seasonality. Keeping data of same month for this period can contribute in better learning to neural network. For simplicity we have assumed sun hour in Rajasthan is of 10 hours throughout the year and all the months have 30 days. These assumptions will have no effect on predicted values and it is done for the purpose of simplicity. So during this tenure all the months will have total $3 \times 20 \times 30 = 1800$ data points each.

The remaining chapter is arranged in the following way. Wavelet transform of solar irradiance is explained in Section 4.2 along with its purpose of applications and advantages. In Section 4.3 adopted input parameters for FFNN along with adopted training procedures are discussed. The forecast procedures are discussed in Section 4.4 and chapter is concluded in Section 4.5.

4.2 WAVELET TRANSFORM OF SOLAR IRRADIANCE

Many a times most desired information is hidden in the frequency content of the signal. The working of Fourier transform (FT) lies in projection of data points on sinusoidal basis functions which extends through the span of time domain. FT represents the frequency components that exist in the signal. However, it doesn't give knowledge about at what time these components are located in signal. In other words FT does not give time and frequency related information simultaneously. Although, this information is not required when the signal

is stationary (Stationary signals are those signals whose frequency content does not change with time). To overcome this shortcoming of FT, STFT was introduced which has constant resolution at all times and frequencies. But a good time resolution and poor frequency resolution at higher frequencies or alternatively a good frequency resolution and poor time resolution at lower frequencies is what is required in practice. Later on WT was introduced which has stated time and frequency resolution and works well in practical scenario. The key feature of WT is that it uses short windows at high frequencies and long windows at low frequencies [Polikar, 1996]. The approach is useful for signals that have high frequency components occurring for short duration or low frequency components for longer duration. This is precisely the case for solar radiation as it is highly fluctuating for shorter period of time when cloud is moving over PV panel. The solar radiation data is usually non-stationary and contains various trends, spikes, and seasonality. Wavelet decomposition can be considered as feature management tool that can conveniently separate the underlying spikes and trend in signals from the spurious short fluctuations. Therefore, solar radiation forecasting, using WT improves the forecast error. There are many wavelet families: Daubechies family, Coef, Haar, Symlet etc. Thus, in this chapter, we have investigated practically how WT helps to get better forecast accuracy in the construction of hybrid model for solar radiation forecasting.

The main idea of wavelet analysis is to measure the degree of similarity between the original signal and the basis function of the WT, also called the mother wavelet [Mallat, 1999]. A discrete wavelet transform (DWT) uses successive high pass and low pass filtering operations to decompose the time series. Mallat's algorithm was developed to implement DWT using filters. This algorithm works in two stages decomposition and reconstruction [Mallat, 1989]. Wavelet decomposition involves low pass filter filtering and downsampling and the wavelet reconstruction involves three steps of upsampling and filtering [Mallat, 1989]. In this work, three-level decomposition has been used; consequently three details ($S_{D_3}(t)$, $S_{D_2}(t)$, $S_{D_1}(t)$) and one approximate ($S_{A_1}(t)$) signals are obtained from the original solar irradiance time series data. A wavelet function of type Daubechies of order 7 (DB7)which is more suitable for solar irradiance forecasting as suggested by [Cao and Cao, 2005, 2006; Cao and Lin, 2008], is used as the mother wavelet. The governing equation for three-level DWT of solar irradiance time series data using DB7 mother wavelet is given by

$$[S_{D_3}(t), S_{D_2}(t), S_{D_1}(t), S_{A_1}(t)] = DWT(S(t)), \tag{4.1}$$

where, S(t) is the original time series, $S_{D_3}(t)$, $S_{D_2}(t)$, $S_{D_1}(t)$ are detailed components and $S_{A_1}(t)$ is the approximation component. Each of these decomposed series consequently will be given as inputs to FFNN for prediction. The detailed description of FFNN is presented in Chapter 2. In the next section, we provide the training details of FFNN.

4.3 FEEDFORWARD NEURAL NETWORKS (FFNN)

ANN has become popular method for various renewable energy applications and forecasting [Mellit and Kalogirou, 2008; Kashyap *et al.*, 2015]. In our work, we have used FFNN to predict the decomposed output obtained after application of DWT on the normalized solar irradiance. We use single layer FFNN. Number of hidden nodes are equal to number of inputs, and a single output neuron. In the training variable step back propagation algorithm is used for the training purpose, the momentum factor is fixed on 0.95. The stop conditions are the limitations of training error and maximum training epoch, which are 0.01 and 1000 respectively. The training procedure is as follows

1. Initialize the weights and biases of feedforward neural network in the range of [0,1].

2. Give an input dataset and the expected output vector (target) to FFNN for training.

3. Update each weight and bias for each iteration of the network using the adaptive variable step back propagation algorithm.

4. The steps 2-3 does not stop until any of the stop conditions is reached.

5. Calculate the error using appropriate error metric (RMSE and MAE).

4.4 FORECAST METHODOLOGY OF WAVELET DECOMPOSED GHI USING FFNN

We have followed the steps given below to forecast GHI time series data.

1. As the requirement of the wavelet transform as well as to enhance the adaptability of the neural network, solar irradiance data is normalized in advance in the range[-0.5-0.5]. If $p(t); t = 0,1,2,...$ denotes a data sequence and $q(t); t = 0,1,2,...$ denotes a normalized data sequence, we have:

$$q(t) = \frac{p(t) - min\ p(\tau)}{max\ p(\tau) - min\ p(\tau)} - 0.5 \tag{4.2}$$

2. The low frequency sequence $S_{A_1}(t)$ and the high frequency sequences $S_{D_3}(t)$, $S_{D_2}(t)$ and $S_{D_1}(t)$ of the solar irradiance are obtained by using WT .

3. Each decomposed output data series from DWT ($S_{D_3}(t), S_{D_2}(t), S_{D_1}(t), S_{A_1}(t)$) is divided in the group of 10 data observations. For example suppose $X_{A_1}(t)$ ($X_{A_1}(1)$, $S_{A_1}(2)$,...,$S_{A_1}(1800)$) represents approximation series then first group of data after grouping will contain $S_{A_1}(1)$, $S_{A_1}(2)$,..., $S_{A_1}(10)$, second group of data will contain $S_{A_1}(2)$, $S_{A_1}(3)$,..., $S_{A_1}(11)$ and so on. Similar grouping procedure is followed for $S_{D_3}(t)$, $S_{D_2}(t)$ and $S_{D_1}(t)$ as shown in Figure 4.1. These

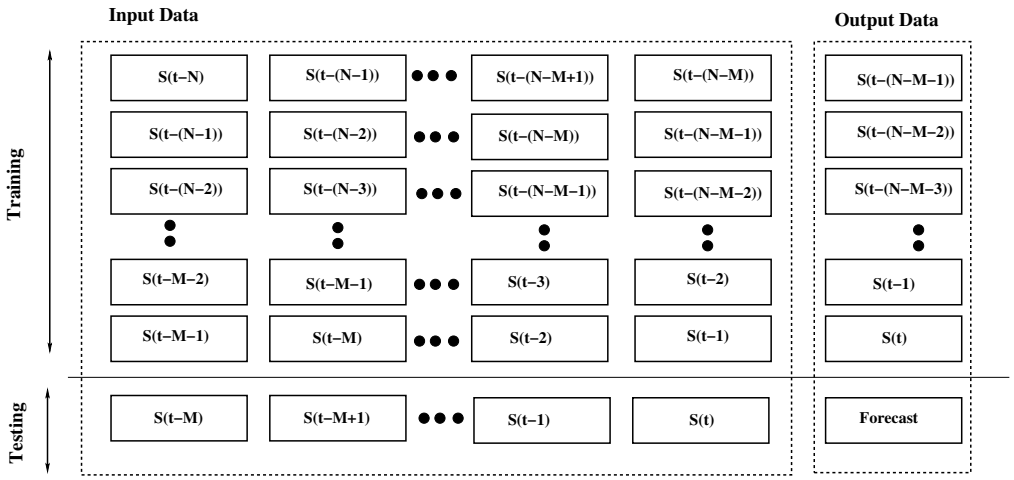


Figure 4.1 : Time series analysis of data

groupings of data help the neural network in better learning of pattern lies in the data series.

4. For the forecasting purpose, for the first time step, FFNN will take first 10 dataset as input and forecast the 11th data, in second time step the FFNN will take second group of 10 dataset and forecast 12th data and the procedure will continue till last data is covered.

5. Forecasted values of all the data sequences, that is, $\hat{S}p_{A_1}(t)$, $\hat{S}_{D3}(t)$, $\hat{S}_{D2}(t)$ and $\hat{S}_{D1}(t)$ are summed up to get the final forecast value of the hybrid model.

4.5 RESULTS

The original data series and its forecast results using FFNN are shown in Figure 4.2. From the figure it is evident that which we do not apply DWT on the data series before prediction, the forecast model is unable to model the sharp fluctuations and results in higher forecast error as seen in Table 4.1. The wavelet decomposed approximation and detailed series and its corresponding forecast of each decomposed series using FFNN is shown in Figure 4.3.

To get the final prediction after application of DWT and FFNN, all the predicted values of the wavelet decomposed series are added together. The final forecast results after applying DWT and FFNN are shown in the Figure 4.4.

Table 4.1 : Error comparison table for different models

Training	Persistence		FFNN		WFFNN	
Data	MAE	RMSE	MAE	RMSE	MAE	RMSE
January	64.43	79.89	21.37	38.04	**3.55**	**5.63**
February	78.69	95.39	33.04	56.74	**5.55**	**9.72**
March	83.45	97.03	23.23	42.74	**5.28**	**8.87**
April	89.95	112.34	34.53	66.06	**7.0**	**11.64**
May	83.51	104.49	21.97	47.45	**6.57**	**11.0**
June	79.65	100.18	30.26	60.83	**6.30**	**10.93**
July	91.35	127.29	78.42	110.76	**11.27**	**16.74**
August	98.63	134.14	80.53	112.90	**12.88**	**17.36**
September	90.03	115.24	52.18	80.17	**8.09**	**12.34**
October	87.56	11.02	41.52	69.91	**7.25**	**11.21**
November	76.49	91.74	12.49	25.82	**4.86**	**7.75**
December	69.39	80.18	8.52	22.34	**3.08**	**5.13**

While applying DWT on the data series before giving input to the FFNN one can notice that the model is able to capture sharp fluctuations present in the data. The forecast error in terms of MAE and RMSE obtained from FFNN and DWT combined with FFNN is presented in Table 4.1.

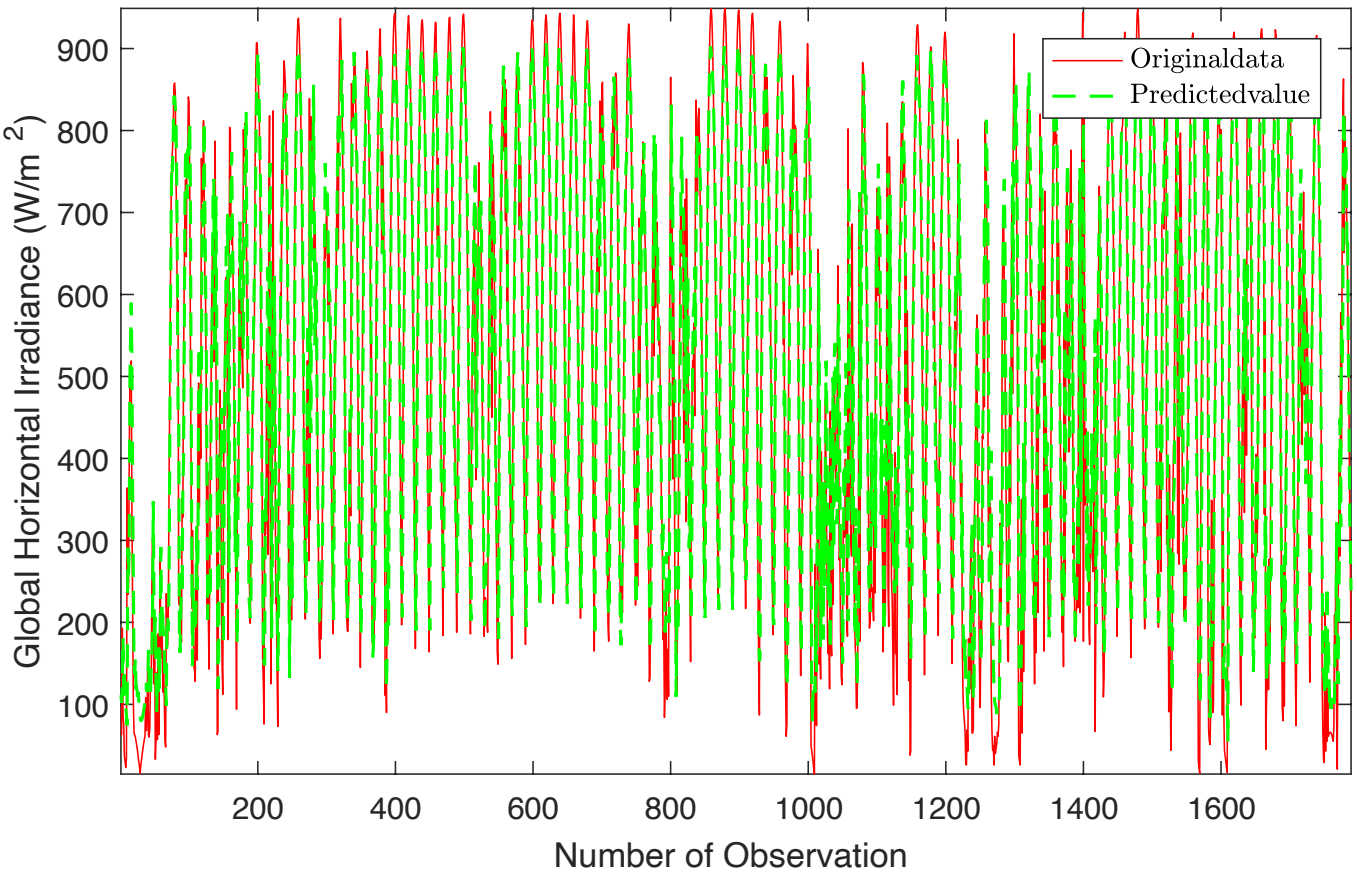
To highlight the improvement in the forecast accuracy the results corresponding to persistence model are presented in the Table 4.1. Here we used Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) as error metric to evaluate the accuracy for comparative statistical analysis of all the comparing models. From the analysis, it is evident that RMSE and MAE for WFFNN is lowest throughout the year for each month compared to FFNN, and the RMSE and MAE for for FFNN is better as compared to persistence model. Moreover, to check whether the accuracy of the WFFNN as compared to FFNN is significantly enhanced or not, we have performed the statistical test. Through statistical test we wanted to know whether the reduction in error is random or it is significant. To test it, we have used t - test to check whether the RMSE of WFFNN is significantly lesser than RMSE of SFFNN. We have performed one tail t - test over RMSE value

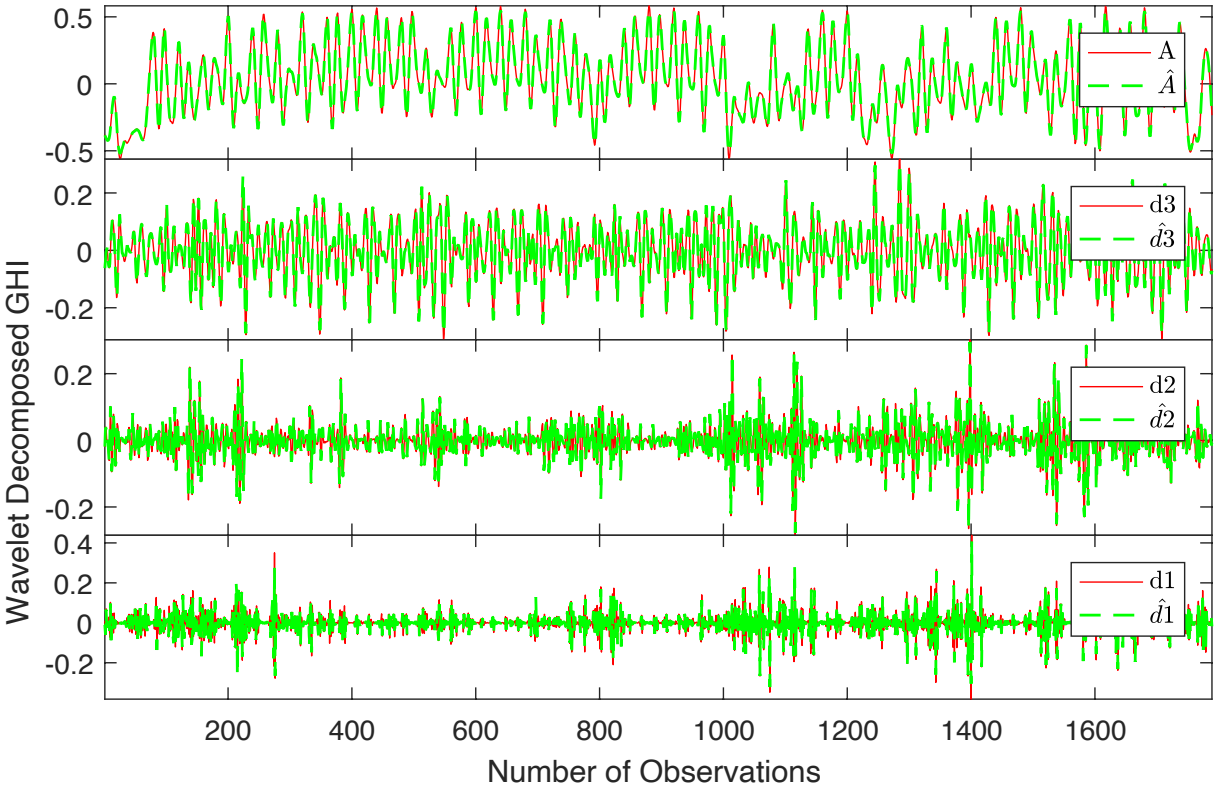
of WFFNN against RMSE value of SFFNN and calculated the p-value and it is coming out to be 0.000014, which is very less. So, we can conclude from this statistical test that the reduction in the error is significant. Here, we want to emphasise that the results which are demonstrated in Figures 4.2, 4.3, 4.4, and 4.5 are corresponding to the month of August because the stated month has least forecast accuracy. Figure 4.5 is demonstrated for few days of August to get better visibility of the results.

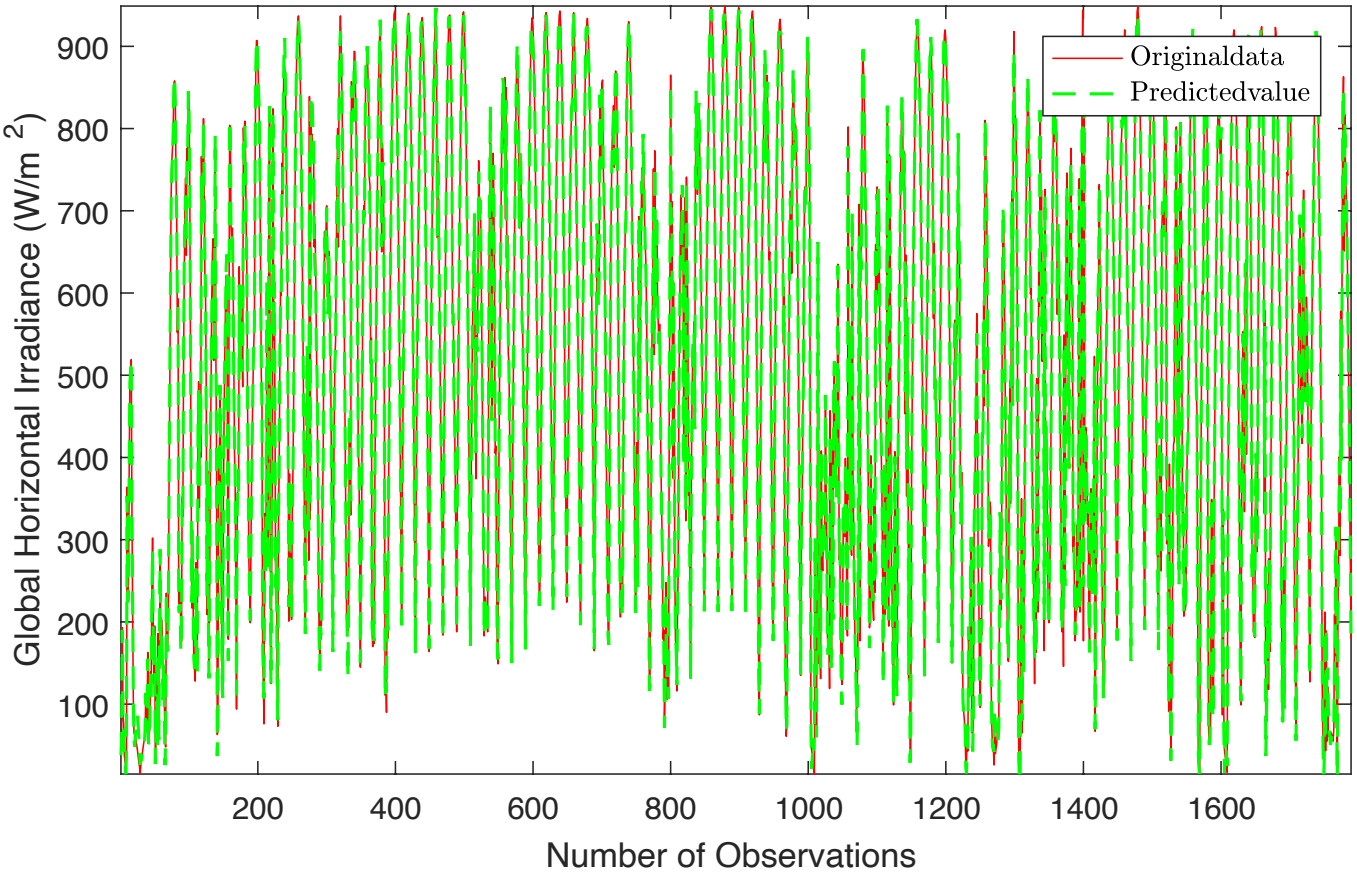
4.6 CONCLUSIONS

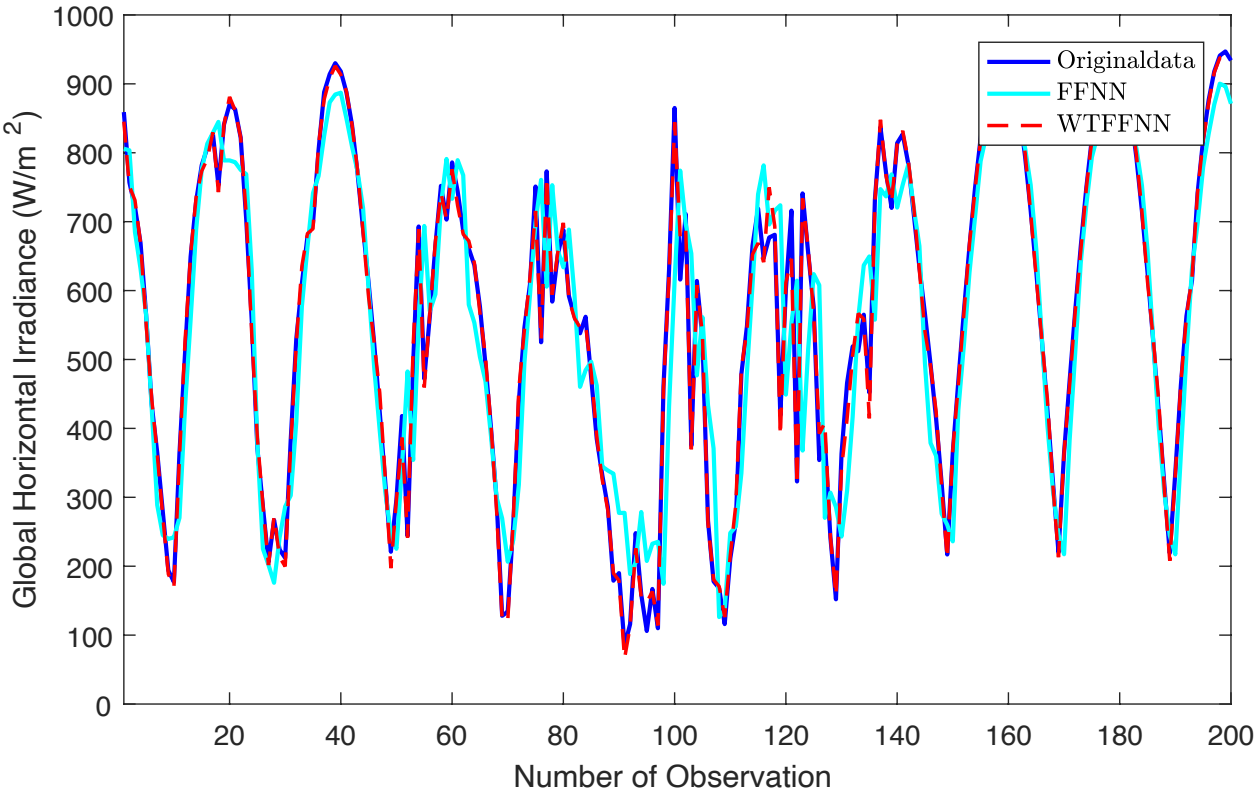
In this chapter a hybrid of DWT and FFNN based model is used for the prediction of GHI time series data on half-hourly frequency. Results indicate that application of FFNN alone can result in better prediction but it can not handle the spikes present in the data series. Wavelet decomposed input improves the learning of the FFNN and results in better forecast accuracy.

…


Originaldata
Predictedvalue
Global Horizontal Irradiance (W/m 2)
Number of Observation
100
200
300
400
500
600
700
800
900
200
400
600
800
1000
1200
1400
1600


Wavelet Decomposed GHI
Number of Observations
A
$\hat{A}$
d3
$\hat{d}3$
d2
$\hat{d}2$
d1
$\hat{d}1$


Originaldata
Predictedvalue
Global Horizontal Irradiance (W/m 2)
Number of Observations
100
200
300
400
500
600
700
800
900
200
400
600
800
1000
1200
1400
1600


Originaldata
FFNN
WTFFNN
Global Horizontal Irradiance (W/m 2)
Number of Observation
0
100
200
300
400
500
600
700
800
900
1000
20
40
60
80
100
120
140
160
180
200