A WAVELET DECOMPOSITION APPROACH TO IMPROVE MODELLING PERFORMANCE OF ARTIFICIAL NEURAL NETWORKS IN SOLAR ENERGY RESEARCH

Sajid Hussain and Ali Al Alili*

Department of Mechanical Engineering, Khalifa University of Science and Technology, Petroleum Institute, P.O.Box 2533, Abu Dhabi, the United Arab Emirates

Abstract

In solar energy research, estimating solar potential over a location of interest is an important step towards the successful planning of renewable energy projects. However, solar data are not available for every point of interest due to the absence of meteorological stations and sophisticated solar sensors. Therefore, the solar radiation has to be estimated with accurate solar radiation models. This paper presents a hybrid technique to increase the modeling performance of artificial neural networks (ANN), commonly used in solar radiation applications. A wavelet multiresolution is performed to decompose the meteorological data into different time and frequency scales before it is presented to the ANN models. As a result, improving the learning process of the ANN models. The proposed approach is compared to traditional ANN model and validated using well-known statistical validation metrics including coefficient of determination (R^2) and root mean square error (RMSE), mean bias error (MBE), mean absolute percentage error (MAPE), and t-statistics. In addition, wavelet cross spectrum (WCS) is used as a visual indicator of the model performance in time, frequency, and phase domains. The results show that the modeling performance of the ANNs is considerably improved from R^2 of 89.21% to 96.34%.

Keywords: *Solar radiation model, artificial neural network, wavelet analysis*

1. Introduction

There is an increasing concern over global warming, emission of greenhouse gases and climate change. This has encouraged the use of renewable energy sources as replacements to coal and oil fired electricity generation processes. Wind, hydro, solar, and geothermal are main sources of renewable energy. The United Arab Emirates, strategically located in the solar belt, has abundance of sunshine throughout the year where the yearly averaged GHI reaches around 6kWh/day (Islam et al. 2009). This gives the UAE an opportunity to harness solar energy and promote clean and green environment. Therefore, accurate solar resource assessment is an inseparable part of an efficient energy planning in the region. Solar radiation reaching the earth's surface are normally measured using different ground based sensors like pyroheliometers, pyranometers and radiometers. However, this process is expensive and cannot be implemented for every geographical region of interest. An alternate way is to model and predict the solar radiations using readily available economical measurements like T, RH, WS, and SSD. Models are trained and validated on historical data collected from ground based sensors at one particular geographical location and used to estimate the solar radiations at nearby locations. This highlights the importance of solar radiation models, their efficiency and accuracy.

Different statistically based (Hassan 2014), (Voyant et al. 2014), (Mohammadi et al. 2015), biologically inspired (Mohandes 2012), (Behrang et al. 2011), (İzgi et al. 2012) and hybrid (Salcedo-Sanz et al. 2014), (Babu and Reddy 2014), (Olatomiwa et al. 2015) models are used for solar radiation modeling. Artificial neural network (ANN) techniques are very common among biologically inspired models. ANN is used to model solar radiation either by using meteorological parameters (Cornaro et al. 2015), (Yacef et al. 2012) or satellite images (Lu et al. 2011), (Linares-Rodriguez et al. 2013).

Often, the modeling performance of the ANN is improved through hybrid techniques (Salcedo-Sanz et al. 2014) or different ANNs are used to model different seasons of the year or more specifically, to cater for temporal nonlinearities in the data. However, none of the considerations is given to spectral non-linearities present in solar and weather time series.

This paper proposes a relatively very simple technique based on wavelet multiresolution analysis. The novelty

doi:10.18086/swc.2017.22.02 Available at http://proceedings.ises.org

of this study is to analyze the time series, both in temporal and spectral domains. A wavelet multiresolution analysis is performed on the signals that decomposes the signals into different frequency and time resolutions, and different ANNs are trained accordingly. Other than modelling, the process filters out the noise present in the signal as noise can degrade the modeling performance of the ANN. This is reflected in the model performance indicators which penalize the model for not correctly modeling the noise components.

The rest of the paper is organized as follows. Section 2 introduces the methodology followed by simulations and discussions in section 3 and conclusions in section 4.

2. Wavelet Decomposition Combined with Artificial Neural Networks

2.1. Wavelet Multiresolution Analysis

In this section, complex time signals are decomposed or transformed into relatively simple parts through wavelet decomposition. Then modeling algorithms are applied on each decomposed part. In the end, all the parts are combined to construct the original signal. A time domain signal can be transformed into frequency domain signal by using the fast Fourier transform (FFT). Although, the frequency domain analysis offers excellent frequency resolution but no time resolution. This problem can be solved by using joint time–frequency analysis (JTFA) techniques like short-time Fourier transform (STFT) and wavelet transform (WT).

The WT applies a strategy where the decomposition of a complex signal is performed using varying time and frequency resolutions. WT is a convolution of a signal with a basis function Ψ(t), called mother wavelet, and its dilated and translated components $\Psi_{d,s}(t)$, called daughter wavelets. Let us denote a time signal as $x(t)$ at time t, the continuous wavelet transform (CWT) of $x(t)$ is $W_x(d,s) = \int_{-\infty}^{+\infty} x(t)\,\Psi^*_{d,s}(t)dt$, where the operator * is a complex conjugate, and $\Psi_{d,s}(t) = \frac{1}{\sqrt{s}}\Psi\left(\frac{t-d}{s}\right)$, where d is translation in time, and s is dilation or scale operator. A discrete wavelet transform (DWT) of a discrete version $x(n)$ of the signal $x(t)$ is computed through a series of high-pass (`h[n]`) and low-pass (`g[n]`) filters, where an *n* level decomposition of the signal $x(n)$ is computed. This step decomposes the original signal into its detailed (D_1, D_2, …, D_n) and approximation (A_1, A_2, …, A_n) wavelet coefficients. A reconstruction step or inverse discrete wavelet transform (IDWT) with high-pass (`H[n]`) and low-pass (`G[n]`) filters is performed to get the original signal back as shown in Fig.1.

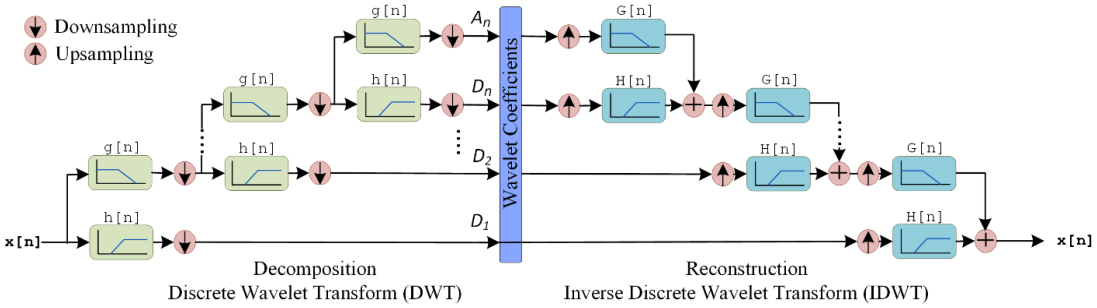

Fig. 1. Decomposition and reconstruction, DWT

The wavelet multiresolution analysis can be combined with ANNs to improve the modeling performance of traditional ANN. Each input variable $x(n)$ and output variable y(n) are decomposed into their detailed and approximation wavelet coefficients through DWT. The decomposition process is stopped where R^2 is maximized. First level detailed coefficients of each input variable are used as input to an ANN to estimate first level detailed coefficient of the output. Similarly, second level detailed coefficients of the input variables are used as input to an ANN to estimate second level detailed coefficient of the output and so on. On the other hand, only n^{th} level approximation coefficients of each input variables are used to estimate n^{th} level approximation coefficient of the output variable. At the end, all estimated detailed and approximation coefficients of the output variable are combined through IDWT to construct an estimation of the original output signal. Figure 2 shows the process.

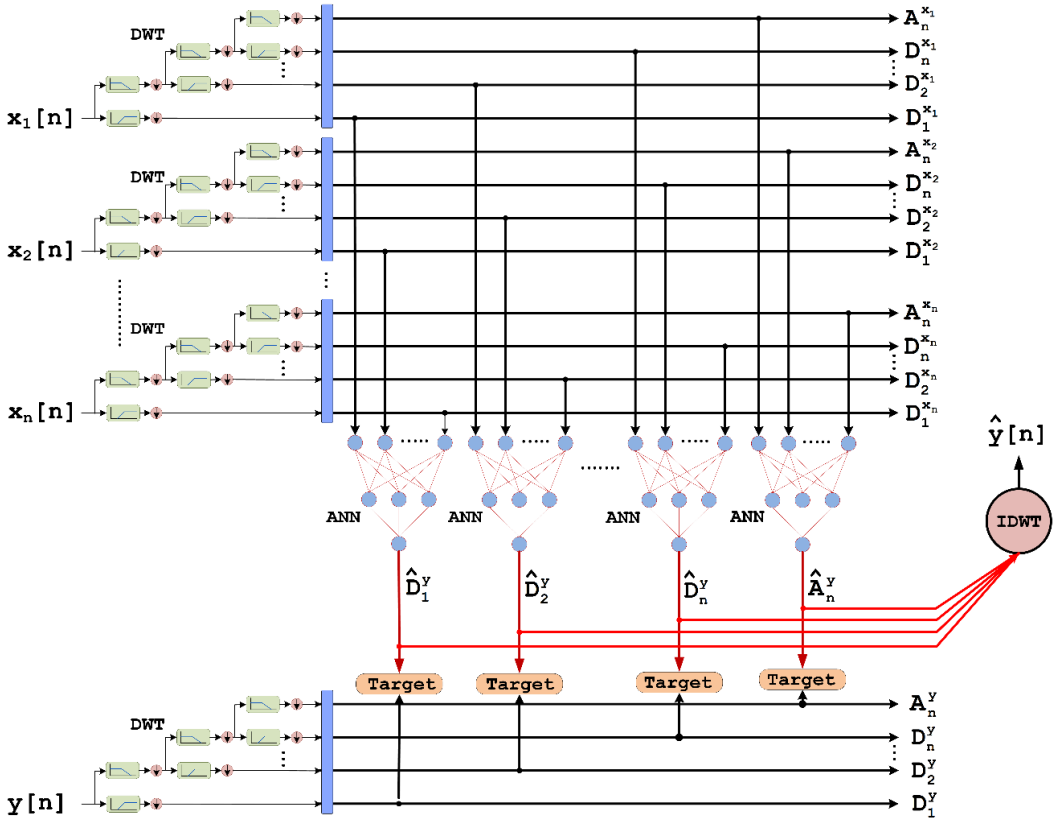

Fig. 2: Wavelet decomposition combined with ANN

2.2. Artificial Neural Network (ANN)

ANN simulates biological nervous systems and tries to mimic information processing capabilities of human brain. ANNs are successfully used in many real-world problems including non-linear function estimation, pattern identification, classification, regression, and forecasting. One of the most common ANN architecture used in research is Multi-layer perception (MLP). Among MLP, the widely used network architecture is feed-forward network with one hidden layer as shown in Fig. 3.

The input–output equations of the k^{th} neuron are as follows (Haykin 1994):

$$y_k = \Phi\left(u_k - \theta_k\right), \qquad u_k = \sum_{j=1}^{n} w_{kj} x_j \,. \qquad \text{(eq. 1)}$$

The input elements of the ANN are $x_1, x_2, ...x_n$ as shown in 3, the output of the summation is u_k , θ_k is a threshold, $w_1, w_2, ...w_n$ are the neuron's synaptic weights and $\Phi(.)$ is an activation function. The neural network has the ability to learn from its environment. In the learning process, synaptic weights of the ANN are updated through enforcement applied to the ANN from the environment. The learning rule has the following form

$$w_{kj}\left(n+1\right) = w_{kj}\left(n\right) + \Delta w_{kj}\left(n\right), \qquad \text{(eq. 2)}$$

where $\Delta w_{kj}\left(n\right)$ is the adjustment of the weight $w_{kj}\left(n\right)$ at n^{th} time-step. The computation of $\Delta w_{kj}\left(n\right)$ is performed through well-known back propagation algorithms including Levenberg-Marquardt optimization, steepest decent techniques and quasi-Newton methods (Haykin 1994).

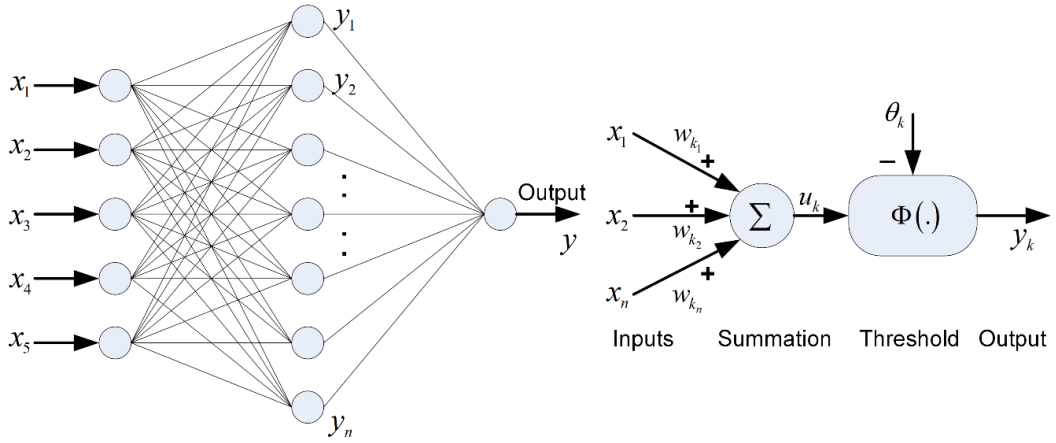

Fig. 3: Feed Forward Artificial Neural Network

2.3. Statistical Evaluation Metrics

The following statistical metrics are used to evaluate the model performance.

$$RMSE = 100\left(\frac{\sqrt{\sum_{n=1}^{N_t}\frac{(GHI-\widehat{GHI})^2}{N_t}}}{\sum_{n=1}^{N_t}\frac{GHI}{N_t}}\right)\%, \quad \text{(eq. 3)}$$

$$MAPE = 100\left(\frac{1}{N_t}\sum_{n=1}^{N_t}\frac{|GHI-\widehat{GHI}|}{|GHI|}\right)\%, \quad \text{(eq. 4)}$$

$$MBE = 100\left(\frac{\sum_{n=1}^{N_t} GHI-\widehat{GHI}}{\sum_{n=1}^{N_t} GHI}\right)\%, \quad \text{(eq. 5)}$$

$$R^2 = 100\left(1-\frac{\sum_{n=1}^{N_t}(GHI-\widehat{GHI})^2}{\sum_{n=1}^{N_t}(GHI-\overline{GHI})^2}\right)\%, \quad \text{(eq. 6)}$$

where root mean square error is RMSE, mean absolute percentage error is MAPE, mean bias error is MBE, and coefficient of determination is R^2. The terms $GHI, \widehat{GHI},$ and $\overline{GHI}$ are original, estimated and mean global horizontal irradiances. The total number of samples in the test data is N_t. In addition, *t*-statistics is performed where it is objectively determined whether a model's estimates are statistically significant. The *t*-statistics is calculated from RMSE and MBE as

$$t = \sqrt{\left(\frac{(N_t-1)MBE^2}{RMSE^2-MBE^2}\right)}. \quad \text{(eq. 7)}$$

The model's performance is better for smaller values of the t-statistics. In order to determine the statistical significance of the model's estimate, a critical t value from standard statistical tables has to be obtained for $\mathbf{1-\alpha}$ confidence level and $\boldsymbol{N_t-1}$ degrees of freedom (DoF).

3. Simulations and Discussions

3.1. Data Collection and Quality Control

Daily averaged data for ten years (2004 – 2013: 3650 samples), collected from Abu Dhabi international airport, UAE at Latitude of 24.22° N and Longitude of 54.65° E are used in this study. The data includes temperature (T : °C), relative humidity (RH : %), wind speed (WS : knots), sunshine duration (SSD : hours), and global horizontal irradiance (GHI : Wh/m^2). The data are checked for missing values by visual inspection. In addition, a multivariate statistical outlier detection using squared Mahalanobis distance is performed. A total of 109 (3%) samples are discarded as being outliers. This leaves a valid data of 3541 samples. Fig. 3 shows the time history of the cleaned data. The data are further normalized in the range of [-1 1], to be used for ANN training.

3.2. ANN Implementation

An MLP is initialized with four input neurons, ten hidden neurons, and one output neuron. The synaptic weights of the MLP are initialized randomly, and Levenberg–Marquardt algorithm is used for training. The stopping criteria are either maximum generations of 300 or MSE value of ≤ 0.02. The number of hidden layer neurons is an important parameter in the modeling process, and in this particular work, an MLP with ten neurons in the hidden layer is found reasonable. This number is found by a simulation experiment where an MLP is initialized with varying number of hidden layer neurons from 1 to 30, and for each MLP, the process of training, validation, and testing is repeated 30 times and the values of RMSE are recorded.

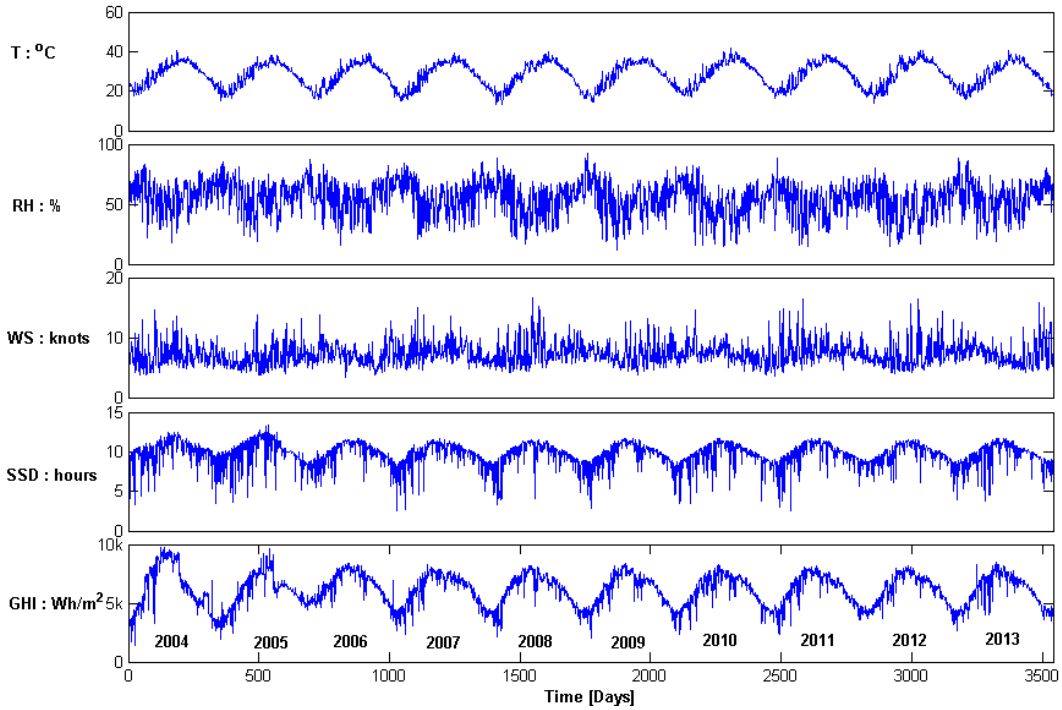

Fig. 4: Time history of the collected data

The MLP with low RMSE is selected for further modeling. The data are divided into training, validation and testing data sets of 2480 (70%), 530 (15%) and 530 (15%) data points, respectively. Fig. 5 shows the results with Fig.5 (a) plotting the time histories of measured and estimated GHIs without wavelet decomposition, using only standalone-MLP. Fig. 5 (d) shows time histories of measured and estimated GHIs with wavelet decomposition, the wavelet approach. Fig. 5 (b) and Fig. 5 (e) plots the results of regression analysis for both the standalone and wavelet approaches, respectively. Fig. 5(c) and Fig. 5 (f) plots the wavelet cross spectrum (WCS) analysis for both approaches. Further analysis of Fig. 5 reveals that wavelet-MLP approach improves the modeling performance of the MLP with R^2 of 96.34% and RMSE of 235.25 Wh m^{-2} day^{-1} (3.88%) as compared to the standalone-MLP approach with R^2 of 89.21% and RMSE of 415.54 Wh m^{-2} day^{-1} (6.83%). A WCS plot in Fig. 5 (c) reveals that in the case of standalone-MLP approach, a frequency component of 0.07×10^{-6} Hz (6-month on the frequency axis or y-axis) is not correctly modeled. A relative phase difference of 45° to 95° (direction of small black arrows inside the WCS plot) between measured and estimated GHIs is observed from the time stamp of 3200 to 3325 days on the time axis (see the red colored arrows indicated on the time axis or x-axis). On the other hand, this is not the problem with the wavelet-MLP approach and the relative phase difference is corrected as shown in Fig. 5 (f). The direction of small black arrows in the WCS plot is towards right most of the time. This concludes that standalone-MLP approach could not model the frequency contents of the GHI signal related to 6-month. However, when the signal is decomposed, with wavelet decomposition, into relatively simple parts, the problem vanishes. The value of R^2 is improved by 6.85%. In addition, looking closely at the time histories in Fig.5 (a), we clearly see the poor modeling performance of the standalone-MLP around time stamps of 3200 to 3325 days at time axis. This corresponds to the months of February, March, and April in the test year of 2013. The regression plot in Fig. 5 (b) reveals that the standalone-MLP model slightly overpredicts the values of GHI in the lower region on 3000 to 5000 Wh m^{-2} day^{-1}. However, this problem is improved in wavelet-MLP regression plot of Fig. 5 (e) because of the removal of

unwanted noise components. Overall, the modeling performance of the wavelet-MLP outperforms the standalone-MLP approach. The optimum wavelet decomposition level where R^2 is maximized is 6.

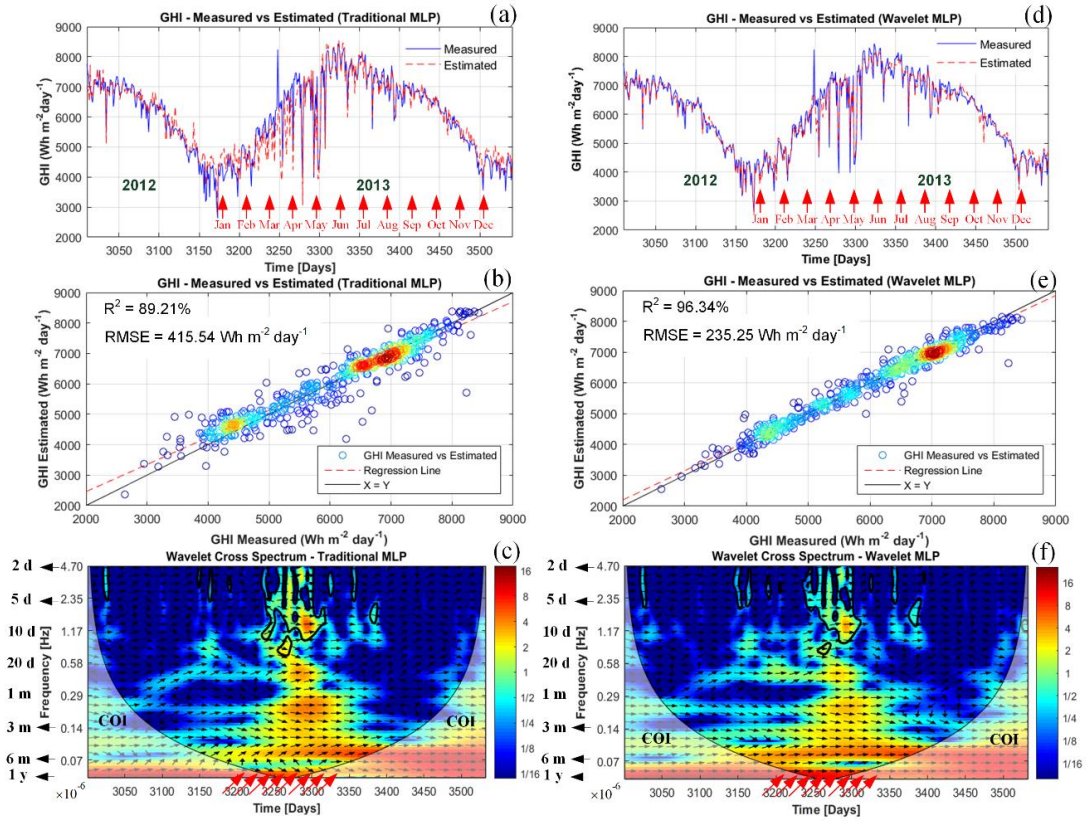

Fig. 5: ANN results (a) measured and simulated GHI, standalone-MLP (b) regression analysis, standalone-MLP (c) Wavelet cross spectrum, standalone-MLP, where relative phase arrows point right for in-phase and left for anti-phase coherence (d) measured and simulated GHI, wavelet-MLP (e) regression analysis, wavelet-MLP (f) Wavelet cross spectrum, wavelet-MLP, where relative phase arrows point right for in-phase and left for anti-phase coherence.

Table1. Model Comparisons

Model	Standalone approach					Wavelet approach				
	R^2 (%) [95% CI]	RMSE (%)	MAPE (%)	MBE (%)	*t*- stat	R^2 (%) [95% CI]	RMSE (%)	MAPE (%)	MBE (%)	*t*-stat
MLP	89.21 [0.87- 0.911]	6.82	5.15	-0.53	1.79	96.34 [0.959- 0.97]	3.89	2.98	0.11	0.59

4. Conclusion

This study models global horizontal irradiation (GHI) over Abu Dhabi, the United Arab Emirates using a widely used soft-computing algorithm in this domain, the artificial neural network (ANN). The suggested model use four sets of inputs: temperature (T), relative humidity (RH), wind speed (WS), and sunshine duration (SSD). The traditional ANN modeling approaches lack the ability to capture temporal and spectral dynamics present in the signals. In this paper, the complex input signals are decomposed into relatively simple time series with varying time and frequency resolutions using wavelet decomposition. ANN modeling is applied to these simple time series, and the modeled time series are combined to form an estimate of the final complex signal. The results show that the proposed strategy significantly improves the performance of traditional ANN. The maximum improvement is observed where the R^2 is improved by 6.83%. In addition, this paper presents a critical analysis of the results in time, frequency and phase domains. It is evident from the analysis that models work better when signals are de-noised, and frequency and phase coherence are achieved between measured and estimated solar radiations for critical and representative frequencies.

This study demonstrates the merits of wavelet-ANN models for solar radiations. However, further testing is required before the suggested techniques are applied to real-life applications. Tests are required for high-resolution data with quantification of potential uncertainties. There need to be rigorous tests at other geographical locations to assess the wider potential and versatility of the proposed wavelet-ANN models for solar energy applications.

5. References

Babu, C. N. and B. E. Reddy (2014). "A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data." Applied Soft Computing **23**: 27-38.

Behrang, M. A., E. Assareh, A. R. Noghrehabadi and A. Ghanbarzadeh (2011). "New sunshine-based models for predicting global solar radiation using PSO (particle swarm optimization) technique." Energy **36**(5): 3036-3049.

Cornaro, C., M. Pierro and F. Bucci (2015). "Master optimization process based on neural networks ensemble for 24-h solar irradiance forecast." Solar Energy **111**: 297-312.

Hassan, J. (2014). "ARIMA and regression models for prediction of daily and monthly clearness index." Renewable Energy **68**: 421-427.

Haykin, S. (1994). Neural Networks: A Comprehensive Foundation. MacMillan, New York.

Islam, M. D., I. Kubo, M. Ohadi and A. A. Alili (2009). "Measurement of solar energy radiation in Abu Dhabi, UAE." Applied Energy **86**(4): 511-515.

İzgi, E., A. Öztopal, B. Yerli, M. K. Kaymak and A. D. Şahin (2012). "Short–mid-term solar power prediction by using artificial neural networks." Solar Energy **86**(2): 725-733.

Linares-Rodriguez, A., J. A. Ruiz-Arias, D. Pozo-Vazquez and J. Tovar-Pescador (2013). "An artificial neural network ensemble model for estimating global solar radiation from Meteosat satellite images." Energy **61**: 636-645.

Lu, N., J. Qin, K. Yang and J. Sun (2011). "A simple and efficient algorithm to estimate daily global solar radiation from geostationary satellite data." Energy **36**(5): 3179-3188.

Mohammadi, K., S. Shamshirband, M. H. Anisi, K. A. Alam and D. Petković (2015). "Support vector regression based prediction of global solar radiation on a horizontal surface." Energy Conversion and Management **91**: 433-441.

Mohandes, M. A. (2012). "Modeling global solar radiation using Particle Swarm Optimization (PSO)." Solar Energy **86**(11): 3137-3145.

Olatomiwa, L., S. Mekhilef, S. Shamshirband, K. Mohammadi, D. Petković and C. Sudheer (2015). "A support vector machine–firefly algorithm-based model for global solar radiation prediction." Solar Energy **115**: 632-644.

Salcedo-Sanz, S., C. Casanova-Mateo, A. Pastor-Sánchez and M. Sánchez-Girón (2014). "Daily global solar radiation prediction based on a hybrid Coral Reefs Optimization – Extreme Learning Machine approach." Solar Energy **105**: 91-98.

Voyant, C., C. Darras, M. Muselli, C. Paoli, M.-L. Nivet and P. Poggi (2014). "Bayesian rules and stochastic models for high accuracy prediction of solar radiation." Applied Energy **114**: 218-226.

Yacef, R., M. Benghanem and A. Mellit (2012). "Prediction of daily global solar irradiation data using Bayesian neural network: A comparative study." Renewable Energy **48**: 146-154.