

Hybrid Solar Forecasting Method Based on Empirical Mode Decomposition and Back Propagation Neural Network

Ahmed Aghmadi^{1,*}, Soumia El Hani, Hamza Mediouni, Nisrin Naseri, and Fatima El Issaoui

¹Energy optimization, Diagnosis and Control, STIS Center, ENSET, Mohamed V University, Rabat, Morocco

Abstract. In order to improve the accuracy of solar radiation prediction and optimize the energy management system. This study proposes a forecasting model based on empirical mode decomposition (EMD) and Back Propagation Neural Network (BPNN). Empirical mode of decomposition (EMD)-based ensemble methods with powerful predictive abilities have become relatively common in forecasting study. First, the existing solar radiation datasets are decomposed into an intrinsic mode function (IMF) and one residue produces fairly stationary sub-series that can easily be modeled on BPNN. Next, both components of the IMF and residue are applied to create the respective BPNN models. Then, the corresponding BPNN is used to predict some sub-series. Finally, the predictive values of the original solar radiation datasets are determined by the sum of each predicted sub-series. Compared with traditional models such as conventional neural network or ARIMA time series, the hybrid EMD-BPNN model shows great results in term of RMSE with 28.13 (W/m²). On the other hand, the result of BPNN and ARIMA was 83.28 (W/m²) and 108.88 (W/m²), respectively. that the non-stationary and non-linear of solar radiation signal has less effect on the accuracy of the prediction.

1 Introduction

Climatologists are increasingly convinced that climate change has been one of the biggest threats in the world over the last few decades [1]. People have therefore begun to think about other possible energy resources and have found renewable energy not only as a strategic solution to reduce the gap between production and consumption, but also a Green and clean energy. Hence, Renewable resources including wind and solar energy, ought to be utilized effectively to overcome the rapidly expanding in energy needs and changing climate [2]. Photovoltaic energy generation in a specific location can be extremely ineffective and highly expensive without an accurate forecasting of solar irradiance, know that solar irradiance is proportional to the production of photovoltaic energy generated by solar panels. Though, PV power generation is challenging as it is strongly linked to numerous environmental variables, such as solar irradiance, temperature, wind speed and cloud movement [3]. Multiple studies have been published on methods for forecasting solar radiation, which can be categorized in three groups: (a) physical models; (b) statistical models; (c) ensemble methods. Every prediction model category has its own features and capabilities. Solar radiation prediction models can be developed using physical parameters such as temperature, pressure, and Sky Imagery for longer-term forecasts. This method is therefore normally developed by

meteorologists and used in large-scale weather forecasts [4]. Unlike statistical models in which there are several sub-models such as artificial neural network, support vector machine, and regression models. The statistical methods rely strongly on historical data, as well as the ability to collect past data to predict time series [5]. The ensemble method or so-called hybrid method corresponds to mixture of statistical or physical methods. This approach is intended to merge various models with specific features in order to overcome the weakness of the particular model, and thereby maximizing the forecast performance [6]. In the field of solar irradiance forecast, a combined Empirical Mode Decomposition (EMD) and Artificial Neural Network (ANN) model may be an effective method. The Empirical Decomposition Mode (EMD) [7], suggested in 1998 by Huang et al. seems to be a recent approach to adaptive data analysis to improve predictability for nonlinear and non-stationary solar radiation data [8]. The key feature of the EMD is that a signal can be totally decomposed into a range of signals called Intrinsic mode functions (IMF) and a residue. Then build up the prediction models based on the BPNN algorithm for each IMF and the residue, sum all the predictive value of the IMFs and residues to obtain the ultimate expected performance. Our contribution to this research represented in establish a structure for short-term solar irradiance prediction based on EMD and hybrid models. The rest of this paper is prepared as follow, section II focuses on Data collection and analysis,

* Corresponding author: ahmed.aghmadi@gmail.com

section III represents the empirical mode of decomposition and explains the machine learning algorithm, section IV has the simulation result analysis and section V the conclusion.

2 Data and analysis

Measured Direct Normal irradiance (DNI) used in this study was extracted from a meteorological ground station located in Rabat, belongs to the research institute for solar energy and new energies (IRESEN). The data collected cover the January – December 2018 duration and consist of averages of one hour in which there are 8760 measurements. For the process of building the proposed models 8000 samples are used and the rest for testing and checking the performance of models separately. The Average of solar radiation intensity is 4.9964 (kWh/m²/day) [9].

3 Methodology

3.1 Empirical mode decomposition (EMD)

The EMD is robust in analyzing non-linear and non-stationary signal sequences, it is based on decomposing any signal into several intrinsic mode oscillations. Each mode of oscillation satisfies the corresponding criteria: (a) The number of extremes and the number of zero crossings in the entire set of data should be the same or at most one differs. (b) At every moment, the mean sum of the local maxima envelope and the local minima envelope is zero. The decomposition of the time series is defined below:

Step 1: Locate all the local maxima and local minima in $s(t)$ and bind all the local maxima by a cubic spline to make the upper envelope configurable. This process is repeated with a local minima to produce.

Step 2: Build the medium $m_1(t)$ envelope with the upper and lower envelopes average.

Step 3: The average envelope is stripped from $s(t)$ to extract the first $h_1(t)$ component:

$$h_1(t) = s(t) - m_1(t) \quad (1)$$

Step4: Test that $h_1(t)$ fulfills the IMF requirements. If not, return to step 1 and use $h_1(t)$ for the second screening as the initial signal:

$$h_2(t) = h_1(t) - m_2(t) \quad (2)$$

Repeat screens for K times until $h(t)$ fulfills IMF specifications if $c_1(t)$ is extracted from the first IMF component:

$$c_1(t) = h_k(t) \quad (3)$$

Stage 5: Subtract $c_1(t)$ to obtain residual $r_1(t)$ from the initial $s(t)$ signal:

$$r_1(t) = s(t) - c_1(t) \quad (4)$$

Step 6: Take $r_1(t)$ as the new initial signal and complete step 1 to step 5 to get new $r_2(t)$ as the residual signal. Repeat for n times the above steps. The IMF can no longer be decomposed, and the whole EMD is executes if n th residual $r_n(t)$ is a monotonic function. The original $s(t)$ signal may be viewed as n -IMF elements and $r_n(t)$ which can be seen in Equation (5):

$$s(t) = \sum_{k=1}^n c_{(k)}(t) + r_{(n)}(t) \quad (5)$$

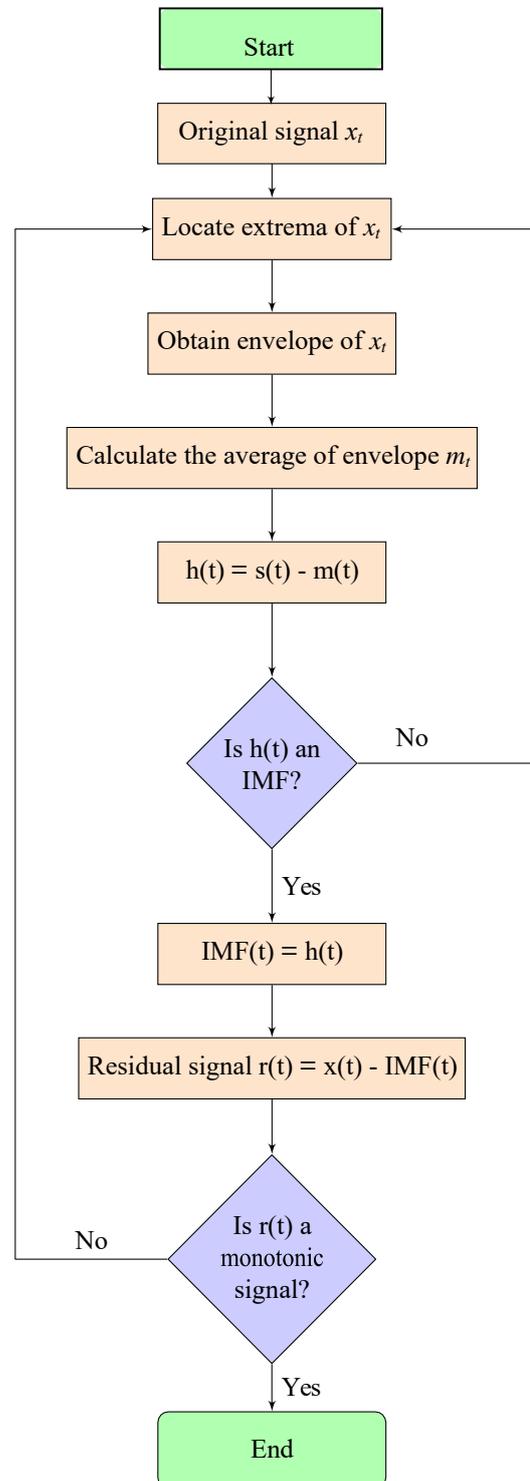


Fig. 1. Flowchart of empirical mode decomposition.

3.2 Back propagation - Artificial neural network (BP-ANN)

ANN method is a simultaneous computing approach motivated by the ability to imitate human thinking; ANN have also been commonly used for non-linear forecasting issues. The back propagation (BP) network is still the most widely used learning algorithm and the original connection weights of the nodes change during the training phase in order to minimize total network errors [10]. In general, an ANN model consists of an input layer (one or more hidden layers) and an output layers by separating nodes into various classes. Each one includes multiple nodes. The number of hidden layers in this paper is one.

3.3 Arima modeling

The value of the dependent variable is represented in the ARIMA model as a linear relation between the previous values of the dependent variable and random errors. The ARIMA model (p, d, q) was implemented for the series, especially because the data sample classifies as a time series. It explains the statistical relation between a sequence and their own findings in the past. It is especially suitable for short-term prediction [11].

The model for ARIMA (p, d, q) is:

$$\Phi_p(B)(1 - B)^d X_t = \theta_0 + \theta_q(B)e_t \quad (6)$$

$\Phi_p(B)$ is a characteristic polynomial of order p, $\theta_q(B)$ is a characteristic polynomial of order q, $(1 - B)$ is the differencing operator. X_t is the observation value at time t, e_t is a white noise process, θ_0 is a constant term, and d is the order of the non-seasonal differencing.

3.4 Hybrid EMD-ANN model

The solar radiation forecasting process with the suggested hybrid method consists of three steps. The signal can be decomposed into practical local time scales in the first step, due to the non-stationary, and stochastic properties of the solar radiation sequence. In the second step, the prediction is done independently for ANN model for the whole decomposed elements. The adaptive algorithm specification for an IMF is focused on the PACF element and the frequency components obtained from the related IMF. The predictions of each IMF are aggregated in the final stage in order to obtain the outcome of the preliminary prediction.

3.5 Evaluation criteria

Three efficiency standards for model reliability and accuracy are used: MAE, MAPE and RMSE. Researchers often use these parameters to check a prediction model's efficiency. RMSE shows the error of the model by evaluating the variations between the observed and the expected values. Small RMSE values indicate that the

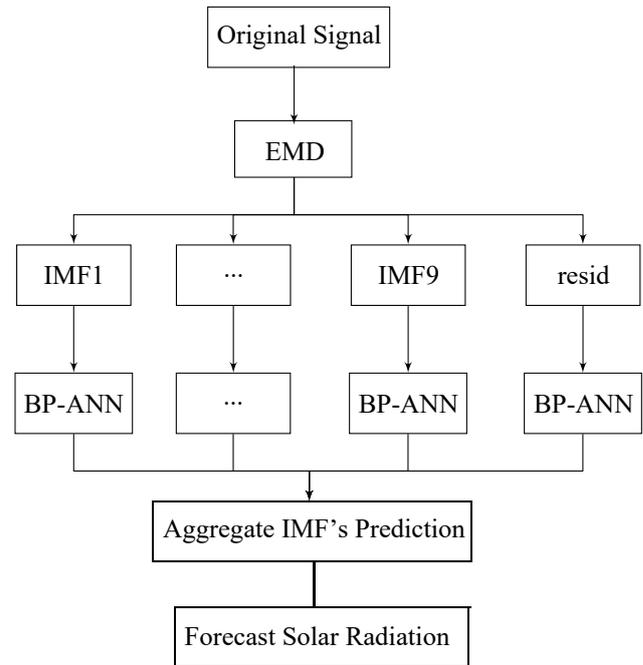


Fig. 2. Framework for hybrid EMD-BPNN model.

model correctly reflects global solar radiation [12]. RMSE is defined in (7) where:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (X_{mes}(i) - X_{pre}(i))^2} \quad (7)$$

where X_{pre} is the predicted solar radiation and X_{mes} is the measured solar radiation. Therefore, mean absolute error (MAE) is the sum of the absolute variations between the forecast and the real observation over the test sample where all variations are of equal weight.

$$MAE = \frac{1}{N} \sum_{n=1}^N |X_{mes}(i) - X_{pre}(i)|^2 \quad (8)$$

The R squared error describes how well the model fits the data:

$$R^2 = 1 - \frac{\sum_{n=1}^N (X_{mes}(i) - X_{pre}(i))^2}{\sum_{n=1}^N (X_{mes}(i) - X_{mean}(i))^2} \quad (9)$$

X_{mes} is the real value, X_{pre} is the predicted value, and X_{mean} the mean value of data.

4 Results and discussion

In this section, the original dataset of direct solar radiation is subdivided into nine IMF's and one residual through Empirical mode decomposition shown in Fig 3 (a), every subset is used to create the appropriate BPNN prediction model, and each BPNN is used to predict the corresponding sub-series, the overall prediction of the original solar radiation is obtained by combining the prediction values of each sub-series. In this analysis, the main BP neural network composed input variables.

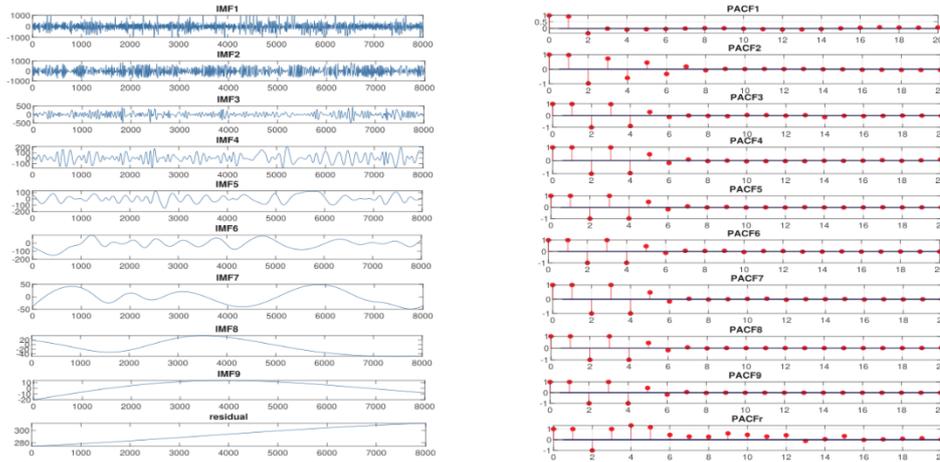


Fig. 3. (a) Decomposition of the solar radiation profile by EMD (b) PACF of respective IMF.

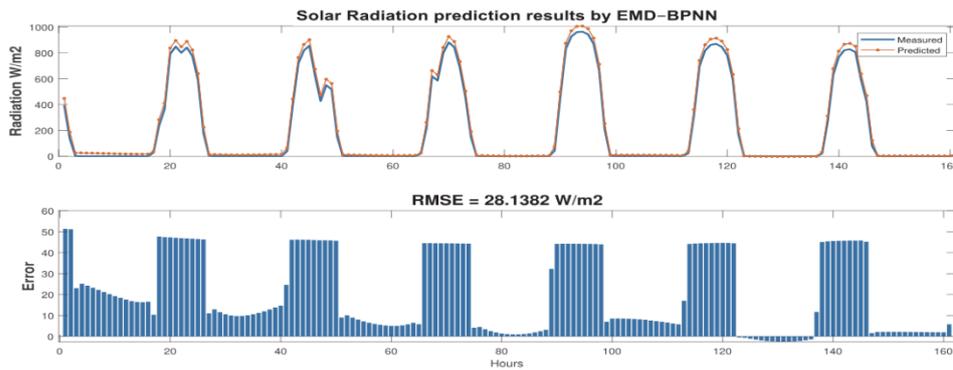


Fig. 4. Solar Radiation results by EMD-BPNN.

We will define them shortly, one output variable, a hidden layer with ten neurons, and a learning rate of 0.03. Training function was Bayesian regularization backpropagation, it takes longer but good for challenge problem.

In addition, the PACF is defined separately for every IMF in Fig 3 (b). Depending on the PACF obtained, the input variables are chosen for the predict of each IMF.

To compare the prediction outcome, several methods have been applied such as the ARIMA time series and Artificial neural network. In this study, ARIMA (1, 1, 2) was employed to model solar radiation in Rabat from January 2018 to November 2018.

Fig 4. shows the solar radiation prediction results by EMD-BPNN in addition to error graph with RMSE of 28.13 W/m^2 the average forecast outcomes by various prediction models and their comparative data for one week are shown in Fig 5.

Performance of prediction models in Table 1 show that EMD and the BP neural network forecasting outcomes are much more reliable than those achieved by the conventional BPNN system. Since the original unpredictable and non-linear data is converted into several fixed frequency and periodical components (i.e. Intrinsic mode functions) after the EMD. The results of these components are more robust when predicted by the BP neural network than the predictive effects of the

original data, and we clearly see that the has the highest R^2 which means that the proposed model fits very well the data with 0.992.

Table 1. Performance of prediction models for Solar Radiation.

Model	RMSE	MAE	R^2
ARIMA	108.88	56.97	0.889
BPNN	83.28	36.7	0.935
EMD-BPNN	28.13	20.99	0.992

5 Conclusion

This study proposes a new methodology focused on EMD and BP neural networks for solar radiation forecasts. Data are split down to various IMFs and a residue by EMD. For the prediction of each IMF and the residue, BP neural network is introduced. Finally, the expected outcomes for every IMF will be simply added and updated. In our case, this method's RMSE and MAE are 28.13 % and 20.99% respectively. much smaller than the conventional BPNN and ARIMA time series method. It is shown that the method suggested can be used efficiently in the prediction of solar radiation.

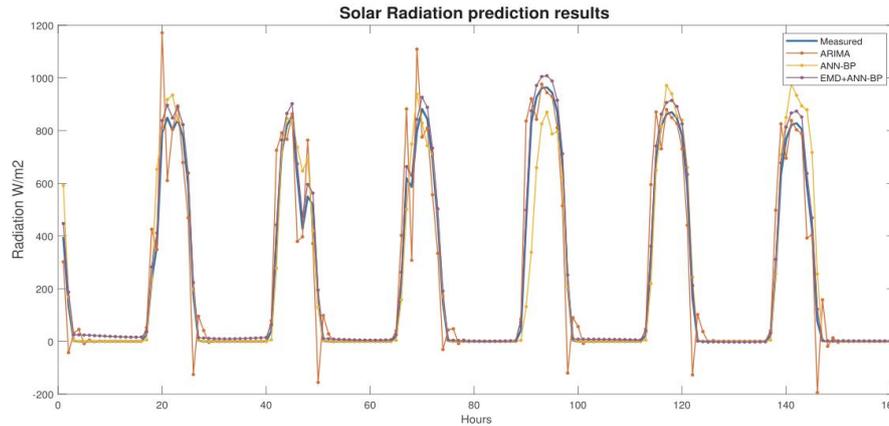


Fig. 5. Solar Radiation results by EMD-BPNN

Acknowledgment. The authors are thankful to the institute IRESEN-Morocco for providing data of the area analyzed for this study.

References

1. Zhu, Bangzhu. "A novel multiscale ensemble carbon price prediction model integrating empirical mode decomposition, genetic algorithm and artificial neural network." *Energies* 5.2 (2012): 355-370.
2. Yadav, Harendra Kumar, Yash Pal, and Madan Mohan Tripathi. "Short-term PV power forecasting using empirical mode decomposition in integration with back-propagation neural network." *Journal of Information and Optimization Sciences* 41.1 (2020): 25-37.
3. Xie, Tuo, et al. "A hybrid forecasting method for solar output power based on variational mode decomposition, deep belief networks and autoregressive moving average." *Applied Sciences* 8.10 (2018): 1901.
4. Liu, Hui, et al. "A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks." *Renewable Energy* 48 (2012): 545-556.
5. Sobri, Sobrina, Sam Koohi-Kamali, and Nasrudin Abd Rahim. "Solar photovoltaic generation forecasting methods: A review." *Energy Conversion and Management* 156 (2018): 459-497.
6. Leva, Sonia, et al. "Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power." *Mathematics and computers in simulation* 131 (2017): 88-100.
7. Huang, Norden E., et al. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences* 454.1971 (1998): 903-995.
8. Qiu, Xueheng, et al. "Empirical mode decomposition based ensemble deep learning for load demand time series forecasting." *Applied Soft Computing* 54 (2017): 246-255.
9. El Mghouchi, Y., T. Ajzoul, and A. El Bouardi. "Prediction of daily solar radiation intensity by day of the year in twenty-four cities of Morocco." *Renewable and Sustainable Energy Reviews* 53 (2016): 823-831.
10. Luo, Xianglong, Guohong Niu, and Qianjiao Wu. "Short-Term Traffic Flow Prediction Based on EMD and Artificial Neural Network." *ICCTP 2009: Critical Issues In Transportation Systems Planning, Development, and Management*. 2009. 1-6.
11. C Nwokike, Chukwudike, et al. "ARIMA Modelling of Neonatal Mortality in Abia State of Nigeria." *Asian Journal of Probability and Statistics* (2020): 54-62.
12. Ali, Mumtaz, and Ramendra Prasad. "Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition." *Renewable and Sustainable Energy Reviews* 104 (2019): 281-295.