Applied Energy 164 (2016) 639–649

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

A new approach for model validation in solar radiation using wavelet, phase and frequency coherence analysis

Sajid Hussain, Ali Al-Alili *

Department of Mechanical Engineering, The Petroleum Institute, P.O. Box 2533, Abu Dhabi, United Arab Emirates

HIGHLIGHTS

- Solar radiation models are usually validated using time domain statistical metrics.
- These metrics have inherent property to penalize models because of noise components.
- This research investigates new validation indices in phase and frequency domains.
- A comparison of different neural network models are presented.
- The proposed techniques reveal an in-depth insight into the model's performance.

ARTICLE INFO

Article history:
Received 28 June 2015
Received in revised form 19 November 2015
Accepted 12 December 2015

Keywords:
Solar models
Artificial neural network
Model validation
Wavelet coherence
Phase and frequency coherence

ABSTRACT

The performance of solar radiation models is evaluated based on time domain statistical metrics. These metrics include root mean square error (RMSE), mean absolute percentage error (MAPE) and coefficient of determination (R^2). However, this study adopts a new approach to model validation in phase and frequency domains. The study proposes frequency coherence and phase synchronization methods quantified into frequency coherence index (FCI) and phase lock value (PLV), respectively. To get more in-depth insight of the model's performance, important visual indicators based on wavelet cross-spectrum (WCS) and wavelet coherence analysis (WCA) are presented. Two different widely used solar radiation models based on artificial neural network (ANN) namely, the multi-layer perceptron (MLP) and the adaptive neuro-fuzzy inference system (ANFIS), are analysed. The proposed techniques are used to exploit weaknesses and strengths of the models. Anomalies in phase and frequency components are identified and the information is used to increase their performance. The MLP is modified into a non-linear autoregressive recurrent exogenous neural networks (NARX-NN) using recursive filtering. Results show that NARX-NN corrects the phase differences, filters out the faulty frequency components and increases the time domain validation metric, i.e. R^2.

1. Introduction

Solar energy is one of the major energy resources of the future. However, climatological factors makes solar energy one of the highly stochastic energy resources in nature. This greatly affects the performance of the electricity distribution networks in case of grid-connected photovoltaic (PV) systems. Therefore, stakeholders and energy planners need to know the amount of solar power that can be generated by a solar system. To this aim, solar potential is measured through ground based sensors and the information is used to calculate power output from the planned solar systems. However, it is often costly and time consuming to install these sensors at different geographical locations and achieve the required spatial resolution. Alternatively, solar radiation is estimated through climatological variables like temperature (T), relative humidity (RH), wind speed (WS) and sun-shine duration (SSD).

Many algorithms have been developed in the past and researchers have used statistically based, biologically inspired, and hybrid models for this purpose. Among statistically based models, autoregressive moving averaging (ARMA) and support vector machine (SVM) regressions are very common. Hassan [1] used an ARMA method for prediction of daily and monthly clearness index for the city of Mosul, Iraq. He used three years of meteorological data to train the regression models. The author compared different structures of the ARMA model and found out that ARMA(2 1 1) model is the best choice in his case with mean percentage error (MPE), root mean squared error (RMSE), mean bias error (MBE),

* Corresponding author.
E-mail address: alialalili@pi.ac.ae (A. Al-Alili).

http://dx.doi.org/10.1016/j.apenergy.2015.12.038

Nomenclature

$\bar{S}_y^W(\cdot)$	smoothed wavelet power spectrum of measured signal
$\bar{S}_{\hat{y}}^W(\cdot)$	smoothed wavelet power spectrum of estimated signal
$\bar{S}_{y\hat{y}}^W(\cdot)$	smoothed wavelet cross spectrum
$\hat{Y}(f)$	estimated irradiance (frequency domain)
$\hat{y}(t)$	estimated irradiance (time domain)
$\mathbf{S}^{-1}$	inverse of covariance matrix
$\phi_y(t)$	instantaneous phase – measured irradiance
$\phi_{\hat{y}}(t)$	instantaneous phase - estimated irradiance
$\Phi_{k,l}$	relative phase
$\psi(t)$	wavelet function (time domain)
$\zeta(t)$	analytical signal
$A(t)$	instantaneous amplitude
$C_{y\hat{y}}(f)$	magnitude squared coherence
D_i^2	squared Mahalanobis distance
R^2	coefficient of determination
$S_{y\hat{y}}^W(\cdot)$	wavelet cross power spectrum
$S_y^W(\cdot)$	wavelet power spectrum
$S_y(f)$	auto spectral density – measured irradiance
$S_{\hat{y}(f)}$	auto spectral density – estimated irradiance
$S_{y\hat{y}}(f)$	cross spectral density
$W_y(\cdot)$	wavelet coefficients
$WC_{y\hat{y}}^2(\cdot)$	wavelet coherence
$Y(f)$	measured irradiance (frequency domain)
$y(t)$	measured irradiance (time domain)
ANFIS	neuro-fuzzy inference system
ANN	artificial neural network
ARMA	auto regressive moving average
BP	backpropagation
COI	cone of influence
EEG	electroencephalogram
FCI	frequency coherence index
FFT	fast Fourier transform
GP	genetic programming
IIR	infinite impulse response
JTFA	joint time frequency analysis
MABE	mean absolute bias error
MAPE	mean absolute percentage error
MBE	mean bias error
MLP	multi layer perceptron
MPE	mean percentage error
MSE	mean square error
NARX-NN	non-linear autoregressive recurrent exogenous neural networks
nRMSE	normalized root mean square error
NSE	NashSutcliffe model efficiency coefficient
PLV	phase lock value
PSO	particle swarm optimization
PV	photovoltaic
RBF	radial basis functions
RH	relative humidity
RMSE	root mean square error
rRMSE	relative root mean square error
SRT	Statistical Regression Techniques
SSD	sun-shine duration
STFT	short time Fourier transform
SVM	support vector machine
T	temperature
WCA	wavelet coherence analysis
WCS	wavelet cross spectrum
WS	wind speed
WT	wavelet transform

and NashSutcliffe model efficiency (NSE) values of 0.9045, 0.2714, 0.0908, and 0.9829, respectively. Mohammadi et al. [2] used SVM regression model to predict global horizontal solar radiation for the city of Isfahan, Iran. The authors used 9 years data for training and 4 years data for testing. The authors reported the best values of mean absolute percentage error (MAPE), mean absolute bias error (MABE), RMSE, relative RMSE (rRMSE), and coefficient of determination (R^2) as 10.4466%, 1.2524 MJ/m^2, 2.0046 MJ/m^2, 9.0343%, and 0.9133, respectively. The authors compared their SVM regression model with particle swarm optimization (PSO) based model and proved that SVM-based model performed better than PSO-based model in their particular case.

Among biologically inspired models, PSO, neuro-fuzzy, genetic programming (GP) regression, and artificial neural network (ANN) are very common. Mohandes [3]modelled global solar radiation using a combination of PSO and ANN for 41 stations in Saudi Arabia. He used meteorological data since 1970. The author compared the performance of PSO–ANN with backpropagation ANN (BP-ANN) and reported the superiority of the PSO–ANN with MAPE of 8.85% as compared to BP-ANN with MAPE of 12.61%. Behrang et al. [4] used PSO to obtain coefficients for the well-known Ångstrom model, empirically. The authors used global solar radiation data for 17 cities in Iran. They compared the performance of the PSO-Ångstrom with statistical regression technique combined with Ångstrom (SRT-Ångstrom) and found out that PSO-Ångstrom was better with an average MAPE reduction of 7.56% as compared to SRT-Ångstrom. Yadev et al. [5,6] used the MLP to model global solar radiation in different states of India. The authors used monthly averaged daily radiation data for 14 years from 26 states in India. The performance of the models was measured using MAPE and the best MLP with MAPE of 6.89% was reported. Mellit et al. [7] used adaptive wavelet ANN to model daily solar radiation in Algeria. The authors used daily total radiation data for 20 years. The MAPE of 6% was reported. Bosch et al. [8] used ANN to model solar radiation over a mountainous area in Spain. Data for 3 years were used and the values of RMSE and MBE were reported as 6% and 0.2%, respectively. Dorvlo et al. [9] used radial basis function (RBF) and MLP networks to model solar radiation in Oman. Data from eight different stations for 10 years were used. The best RBF and MLP were reported with RMSE of 0.83 MJ/m^2/day and 1.01 MJ/m^2/day, respectively. Alam et al. [10] developed 16 different ANN models to estimate solar radiation for different stations in India. The authors reported three best networks with R^2 values of 0.931, 0.907 and 0.923 for summer, raining, and winter seasons, respectively. Fadare [11] used different ANNs to model solar energy potential for 195 cites in Nigeria over the period of 10 years. He reported the values of R^2 as 97.8%, 97% and 95.6% for training, testing and the whole dataset, respectively. Amrouche and Pivert [12] used ANN to model daily global solar radiation in France. They used high resolution data (5 min) from June 2008 to May 2009. For the best ANN, MSE and RMSE were reported as 16.45 W/m^2 and 33.10 W/m^2, respectively. Senkal and Kuleli [13] estimated solar radiation over Turkey using ANN. The authors used meteorological data from August 1997 to