Available online at www.sciencedirect.com
ScienceDirect
Solar Energy 111 (2015) 297–312
SOLAR ENERGY
www.elsevier.com/locate/solener

# Master optimization process based on neural networks ensemble for 24-h solar irradiance forecast

C. Cornaro [a,b,*], M. Pierro [a], F. Bucci [a]

[a] Department of Enterprise Engineering, University of Rome Tor Vergata, Via del Politecnico, 1 00133 Rome, Italy
[b] CHOSE, University of Rome Tor Vergata, Via del Politecnico, 1 00133 Rome, Italy

Received 3 March 2014; received in revised form 10 October 2014; accepted 24 October 2014
Available online 25 November 2014

Communicated by: Associate Editor David Renne

**Abstract**

In the paper two models implemented to forecast the hourly solar irradiance with a day in advance are described. The models, based on Artificial Neural Networks (ANN), are generated by a master optimization process that defines the best number of neurons and selects a suitable ensemble of ANN.

The two models consist of a Statistical (ST) model that uses only local measured data and a Model Output Statistics (MOS) that corrects Numerical Weather Prediction (NWP) data. ST and MOS are tested for the University of Rome "Tor Vergata" site. The models are trained and validated using one year data. Through a cross training procedure, the dependence of the models on the training year is also analyzed.

The performance of ST, NWP and MOS models, together with the benchmark Persistence Model (PM), are compared. The ST model and the NWP model exhibit similar results. Nevertheless different sources of forecast errors between ST and NWP models are identified. The MOS model gives the best performance, improving the forecast of approximately 29% with respect to the PM.


*Keywords:* Solar irradiance; Forecast; Neural network; MOS; Ensemble

## 1. Introduction

The rapid growth of electricity produced by PV has introduced some criticisms into the grid due to the fluctuating nature of energy source which is dependent on meteorological conditions. Thus reliable forecast models are required for management and operation strategies. In particular the 24/72 h horizon forecast is essential for transmission scheduling and day ahead energy market. In Italy the PV energy production during the 2012 reached 18.9 GW h produced by 16.4 GW installed providing an average of 7% of yearly electrical consumption with a monthly peak of 9% (Statistical data from the Italian Manager of Electrical Services (GSE): "Solare Fotovoltaico - Rapporto Statistico 2012" and Statistical data from the Italian Manager of the National Electrical Transmission Grid: "produzione 2012").

The techniques to forecast the solar radiation or PV production on the 24/72 h horizon can be divided in three mains groups:

(1) Numerical Weather Prediction models (NWP)
(2) Statistical models (ST)
(3) Model Output Statistic (MOS)

* Corresponding author at: Department of Enterprise Engineering, University of Rome Tor Vergata, Via del Politecnico, 1 00133 Rome, Italy. Tel.: +39 0672597233.
*E-mail address:* cornaro@uniroma2.it (C. Cornaro).

http://dx.doi.org/10.1016/j.solener.2014.10.036
0038-092X/

**Nomenclature**

*Abbreviations*

| | |
|---|---|
| NWP | Numerical Weather Prediction model |
| ST | Statistical model |
| MOS | Model Output Statistic model |
| ANN | Artificial Neural Network |
| MLP NN | Multi-Layer Perceptron Neural Network |
| STNN | Developed Statistical model based on ANN |
| ECMWF-MOSNN | Developed Model Output Statistic model based on ANN and ECMWF NWP data |

*Variables and dimensions*

| | |
|---|---|
| $Gh$ | Global Horizontal Irradiance W/m$^2$ |
| Gpoa | Global Irradiance on the Plane of Array W/m$^2$ |
| NADV | Normalized Absolute Daily Variation of the solar radiation between the day ($t$) and the day ($t-1$) Dimensionless |
| NMHV | Normalized Maximum Hourly Variation of the solar irradiance Dimensionless |
| $K_t$ | Clearness index Dimensionless |
| $P_{\text{index}}$ | persistence index Dimensionless |
| Hh | daily horizontal irradiation W h/m$^2$ day |
| Ta | average daily temperature °C |
| OD | Ordinal Day number Dimensionless |
| $H_{\text{NWP}}$ | Daily irradiation forecast W h/m$^2$ day |
| $Gh_{cs}$ | Clear Sky Global Horizontal Irradiance (Ineichen/Perez model) W/m$^2$ |
| $P(X)$ | quantile trajectory W/m$^2$ |
| MSE | Mean Square Error (W/m$^2$)$^2$ |
| Corr | Pearson correlation index Dimensionless |
| RMSE | Root Mean Square Error W/m$^2$ |
| MAE | Mean Absolute Error W/m$^2$ |
| MBE | Mean Bias Error W/m$^2$ |
| Irmse | skill score with respect to the RMSE metric Dimensionless |
| $\Delta G_{\text{rel}}$ | mean width of the prediction intervals W/m$^2$ |

The numerical weather prediction models are essentially based on the numerical integration of coupled differential equations that describe the dynamics of the atmosphere and radiation transport mechanisms. The main advantage of these forecasting methods is that they are based on deterministic physical models. On the other hand, the main problem, in addition to the non-linearity of the used equations, is the spatial resolution of the integration grid: from 100 km (Global models) to few km (Mesoscale models) that is too wide with respect to the PV plants size. Inside the grid cell the cloud cover and aerosols are homogeneously fixed at their average values thus great errors could be induced both in the amount and in the time of the forecast irradiance on the PV site. Besides many NWP models have a temporal output interval greater than one hour while, as in this case, the hourly irradiance forecast is required. A comparison of main numerical weather prediction models (Global and Mesoscale models) for solar irradiance forecast in different locations can be found in Perez et al. (2010, 2013) and Muller and Remund (2010).

The statistical models are based on methods to reconstruct the relations between the hourly irradiance and past meteorological parameters (cloud ratio, air temperature, relative humidity, pressure, etc.) or past irradiance observations. The most used models for the one day horizon irradiance forecast are based on Artificial Neural Networks (ANN). Thus spatial and temporal resolution problems are overcome since these methods use ground measurements taken directly on the PV plant site with a temporal resolution less than one hour. On the other hand these methods are not able to provide a good forecast in unstable weather conditions since in these cases the correlation between the irradiance and the meteorological variables rapidly falls down. Several statistical models for 24 h forecast of global irradiance or PV power can be found in the literature presenting different ANN architectures. In Di Piazza et al. (2013) and Chaouachi et al. (2009) time series ANN Focused Time-Delay Neural Network (FTDNN) and the Nonlinear Autoregressive Network with exogenous inputs (NARX Network), are used. Radial Basis Function Neural Network (RBFNN) is also implemented in Chaouachi et al. (2009), while Multi-Layer Perceptron Neural Network (MLPNN) are developed in Chaouachi et al. (2009), Mellit and Massi (2010) and Voyant et al. (2013).

The model output statistics approach combines both NWP and ST models. The first one is used for the forecast while the second is used to correct the site effects through local ground measurements. The ST models are essentially used to down scale the irradiance forecast, reducing the systematic errors of the physical models. A variety of model output statistic models that use statistical post processing of the NWP output and stochastic learning techniques have been developed by various authors. In Perotto et al. (2013) a post-processing algorithm to correct the radiation schemes used by the WRF-NWP model is described. This is a physical based algorithm that improves the forecast of atmosphere water vapor profile. It uses regression coefficients that should be calculated from ground measurements. A statistical post-processing correction of the bias errors of the ECMWF-NWP data was proposed by Lorenz et al. (2009a). This seems to be