Available online at www.sciencedirect.com

ScienceDirect

Solar Energy 117 (2015) 99–113

SOLAR ENERGY

www.elsevier.com/locate/solener

Model output statistics cascade to improve day ahead solar irradiance forecast

M. Pierro [a], F. Bucci [a], C. Cornaro [a,*], E. Maggioni [b], A. Perotto [b], M. Pravettoni [c], F. Spada [b]

[a] Department of Enterprise Engineering, University of Rome Tor Vergata, Via del Politecnico, 1, 00133 Rome, Italy
[b] IDEAM srl, Via Frova 34, Cinisello Balsamo, MI, Italy
[c] Institute of Applied Sustainability to the Build Environment, University of Applied Sciences and Arts of Southern Switzerland, Canobbio CH-6952, Switzerland

Received 29 September 2014; received in revised form 9 April 2015; accepted 26 April 2015

Communicated by: Associate Editor Christian A. Gueymard

Abstract

In this paper a new hybrid Model Output Statistics (MOS), named MOS cascade, is developed to refine the day-ahead forecast of the global horizontal irradiance provided by the Weather Research and Forecast (WRF) model. The proposed approach is based on a sequence of two different MOS. The first, called MOSRH, is a new physically based algorithm, built to correct the treatment of humidity in the WRF radiation schemes. The second, called MOSNN, is based on artificial intelligence techniques and aims to correct the main systematic and learnable errors of the Numerical Weather Prediction output. The 1-day and 2-day forecast accuracies are analyzed via direct comparison with irradiance data measured in two sites, Rome and Lugano. The paper shows that a considerable reduction in error was achieved using MOSRH model and MOS cascade. The differences between the two sites are discussed in details. Finally, the results obtained are compared with the benchmark accuracy reached for the data available for the average climate in Southern Spain and Switzerland.

Keywords: Forecast; Solar irradiance; Photovoltaic; Model output statistics

1. Introduction

Intermittent renewable energy presents a big challenge specifically in the reliability of the electric power grid operation. The development and optimization of smart grids has come to the attention of many research centres and recently the European Commission, together with other national and international entities, has boosted its funding towards research in this field. Solar photovoltaic (PV) energy, along with other renewable energies, requires a precise daily or even hourly weather forecast to prevent both power failures and overloads and to manage the power share between the utilities.

At the same time, power produced from PV systems continues to increase, grabbing a bigger slice of the energy

* Corresponding author at: Department of Enterprise Engineering, University of Rome Tor Vergata, Via del Politecnico, 1, 00133 Rome, Italy. Tel./fax: +39 0672597233.
E-mail addresses: marco.pierro@gmail.com (M. Pierro), frabucci@gmail.com (F. Bucci), cornaro@uniroma2.it (C. Cornaro), enrico.maggioni@ideamweb.com (E. Maggioni), alessandro.perotto@ideamweb.com (A. Perotto), mauro.pravettoni@supsi.ch (M. Pravettoni).

http://dx.doi.org/10.1016/j.solener.2015.04.033
0038-092X/

market day by day. For these reasons vendors need the forecast of Global Horizontal Irradiance (GHI) to be as accurate as possible in order to calculate the energy production of the plants and for critical decision making in energy trading.

The most commonly used methods for the day ahead solar irradiance forecast (24/72 h horizon) are based on Numerical Weather Prediction models (NWP). These models are simulation software able to provide the numerical integration of the coupled differential equations describing the dynamics of the atmosphere and radiation transport mechanisms. The NWP models can be classified in two main categories: the global models, such as IFS (European Centre for Medium-range Weather Forecast), GFS (Environmental Modeling Centre) and GEM (Environment Canada Center), that simulate the whole earth atmosphere and the mesoscale models, like MM5 and WRF, which cover a smaller geographical area and have a higher resolution using initial boundary conditions extracted from the global models.

A comparison of main Numerical Weather Prediction models (global and mesoscale) for solar irradiance forecast in different locations can be found in Perez et al. (2010, 2013) and Muller and Remund (2010).

Even if the NWP models take all the main physical processes that govern the atmosphere into account, these processes are only an approximation. Moreover the non-linearity of the governing equations leads to a strong dependence on the initial conditions used for the numerical integration. This chaotic behaviour of the system is amplified even more because the initial conditions come from a heterogeneous and irregularly distributed observational network that can be too sparse in some regions of the world therefore providing insufficient data. In addition, the spatial resolution of the integration grid is too coarse with respect to the PV plant size and the temporal output interval may be greater than one hour. Furthermore, NWP models are not finely tuned for radiation forecast even considering the most recent sophisticated radiation schemes, which are implemented in current meteorological models. Indeed, the radiation schemes usually run independently from the governing equations of a NWP model making rigid assumptions for sub-grid cloud variability and showing little ability to translate quick changes in meteorological conditions, especially for cloud cover and humidity. As a result the NWP outputs still do not reach the desired accuracy and they are often affected by systematic errors.

For these reasons, outperforming forecasts could be achieved by post processing techniques called Model Output Statistics (MOS) that use ground measurements to remove bias and learnable errors from the NWP data.

Several MOS that use statistical analysis or stochastic learning techniques have been developed by various authors. A pure statistical post-processing correction of the bias errors of the NWP data of ECMWF was proposed by Lorenz et al. (2009a). This seems to be the best performing MOS model for global horizontal irradiance forecast (Lorenz et al., 2009b). Perez et al. (2007) developed a semi empirical model that correlates the NWP sky cover (provided by the National Digital Forecast Database, USA) with the global horizontal irradiance.

Other authors set up MOS models based on stochastic learning approaches. In Marquez and Coimbra (2011), Huang et al. (2010) and Yona et al. (2008) the Multi Layer Perceptron Neural Networks (MLPNN) were used, and in Wang et al. (2011) this Artificial Neural Network (ANN) architecture was coupled with the Gray Model. The Radial Basis Function Neural Network is developed in Chen et al. (2011) and Yona et al. (2008) while the non-linear autoregressive network with exogenous inputs in Cai et al. (2010). Recurrent Neural Networks and Diagonal Recurrent Wavelet Neural Networks were implemented in Cao and Lin (2008) and Yona et al. (2008).

An overview on solar irradiance and PV power forecast techniques can be found in Paulescu et al. (2013), Kleissl (2013) and "Photovoltaic and Solar Forecasting: State of the Art", IEA PVPS Task 14 (2013), while a complete study on solar radiation benchmarks is reported in Lorenz et al. (2009b), Beyer et al. (2009) and Traunmuller and Steinmaurer (2010).

One of the most used Mesoscale models is Weather Research and Forecast (WRF) – Advanced Research WRF (ARW) developed by the National Center of Atmospheric Research (NCAR), USA.

The aim of this work was to refine the irradiance forecast coming from WRF using a sequence of two different MOS.

The first post processing technique (MOSRH) is a new physical based algorithm that improves the forecast of the concentration of water vapour in the atmosphere. It uses regression coefficients that can be calculated from ground measurements, with the intent of reproducing more realistic absorption curves that take into account the entire vertical column of the atmosphere.

The second post processing approach (MOSNN) is based on stochastic learning algorithms that use Artificial Neural Networks ensemble to correct the NWP bias error. It was effectively used in a previous study (Cornaro et al., 2014) to refine the irradiance prediction of ECMWF reaching a considerable level of improvement in forecast accuracy. Thus the MOSNN has been used to refine the MOSRH output generating a hybrid MOS called "MOS cascade".

There are two main reasons for using this MOS cascade. Firstly, the MOSNN, when applied to the MOSRH output, can remove all the learnable errors not directly related to the forecast of humidity. Secondly, the MOSNN can directly provide the PV energy forecast, without the need of any further calculations, if trained with the power data produced by a PV plant. Indeed one of the advantages of the MOS based on stochastic learning techniques is that the variable predicted by the MOS (as PV energy) can be different from the input data provided by the NWP forecast (as GHI).

The MOS cascade was tested on different locations and periods considering the solar irradiance forecast on the horizon over one and two days. The forecast accuracy obtained with and without the data post processing was measured for the sites of Rome, during the years 2008–2012, and of Lugano during the years 2010–2011.

In Section 2 the data used (ground measurements and NWP forecast) are reported. In Sections 3 and 4, the basic features of the two MOS are described. In Section 5, the forecast accuracy reached by the MOSRH and by the MOS cascade is analyzed and discussed in detail.

2. Data

2.1. Experimental data from Rome

The experimental data used for Rome (as input to train and test the models) come from the ESTER outdoor Laboratory at the University of Rome "Tor Vergata" (Spena et al., 2008). The facility consists of a test unit for the outdoor monitoring of PV modules and a solar-weather unit (in operation since 2004) that collects solar irradiance data together with the typical weather variables such as air temperature, relative humidity etc. Five years (from 2008 till 2012) of GHI and air temperature measurements were used for the present work. Global horizontal irradiance was measured with a pyranometer Kipp&Zonen CM21, while for air temperature a Rotronic thermo hygrometer was used. Data are recorded by a CR100 data logger at a one-minute time rate. According to the WMO guidelines (Zahumenský, 2004), a data filtering procedure was applied to remove the physical inconsistencies in measurements that are caused by monitoring problems. After this operation, the variables' hourly and daily values were calculated.

2.2. Experimental data from Lugano

The experimental data used for Lugano (as input and to train and test the models) come from the outdoor test field at the University of Applied Sciences and Arts of Southern Switzerland (SUPSI). The facility consists of a PV test unit for the outdoor monitoring of PV modules and a solar-weather unit. GHI is measured with a Kipp&Zonen CMP21 pyranometer. The instruments are calibrated yearly within the framework of an international measurement campaign that has taken place since 2010 along with several international partners (Galleano et al., 2014). Data are recorded every minute from 5 AM to 7 PM, UTC + 1 by a HP Agilent data logger and then filtered daily to check for outliers or system failures. Ground measurements over two years, (2010–2011) were used for the present work.

2.3. Numerical weather prediction data from WRF model

2.3.1. WRF overview

The NWP model used is WRF–ARW version 3.4 developed by NCAR. The model is run operationally by the US National Weather Service and, being open-source and easily portable, it is widely utilized around the world for research and weather forecasts (Skamarock et al., 2008).

WRF was originally built as a Mesoscale or local area model (LAM), i.e. it simulates the weather in geographical areas smaller than the entire globe. The governing equations, describing the dynamic and thermodynamic evolution of the atmosphere (Pielke, 2002), are solved numerically on grid points over the area of interest that constitutes the domain of the model. The domain is three-dimensional. On the horizontal plane models may allow different map projections and utilize a specific grid structure. In the present case WRF-ARW employs the Arakawa C-grid (Arakawa and Lamb, 1977). The Arakawa C-grid is staggered, which means that not all governing equations are solved on the same grid points, dynamic physical quantities are solved on so-called "flux points", while mass related physical quantities are solved on "mass points". Vertically a terrain-following coordinate is used (Skamarock et al., 2008; Pielke, 2002). Vertical levels, the number of which can be chosen by the user, are denser near the Earth's surface and grow more distant from one another when moving towards the top of the atmosphere. The diagonal distance between two mass grid points on the same vertical level defines the model's horizontal resolution, which can be chosen by the user too. Local area models require initialization and contour data from a larger LAM or a global model. Initialization data provide the values of all the necessary atmospheric and surface fields at the beginning of the simulation; contour data provide the evolution of these fields outside of the LAM domain.

In the WRF model, the physical phenomena not solved explicitly by the governing equations are calculated by algorithms independent on the governing equation solver and "plugged in" through a physical interface (Skamarock et al., 2008). For this reason, algorithms for the calculations of the same physical quantities can be non-unique and, for this specific study, a choice between several radiation schemes is available.

The WRF model radiation schemes provide atmospheric heating due to radiative flux divergence and downward surface long and short radiation wavelengths for the ground heat budget. The long wavelength radiation (above 4000 nm) includes infrared or thermal radiation absorbed and emitted by gases and surfaces. The short wavelength radiation (between 300 and 4000 nm) includes visible and surrounding wavelengths that make up the solar spectrum. Absorption, reflection and scattering processes in the atmosphere and on the surface are simulated. Upward shortwave radiation is the reflection due to surface albedo (Skamarock et al., 2008). Shortwave radiation in the WRF model corresponds to GHI.

2.3.2. Model setup

Daily hindcasts are performed, running 24 h forecasts from 00 UTC to 00 UTC, with a 6 h spin up time, for all days between 2008 and 2012.

The domain of the model extends over Italy and Switzerland, including most of the Alpine region (Fig. 1).

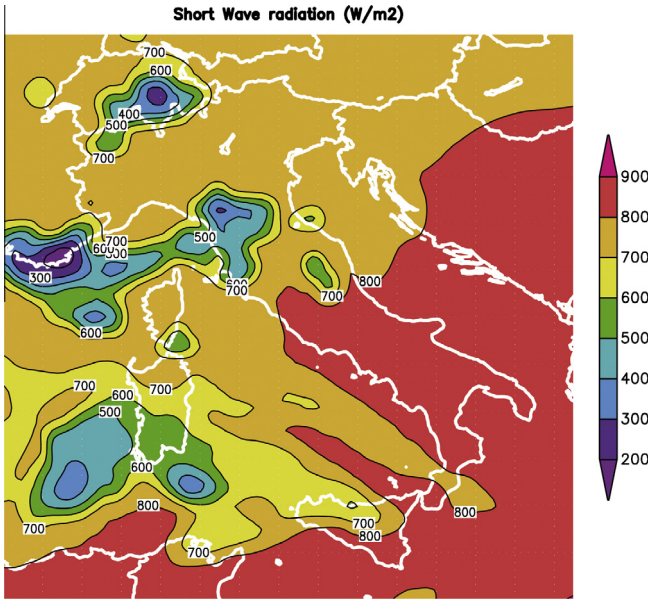

Fig. 1. WRF Model domain.

A variety of climate regions are situated inside the chosen domain.

The model horizontal resolution is 32 km, approximately the same of ECMWF – reanalysis data (ERA-40, 2004), whose performance was analyzed in a previous study (Cornaro et al., 2014).

Different initialization and contour data have been tested in order to analyse the impact on GHI forecasts.

In the first phase, NCEP-DOE reanalysis 2 data (NCEP-DOE AMIP-II Reanalysis, 2002) was used.

Subsequently, in order to simulate a process more similar to operational forecasting, the global NWP model GFS (Global Forecasting System) (Environmental Modeling Center, 2003) provided initial and contour data. A typical one-day-ahead operational forecast of GHI (from now on called "1-day forecast") and two-day-ahead operational forecast of GHI (from now on called "2-day forecast") were simulated. The results were analyzed separately.

For the WRF model a physics configuration suitable for operational forecasting was made, balancing accuracy in the results with computational efficiency.

For microphysics the "Eta microphysics scheme" was chosen as is also suggested for coarse resolutions (Rogers et al., 2001). For the surface layer the "MM5 similarity scheme" is used (Paulson, 1970), while as a surface scheme the 5 layer thermal diffusion scheme is used (Dudhia, 1996) when initial data derives from the NCEP-DOE reanalysis and the Noah Surface model (Tewari et al., 2004) when initial data derives from GFS. The YSU boundary layer scheme was chosen (Hong et al., 2006). The Kain-Fritsch cumulus scheme was implemented (Kain, 2004).

The Rapid Radiative Transfer Model (RRTMG) (Mlawer et al., 1997; Fu and Liou, 1992; Oreopoulos and Barker, 1999) long wave radiation scheme was chosen accordingly to the short wave radiation scheme as explained in the following paragraph.

In a preliminary phase, several short wave (SW) radiation schemes were tested. The RRTMG was chosen for its skill to describe sub grid cloud variability through a Monte-Carlo independent column approximation and due to its ability to distinguish near-infrared, visible and UV wavelengths. Radiation absorption and scattering by other gases and aerosols are calculated through various parameterizations. Concentration and chemical composition come from parametric values taken from pre-built tables. To obtain more accurate results on aerosol concentration and chemical composition of the atmosphere the WRF model could be coupled with an atmospheric chemistry model (WRF-Chem distribution) (Grell et al, 2005). Typically coupling is not performed for daily operational weather forecasts because of the computational time it requires and the lack of accurate source data needed for the entirety of the domain being considered. For these reasons, no coupled chemical model was run for these simulations.

Output model gridded data were interpolated to produce a point forecast for the location of Rome -Tor Vergata and Lugano for a direct intercomparison with ground measurements.

3. Physical post processing (MOSRH)

This section describes the Model Output Statistics (MOSRH) that provide the correction of the GHI forecast as calculated directly by the WRF model.

The RRTMG radiation scheme was chosen because it provides a more detailed output since it differentiates between direct and diffuse radiation and provides UV, visible and near infrared wavelengths as distinct output fields. Nevertheless, GHI forecasts are affected by elevated errors, which in part derive from the unrefined process of damping GHI due to the humidity present in the atmosphere column.

The majority of radiation schemes used by the WRF model compute the GHI at the bottom of the atmosphere by damping the top of the atmosphere GHI by a factor proportional to the highest relative humidity value on a specific level in the atmospheric column (Lara-Fanego et al., 2011).

The main objective of MOSRH is to improve the treatment of the amount of humidity in the atmosphere used by the radiation scheme. Besides humidity, other factors are important to a good forecast of GHI. In this study, WRF was not coupled with a chemical model, limiting some capabilities of the radiation scheme. Since aerosols and gas concentrations are given by fixed tables that only take into consideration date and geography, the variation of diffuse and direct components due to these aspects is rigid and predefined.

Nevertheless, the errors in forecasting humidity and in the treatment of this variable by the radiation scheme have a direct impact on the GHI forecast accuracy.

Analysing the behaviour of the model's built-in radiation schemes, GHI experiences a considerable damping only if one or more vertical model levels present water vapour content at the saturation level. This procedure, which is a direct application of the physical laws, can be very problematic if applied in a meteorological model that describes the atmosphere creating a finite number of vertical layers and a finite spacing horizontal square grid. In cases where the forecast for relative humidity of the entire column is very close but not at the saturation point, the model GHI damping is very weak, whereas, generally in this situation in the real atmosphere there is almost certainly condensation of the water vapour at some height, with a consequent GHI damping at the surface (Houghton, 2002).

The idea behind MOSRH is based on the fact that GHI, through the atmosphere, experiences a damping at every layer with a non-negligible value of water molecules in the liquid phase. This situation occurs when relative humidity is near or at the saturation level.

Because of the NWP limitations described above, MOSRH does not only take into consideration vertical levels where relative humidity saturates but all levels where relative humidity is greater than a predefined threshold. This value is not unique and may be dependent on the altitude. A value for every vertical level was calculated empirically: a linear regression that took into account the entire time interval of this study (2008–2012) sought the values that minimized the error of forecast GHI compared to observational data for both sites.

MOSRH was built to avoid the "on/off switch" of typical radiation schemes, which behaves like a step function in damping GHI i.e., saturation equals almost complete absorption and below saturation there is no absorption. In the MOSRH approach, damping increases linearly for the levels where relative humidity is higher than its threshold to reach complete absorption only when relative humidity reaches 100%. Subsequently the impact of damping is weighed accordingly to the quantity of humidity in the single vertical level, now independently from their altitude. The weights decrease exponentially with decreasing humidity and are derived from the application of the Beer-Lambert Law (Houghton, 2002). Subsequently the weights are normalized in order to obtain a value between 0 and 100, similar to the calculation of total cloud cover percentage. The result was named Pseudo Cloud Cover (PCC).

In synthesis: the integral produces a value proportional to the "damping capacity" of the atmosphere that can be used for GHI correction.

Two of the key problems in GHI forecasting were considered by MOSRH: not only saturation values of water vapour are taken into consideration, but all levels with relative humidity over a certain threshold, and not only the level with the highest water vapour content is taken into account, but all the atmospheric levels.

MOSRH consists of three steps.

During the first step the integral of relative humidity values for an atmospheric column thus the PCC is calculated by Eq. (1).

$$\mathrm{PCC} = \frac{\sum_j (\mathrm{RH}_j w_j)}{\sum_j (w_j)} \tag{1}$$

where RH_j is the relative humidity value of level j and the sum is calculated from bottom to the top of the atmosphere; w_j is the weight of the level j and is equal to zero for the levels below their specific threshold.

During the second step GHI is computed at the bottom of the atmosphere. This is done by damping the forecast clear-sky GHI by a value proportional to PCC. Clear-sky GHI is the GHI radiation that reaches the surface considering the absorption and scattering by aerosols and gases other than water vapour, as calculated by the chosen radiation scheme for clear-sky days.

Various types of regressions were tested during the study. Eq. (2) was chosen for the simplicity and linearity of the formula and because it suits observational data better. The resulting corrected GHI at the bottom of the atmosphere is then given by Eq. (2).

$$\mathrm{GHI}^f = \mathrm{dGHI}_{\mathrm{cs}}(1 - \mathrm{aPCC}^b) + c \tag{2}$$

$\mathrm{GHI}_{\mathrm{cs}}$ is the clear-sky GHI at the bottom of the atmosphere, a, b, c and d are the parameters of the regression.

The parameter values are different for different geographical locations. As a result, two sets of coefficients, for Rome and Lugano, were calculated.

Eq. (2) is a general equation that includes two different cases that exclude one another, based on the forecast meteorological conditions. Clear sky conditions are defined by $\mathrm{PCC} < 0.05$, in this case a systematic overestimation of GHI was observed, probably due to the oversimplification in evaluating aerosols and other damping substances in the atmosphere. To remove this bias coefficients a and b are fixed to the values of 0 and 1 respectively, while c and d are provided by the regression with observational data from Rome and Lugano, within the time interval considered (Table 3.1). In the second, PCC is greater or equal to 0.05 i.e. GHI is damped by cloud cover, coefficients c and d are fixed respectively to 0 and 1, while a and b are provided by the regression with observational data from Rome and Lugano, within the time interval considered (Table 3.1).

Table 3.1
Coefficient values for both sites.

		a	b	c	d
Rome	$\mathrm{PCC} < 0.05$	0	1	9.68	0.922
	$\mathrm{PCC} > 0.05$	0.541	0.692	0	1
Lugano	$\mathrm{PCC} < 0.05$	0	1	10.2	0.932
	$\mathrm{PCC} > 0.05$	0.471	0.562	0	1

4. Stochastic learning post processing (MOSNN)

The MOSNN is a refinement technique of the NWP data that uses ANN and it was implemented by the ESTER laboratory (Cornaro et al., 2014).

ANN is a mathematical model that invokes the structure of the biological neural connections. The model is made up of a group of neurons, and it processes information using a connectionist approach to computation.

ANN has the ability to imitate natural intelligence in its learning from existing sample data. The algorithm learns from sample data by constructing input-output connections without explicit analytical expressions of the model equation and for this reason it is typically used to model complex relationships between inputs and outputs. This technique is often employed in solving forecasting problems; an extensive review can be found in (Zhang et al., 1998). For what concerns solar radiation forecast using ANN, a review can be found in (Mellit, 2008).

MOSNN uses Multi-Layer Perceptron Neural Network architecture (MLPNN), with one hidden layer.

This architecture uses meteorological inputs and NWP prediction to forecast the one or two days ahead hourly irradiance as shown in Eq. (3).

$$[\mathrm{GHI}_1(t+x),\ldots\mathrm{GHI}_{24}(t+x)] = f(\text{past meteorological parameters, NWPdata}) \quad (3)$$

where x is the forecast horizon.

The inputs come from past local measurements and from NWP forecasting data.

Fig. 2A synthetically reports the adopted MLPNN, where $\mathbf{P^1}$ is the input data vector; $\mathbf{IW^1}$, $\mathbf{LW^1}$ and $\mathbf{b^i}$ are the input and layer weights matrix and the bias vectors that should be calculated by a training and validation procedure; $\mathbf{S^1}$ is the number of neurons in the hidden layer that should be optimized; $\mathbf{a^i}$ are the output vectors and $\mathbf{f^i}$ are the chosen transfer functions (tansigmoid and pureline).

It has to be considered that the performance of the ANN is strongly dependent on the internal structure of the net, on the training method adopted and on the input variable chosen.

The input variables chosen for the considered built neural network are listed in Table 4.1, while the ANN output is the 1 day or 2 day ahead hourly solar horizontal global irradiance, as reported in Fig. 2B.

The optimization procedure used to generate the MOSNN is described in detail in (Cornaro et al., 2014) and can be briefly summarized as follows.

The network under examination was implemented using the ANN Matlab toolbox. The Levenberg-Marquardt technique was used to train the chosen net coupled with the repeated random sample validation procedure. The net structure was identified through an optimization process that provided the best number of neurons in the hidden layer (S) through the Mean Square Error (MSE) minimization procedure. The optimized net consisted of 6 input neurons, 24 output neurons and 6–14 neurons in the hidden layer depending on the year of data chosen for training and validation.

Moreover, instead of choosing an optimal network, the technique of the ANN Ensemble was adopted. The ANN was repeatedly trained with one year of data randomly sorting 60% of data for training and 40% for validation generating over 500 Neural Networks. The ensemble was originated choosing the ANN with the MSE lower than the average MSE of the 500 networks. The qualified ensemble of 300–350 networks was used to calculate the quantile trajectories P(X): for instance, the P(50) point is the value GHI(50) such that the probability of GHI $\leqslant$ GHI(50) is equal to 50%. The P(60) trajectory was used as GHI forecast to evaluate the model accuracy, since the ensemble average and the P(50) tend to underestimate the irradiance. Thus the P(60) provided a further bias reduction. This technique is known as Probabilistic Forecast and it is commonly used for wind prediction.

The ANN test was made using another year of data, never seen by the net. For Rome, five years of data were used. In particular, the ANN was trained using the 2008 data for testing 2009, 2010, 2011 and 2012 while 2011 was used to test the year 2008. For Lugano, the ANN trained on the 2010 was used to forecast the 2011 and vice versa.

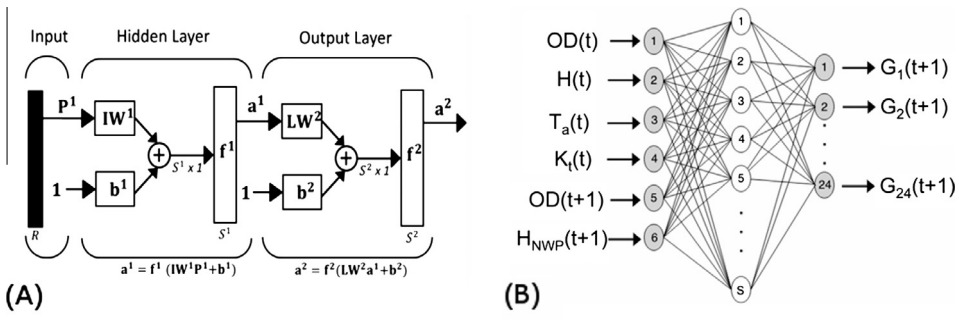

Fig. 2. (A) Sketch of the MultiLayer Perceptron Neural Network (MLPNN) architecture, and (B) Sketch of the input and output variables used for the 1 day forecast.

Table 4.1
Input variables given to the NNMOS, HNWP is referred to the next day.

Name	Symbol	Units
Ordinal day	OD	–
Daily Irradiation	H	Wh/m^2
Average daily temperature	Ta	°C
Clearness index	Kt	–
Daily irradiation forecast	HNWP	Wh/m^2

In Fig. 3 an example of the P(60) trajectory together with the quantile trajectories between the P(5) and P(100) curves is shown for a sample of days. It can be noted as the prediction interval delimited by P(5) and P(100) is larger where a major forecast uncertainty is present (overcast days, for example).

5. Results

Verification of model data and MOS data was carried out using basic statistical techniques, such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Bias Error (MBE) and the skill score with respect to a reference model, as defined in Table 5.1. Moreover, in this work the following Persistence Model (PM) was adopted as main reference (Eq. (4)).

$$[\mathrm{GHI}_1(d+1),\ldots\mathrm{GHI}_{24}(d+1)] = \langle K_t^*(d)\rangle[\mathrm{GHI}_{cs1}(d+1),\ldots\mathrm{GHI}_{cs24}(d+1)] \quad (4)$$

where

$[\mathrm{GHI}_1(d+1),\ldots\mathrm{GHI}_{24}(d+1)]$ = hourly global horizontal irradiance of the day $(d+1)$.

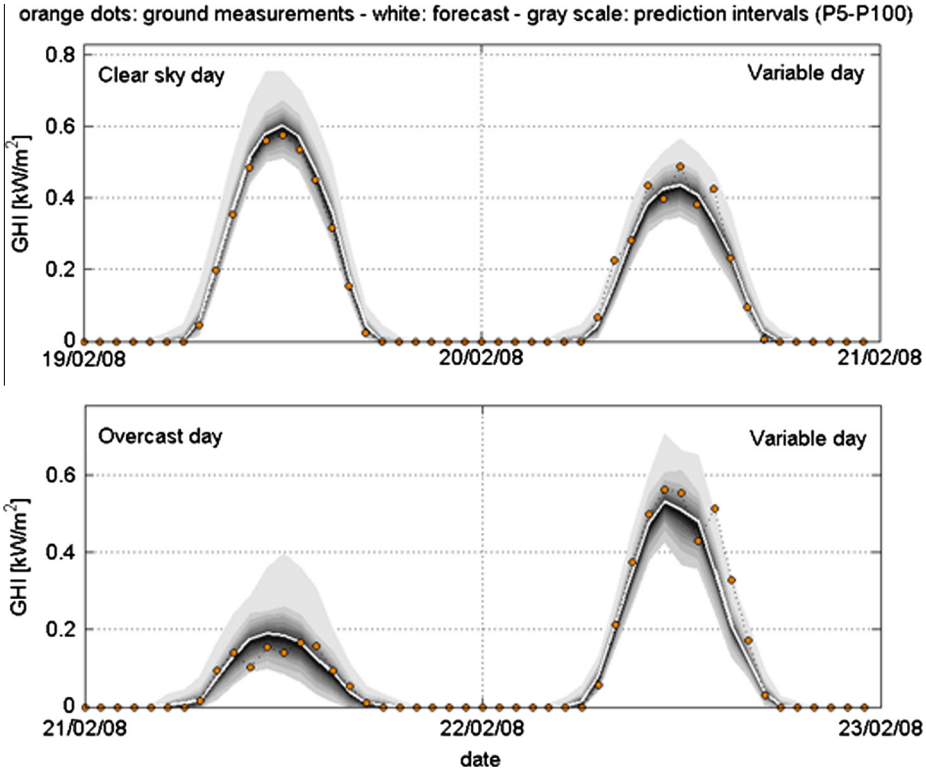

Fig. 3. Prediction intervals from P(5) to P(100) (gray) – forecast P(60) (white) and ground measurements (dots).

Table 5.1
Performance indexes used for model comparison.

Name	Acronym and formulae
Root Mean Square Error	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(GHI_i^m - GHI_i^f)^2}{n}} W/m^2$
Mean Absolute Error	$MAE = \frac{\sum_{i=1}^{n}\lvert GHI_i^m - GHI_i^f\rvert}{n} W/m^2$
Mean Bias Error	$MBE = \frac{\sum_{i=1}^{n}(GHI_i^f - GHI_i^m)}{n} W/m^2$
RMSE Skill Score or Improvement	$Skill\ Score = 100\left(\frac{RMSE(reference\ \ model) - RMSE(test\ \ model)}{RMSE(reference\ \ model)}\right)$ [%]

GHI_i^m = measured hourly global horizontal irradiance W/m^2.
GHI_i^f = forecast hourly global horizontal irradiance W/m^2.

$[GHI_{cs1}(d+1), \ldots GHI_{cs24}(d+1)]$ = hourly clear sky global horizontal irradiance of the day $(d+1)$ calculated using the clear sky model (Ineichen and Perez, 2002; Reno et al., 2012).

$K_t^*(d)$ = daily average of the hourly clear sky index of the day (d) defined as the ratio of the measured global horizontal irradiance and the global horizontal irradiance calculated using the clear sky model.

The accuracy of the PM can be considered as a measurement of the difficulty of the irradiance forecast for a specific site and period so that the skill score with respect to the PM states the quality of adopted forecast models.

All the metrics were calculated over one year of data. Night time hours were eliminated by only considering global horizontal irradiance as greater than zero.

Table 5.2
Results of the various forecast approaches for Rome.

Rome		2008	2009	2010	2011	2012
MAE (W/m^2) *(NMAE (%))*	Raw	92	82	90	69	75
		(26.9)	*(23.4)*	*(26.7)*	*(19.0)*	*(19.4)*
1 day	Reanalysis	96	84	92	68	77
		(28.1)	*(23.9)*	*(27.3)*	*(18.9)*	*(20.0)*
	Persistence	90	84	95	77	82
		(26.7)	*(24.5)*	*(30.0)*	*(22.0)*	*(23.6)*
	MOSRH	64	61	69	60	58
		(19.1)	*(17.9)*	*(21.7)*	*(17.1)*	*(16.7)*
	MOSNN(MOSRH)	67	63	71	60	57
		(20.0)	*(18.3)*	*(22.3)*	*(17.1)*	*(16.3)*
MAE (W/m^2) *(NMAE (%))*	Raw	94	85	93	69	78
		(27.3)	*(24.0)*	*(27.6)*	*(18.8)*	*(20.3)*
2 day	Persistence	98	92	110	84	89
		(28.9)	*(26.7)*	*(34.7)*	*(24.0)*	*(25.6)*
	MOSRH	67	64	72	58	60
		(19.7)	*(18.6)*	*(22.7)*	*(16.6)*	*(17.4)*
	MOSNN(MOSRH)	67	66	73	61	62
		(19.6)	*(19.0)*	*(22.9)*	*(17.5)*	*(17.9)*
RMSE (W/m^2) *(NRMSE (%))*	Raw	153	140	149	119	131
		(44.9)	*(39.7)*	*(44.0)*	*(32.8)*	*(33.9)*
1 day	Reanalysis	157	143	153	121	138
		(46.0)	*(40.6)*	*(45.3)*	*(33.3)*	*(35.9)*
	Persistence	148	136	149	130	140
		(44.0)	*(32.0)*	*(47.1)*	*(37.1)*	*(40.4)*
	MOSRH	104	100	105	96	92
		(31.0)	*(29.1)*	*(33.0)*	*(27.4)*	*(26.4)*
	MOSNN(MOSRH)	115	102	107	96	97
		(34.1)	*(29.7)*	*(33.7)*	*(27.5)*	*(28.0)*
RMSE (W/m^2) *(NRMSE (%))*	Raw	155	142	154	118	138
		(45.3)	*(40.3)*	*(45.5)*	*(32.2)*	*(35.6)*
2 day	Persistence	159	150	168	139	149
		(46.9)	*(43.5)*	*(53.0)*	*(39.7)*	*(42.9)*
	MOSRH	108	106	110	99	95
		(31.7)	*(30.6)*	*(34.7)*	*(28.2)*	*(27.5)*
	MOSNN(MOSRH)	113	107	110	101	96
		(33.3)	*(30.9)*	*(34.7)*	*(28.9)*	*(27.8)*
MBE (W/m^2) *(NMBE (%))*	Raw	80	64	70	52	64
		(23.4)	*(18.2)*	*(20.0)*	*(14.0)*	*(16.2)*
1 day	Reanalysis	78	66	75	47	61
		(22.8)	*(18.2)*	*(22.0)*	*(12.0)*	*(15.0)*
	MOSRH	15	5	13	−5	0
		(4.4)	*(1.4)*	*(4.1)*	*(−1.4)*	*(0)*
	MOSNN(MOSRH)	4	12	20	0	22
		(1.2)	*(3.5)*	*(6.3)*	*(0)*	*(6.3)*
MBE (W/m^2) *(NMBE (%))*	Raw	77	66	73	51	67
		(22.5)	*(18.7)*	*(21.6)*	*(14.0)*	*(17.3)*
2 day	MOSRH	12	7	10	−2	−1
		(3.5)	*(2.0)*	*(3.2)*	*(−0.6)*	*(−0.3)*
	MOSNN(MOSRH)	24	7	5	4	4
		(7.1)	*(2.0)*	*(1.6)*	*(1.1)*	*(1.2)*

It should be remarked that the accuracy values reported in the following tables cannot be directly compared since the PM accuracies are different between the different sites and years. However, for the comparison of the results it is possible to use the Skill Score calculated with respect to the reference model, that is much less site and time dependent.

All the results obtained are reported in Table 5.2 for Rome and 5.3 for Lugano. The tables also show relative errors, which are obtained dividing by the yearly average horizontal irradiance.

5.1. Direct WRF model output results (Raw data)

During a first data analysis, the accuracy of the d*irect WRF forecast initialized with GFS + 1 day and + 2 day data* (*R*aw 1-day, Raw 2-days) together with the *WRF forecast initialized* with NCEP-DOE II Reanalysis data (Reanalysis 1-day) were calculated and compared. Direct model output provides information on the quality of the operational forecast obtainable by WRF initialized with GFS, while WRF initialized with Reanalysis data provides information on the forecast dependence on the time horizon and grid resolution of the boundary conditions used for the forecast. During the run, the Reanalysis model assimilates the observed data at regular intervals, limiting errors for meteorological variables, but at the same time, both the vertical and horizontal grid resolutions are sensibly lower and physical schemes are simplified compared to GFS.

Fig. 4 shows the skill score with respect to the RMSE of PM obtained by the d*irect WRF forecast* (*R*aw 1-day, Raw 2-days) and by the Reanalysis data (Reanalysis 1-day) for both the sites under investigation. For Rome, the three data sets achieve no or very little improvement with respect to the persistence model, with a skill score between −6% and 8% for the 1 day direct model forecast and Reanalysis and between 2.5% and 15% for the 2 days Direct model forecast. Table 5.2 shows that the Direct WRF output and the Reanalysis have an accuracy range of 69–93 W/m^2 in MAE and 118–157 W/m^2 in RMSE comparable to the persistence that exhibits 77–110 W/m^2 in MAE and 130–168 W/m^2 in RMSE.

The 2-day forecast skill score is higher than that of the 1-day forecast, only because the 1-day accuracy of the PM is higher than the 2-day PM accuracy.

What was even less to be expected was that the forecast initialized with Reanalysis data showed slightly higher errors than the direct model output. Thus, for the 1–2 day GHI forecast the greater resolution of the initial and contour data and more complex physics seems to have more importance than boundary conditions derived from observational data.

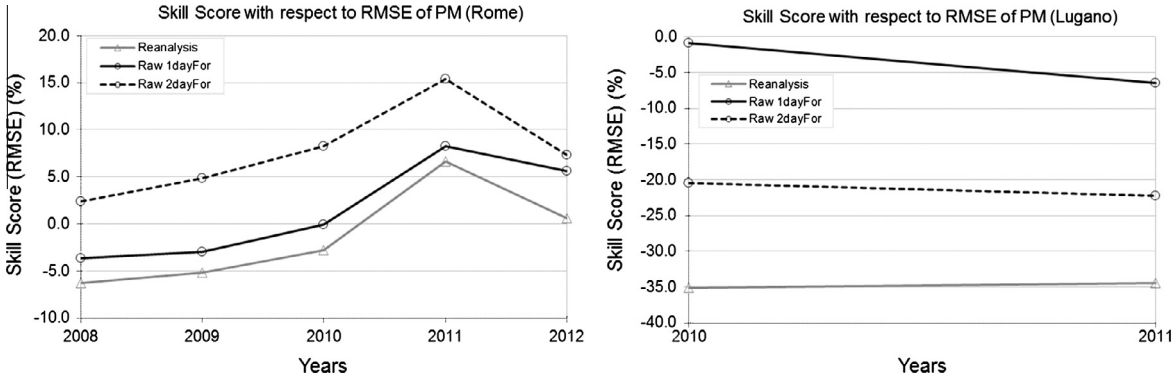

Fig. 4. Skill score with respect to the RMSE of PM obtained for the *1 and 2 day Direct WRF forecast* (Raw 1-day, Raw 2-days) and for the Reanalysis for both the Rome and Lugano sites.

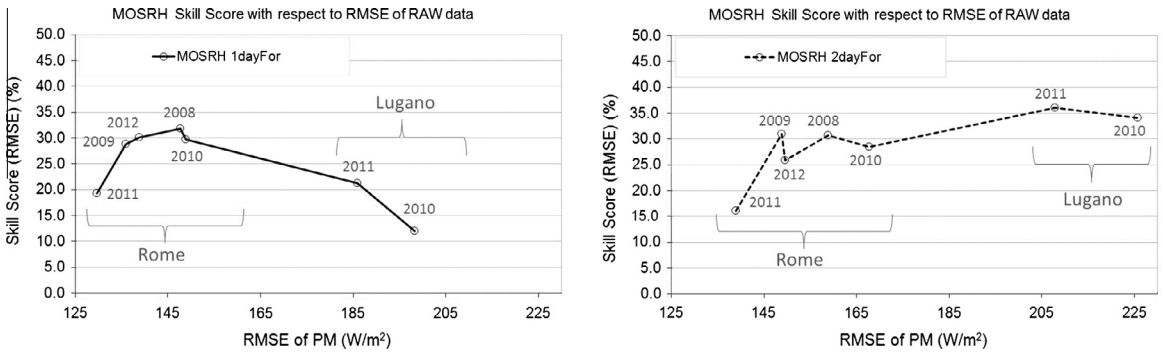

Fig. 5. Skill score with respect to the RMSE of RAW data of the *1 day (left) and 2 day (right) MOSRH* for Rome and Lugano.

Moreover, the yearly MBE between 51 and 80 W/m^2 (Table 5.2) underlines the evident overestimation of the direct model output GHI for all three data cases. The bias seems dependent on the model and radiation scheme used, but not on the initialization and contour data. Lugano has greater meteorological variability: weather variations are more frequent and rapid, hence a much greater error in persistence data. In this case all the three data sets have worse performance than the PM, obtaining a negative skill score for almost all the cases studied. The raw 1-day forecast obtains a yearly average MAE and RMSE around 133–199 W/m^2 while the PM achieves a MAE and RMSE of 130–210 W/m^2. The raw 2 day forecast shows a yearly average MAE and RMSE of 193–263 W/m^2 greater than 153–215 W/m^2 of the Persistence model. Thus in this case, differently from Rome, the 1-day forecast is much better than the 2-day forecast mostly because the high variability of weather conditions makes longer forecast more difficult. As for Rome, the performance shows a greater dependence on initialization and contour conditions. Also in this case, the direct model forecast is affected by greater systematic errors with a MBE between 107 and 171 W/m^2.

Tables 5.4 and 5.5 show the capability of the models to provide "good forecast days" (MAE $<$ 10 W/m^2) and "bad forecast days" (MAE $>$ 80 W/m^2). As for Rome, having more frequent longer periods with stable weather, the persistence model has the highest number of "good forecast days", while the number of days with errors greater than 80 W/m^2 is comparable to the one obtained by the three data sets (Table 5.4). In Lugano however, the WRF model initialized with GFS data presents notable errors for every type of weather, providing a lower number of "good forecast days" and the highest number of "bad forecast days" with respect to the PM (Table 5.5).

5.2. Post processing results (MOSRH and MOS cascade)

The physical based MOSRH, described in Section 3 was built to increase the very poor forecast capability of the direct WRF/GFS model.

Fig. 5 shows the skill score of the MOSRH with respect to the RMSE of Direct model output (RAW data) plotted versus the RMSE of the Persistence model.

For Rome, the MOSRH obtains a skill score with respect to the RMSE of the raw data between 16% and 32% depending on forecast horizon and on the persistence of the weather conditions. This is mainly due to the yearly mean bias correction, from 51–80 W/m^2 to 1–15 W/m^2 (see Table 5.2). The MOS increases the forecast performance on both sunny and cloudy days, since the number of days with MAE $<$ 10 W/m^2 increased by 30–50%, while the number of days with MAE $>$ 80 W/m^2 decreased by the same amount (see Table 5.2).

For Lugano, the MOSRH improves the 1-day forecast Direct model accuracy by 12–21% and the 2-day forecast accuracy by 35% (see Fig. 5). This difference in skill score is mainly due to the very low capability of the Direct WRF/GFS model in providing a 2-day forecast in this climatic conditions (see Fig. 4). Due to the greater meteorological variability, MOSRH technique shows a better performance in days with variable and perturbed weather, while no significant improvement is noted on mostly sunny days. Indeed, days with MAE $>$ 80 W/m^2 are reduced by about 20–30% with respect to the raw data, but no

Table 5.3
Results of the various forecast approaches for Lugano.

Lugano		2010	2011
MAE (W/m^2)	Raw	134	132
(NMAE (%))		*(39.4)*	*(36.1)*
1 day	Reanalysis	196	181
		(57.4)	*(49.3)*
	Persistence	137	122
		(40.2)	*(32.7)*
	MOSRH	126	106
		(37.1)	*(28.3)*
	MOSNN(MOSRH)	103	91
		(30.2)	*(24.4)*
MAE (W/m^2)	Raw	200	185
(NMAE (%))		*(58.4)*	*(50.4)*
2 day	Persistence	164	142
		(48.2)	*(38.0)*
	MOSRH	128	109
		(37.5)	*(29.2)*
	MOSNN(MOSRH)	112	103
		(33.0)	*(27.6)*
RMSE (W/m^2)	Raw	200	198
(NRMSE (%))		*(58.5)*	*(54.0)*
1 day	Reanalysis	268	250
		(78.3)	*(68.1)*
	Persistence	221	201
		(64.9)	*(53.9)*
	MOSRH	176	156
		(51.7)	*(41.8)*
	MOSNN(MOSRH)	142	131
		(41.8)	*(35.1)*
RMSE (W/m^2)	Raw	272	254
(NRMSE (%))		*(79.4)*	*(69.2)*
2 day	Persistence	226	208
		(66.4)	*(55.7)*
	MOSRH	179	162
		(52.6)	*(43.5)*
	MOSNN(MOSRH)	152	146
		(44.7)	*(39.1)*
MBE (W/m^2)	Raw	107	115
(NMBE (%))		*(31.2)*	*(31.0)*
1 day	Reanalysis	174	156
		(50.1)	*(42.0)*
	MOSRH	84	86
		(24.6)	*(17.6)*
	MOSNN(MOSRH)	15	5
		(4.5)	*(1.3)*
MBE (W/m^2)	Raw	171	163
(NMBE (%))		*(49.9)*	*(44.4)*
2 day	MOSRH	85	67
		(25.0)	*(17.9)*
	MOSNN(MOSRH)	22	−4
		(6.5)	*(−1.2)*

substantial change in the number of days with MAE < 10 W/m^2 is noticed (see Table 5.5).

The stochastic learning MOSNN, described in Section 4, was applied to the output of the MOSRH to understand if it was possible to reach further improvement in the forecast accuracy.

Fig. 6 compares the skill score with respect to the PM model, achieved by the MOSRH with the skill score obtained by the MOS cascade, named MOSNN (MOSRH), for the two different sites under investigation.

For Rome, the first post processing MOSRH produces a skill score in MAE and RMSE between 25% and 30% for the 1-day forecast and between 30% and 35% for the 2-day forecast. For the Lugano site, similar accuracy can be obtained when the MOS cascade alone is used. Indeed in this case, the MOSNN(MOSRH) allows us to reach a Persistence skill score of around 25% in MAE and 30% in RMSE for the 1-day forecast and around 30% in MAE and RMSE for the 2-day forecast.

Fig. 7 shows the skill score of the MOS cascade with respect to the RMSE of MOSRH output plotted versus the RMSE of the Persistence model. It appears that the MOS cascade achieves the worst performance when applied to the Rome post-processed irradiance data while it produces a considerable skill score improvement in the Lugano climatic conditions.

For Rome the cascade could reduce the MOSRH accuracy upto 10% while for Lugano it enhances the MOSRH performance to 15–20% for the 1-day forecast and upto 10–15% for the 2-day forecast.

This means that for the Rome climatic situation the main systematic error of the WRF model is related to the water vapour treatment in the built-in radiation schemes. Since this error was removed by the physical post-processing method (MOSRH), no further bias reduction could be achieved with the MOS cascade (MOSNN(MOSRH)). On the contrary, for Lugano, the very low persistence of the irradiance conditions is also due to the complex orography, so that the MOSRH water vapour modeling might not be reliable enough for the site and other sources of bias could have a great impact on the WRF forecast. Thus the Artificial Neural Network approach (MOSNN) could effectively correct the residual systematic errors resulting in a considerable improvement of the forecast accuracy.

One the whole, for the one day forecast for Rome, the MOSRH post processing reduces the MAE and RMSE of Direct WRF/GFS model from (81.6–138.4) W/m^2 to (62.4–99) W/m^2 and for the two day forecast from (83.8–141.4) W/m^2 to (64.3–103) W/m^2. For the one day forecast for Lugano, the MOS cascade reduces the same errors from (133–199) W/m^2 to (97–137) W/m^2 and for the two day forecast from (192.5–263) W/m^2 to (108–149) W/m^2.

It is should be noted that, for Rome, the MOSNN approach, when directly used to correct WRF raw data, was quite underperforming with respect to the MOSRH approach. On the other hand, in a previous work (Cornaro et al., 2014), the MOSNN post-processing was used to refine the ECMWF – ERA40 irradiance forecast of the years 2008-2011 obtaining a RMSE of 106 W/m^2 and a MAE of 65 W/m^2. These values are very similar to the RMSE of 102 W/m^2 and the MAE of 64 W/m^2 obtained by the MOSRH correction of the WRF/GFS raw data over the same period. Nevertheless, in the case of ECMWF – ERA40, the main source of bias errors was identified to be the time resolution of NWP irradiance forecast (each 3 h instead of 1 h). Thus the MOSNN approach based on an Artificial Intelligence technique is much more effective in the correction of this kind of systematic errors than in removing the bias coming from a rough approximation of the water vapour concentration.

5.3. Comparison with the benchmarks

To better understand the quality of the irradiance forecast provided by a specific technique, it is useful to compare the results discussed above with some benchmark accuracy values.

Lorenz et al. (2009b) report the forecast performance obtained with different methods for different European countries providing a range of accuracy that can be used as a benchmark for the horizontal irradiance forecast on the horizon of 24/72 h. Unfortunately, for the results in this paper, there is no benchmark for Italian climatic

Table 5.4
Numbers of "good days" and "bad days" in Rome.

Rome		2008	2009	2010	2011	2012
Days with Mae > 80 W/m^2	Persistence 1-day	61	47	49	30	35
	Reanalysis	65	48	63	33	37
	Raw 1-day	73	54	70	39	44
	Raw 2-day	69	62	74	39	43
	MOSRH 1-day	39	30	31	17	25
	MOSRH 2-day	40	44	35	18	28
Days with Mae < 10 W/m^2	Persistence 2-day	69	68	46	104	83
	Reanalysis	49	57	61	89	85
	Raw 1-day	43	49	50	90	78
	Raw 2-day	41	49	49	83	80
	MOSRH 1-day	78	80	63	106	106
	MOSRH 2-day	81	79	62	102	102

Table 5.5
Numbers of "good days" and "bad days" in Lugano.

Lugano		2010	2011
Days with Mae > 80 W/m^2	Persistence	111	72
	Reanalysis	142	95
	Raw 1-day	105	79
	Raw 2-day	144	105
	MOSRH 1-day	82	58
	MOSRH 2-day	119	85
Days with Mae < 10 W/m^2	Persistence	25	31
	Reanalysis	1	6
	Raw 1-day	10	12
	Raw 2-day	2	3
	MOSRH 1-day	8	20
	MOSRH 2-day	7	13

conditions. However, comparing the errors of the PM calculated by Lorenz et al. (2009b) with the persistence accuracy reported in Tables 5.2 and 5.3, it appears that Rome shows a solar irradiance variability that falls between those of Southern Spain and Switzerland. Thus, Tables 5.6 and 5.7 compare the benchmark accuracy range found by Lorenz et al. (2009b) for the Southern Spain and Swiss stations (in bold in the tables), with the performance obtained by the MOSRH and the MOSNN(MOSRH) cascade for Rome and Lugano (in bold in the tables). Regarding the station in Rome, the MOSRH gives a skill score based on PM perfectly inside the benchmark. In the case of the 1-day forecast, the MOSRH improves the PM accuracy of 29% that is inside the benchmark windows of 36–1% (Southern Spain) and 49–22% (Switzerland). In the

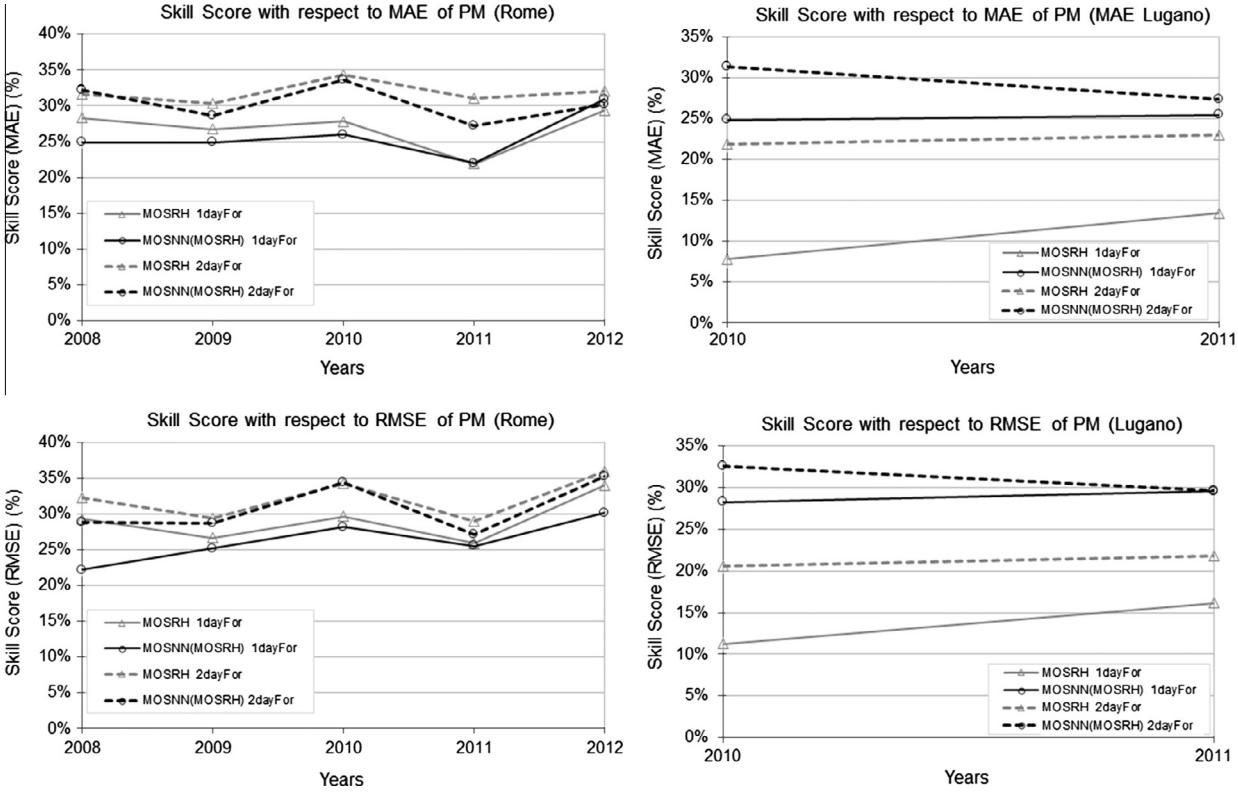

Fig. 6. Skill score with respect to the MAE and RMSE of PM for Rome and Lugano sites.

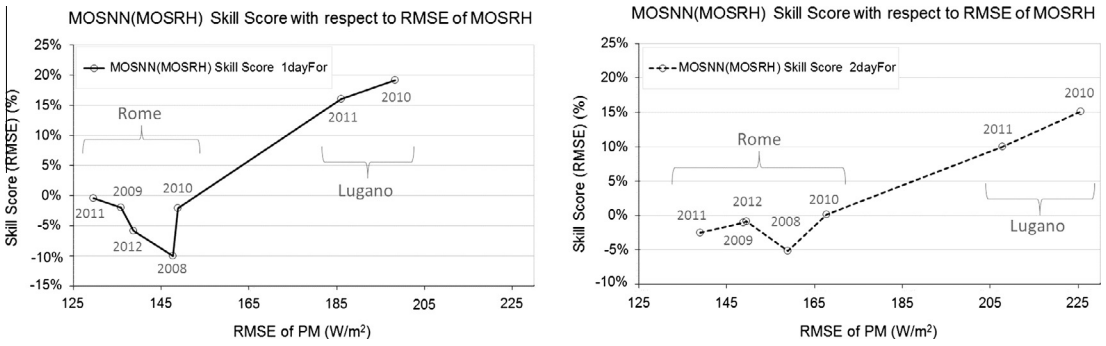

Fig. 7. Skill score of the MOS cascade with respect to the RMSE of MOSRH output plotted versus the RMSE of the Persistence model for 1 day forecast (left) and 2 day forecast (right).

Table 5.6
Performance indexes and improvement of MOSRH and MOSNN compared to the Rome benchmarks available.

1 day forecast		RMSE (W/m^2)	MAE (W/m^2)	PM skill score (RMSE)
Southern Spain ([a])	**Benchmark forecast models**	**81–124**	**48–79.8**	**36–1%**
		20.8–31.7%	*12.2–20.4%*	
	Persistence	125.9	65.0	
		32.1%	*16.6%*	
Rome ([b])	**MOSRH**	**99**	**62.4**	**29%**
		29%	*18%*	
	MOSNN(MOSRH)	**103**	**63**	**26%**
		31%	*19%*	
	Persistence	140.2	85.4	
		47%	*25%*	
Switzerland ([a])	**Benchmark forecast models**	**107–122**	**70–85**	**49–22%**
		39.6–45%	*25.8–31.5%*	
	Persistence	158.0	104.0	
		58.4%	*38.7%*	
Lugano ([c])	**MOSRH**	**166**	**116**	**14%**
		47%	*32%*	
	MOSNN(MOSRH)	**137**	**97**	**29%**
		38%	*27%*	
	Persistence	192.2	129.4	
		54%	*36%*	

[a] Average performance of different meteorological stations for the period July 2007–June 2008.
[b] Average performance over different years 2008–2012.
[c] Average performance over different years 2010–2011.

Table 5.7
Performance indexes and improvement of MOSRH and MOSNN compared to the Lugano benchmarks available.

2 day forecast		RMSE (W/m^2)	MAE (W/m^2)	PM skill score (RMSE)
Southern Spain ([a])	**Benchmark forecast models**	**83–144**	**49–94**	**41–11%**
		21.3–36.8%	12.6–23.9%	
	Persistence	140.0	75.0	
		35.8%	19.1%	
Rome ([b])	**MOSRH**	**103**	**64.3**	**32%**
		30%	19%	
	MOSNN(MOSRH)	**105**	**66**	**31%**
		31%	19%	
	Persistence	152.8	94.5	
		51%	28%	
Switzerland ([a])	**Benchmark forecast models**	**113–125**	**74–87**	**35–28%**
		41.8–46.3%	27.4–32.4%	
	Persistence	173.0	116.0	
		64.0%	42.9%	
Lugano ([c])	**MOSRH**	**171**	**118**	**21%**
		48%	33%	
	MOSNN(MOSRH)	**149**	**108**	**31%**
		42%	30%	
	Persistence	216.8	152.6	
		61%	43%	

[a] Average performance of different meteorological stations for the period July 2007–June 2008.
[b] Average performance over different years 2008–2012.
[c] Average performance over different years 2010–2011.

2-day forecast the MOSRH approach obtains an improvement of 32% with respect to the reference ranges of 41–11% (Southern Spain) and 35–28% (Switzerland). Regarding Lugano, it should be pointed out that the irradiance variability of this site is much higher than the Switzerland average variability. Indeed the RMSE of the PM in Lugano is between 20% and 25% higher than the PM RMSE averaged on the sixteen Swiss stations. Nevertheless, also in this case study, the PM skill score obtained using the MOS cascade is inside the

Switzerland benchmark. For the 1-day forecast, the MOSNN(MOSRH) gives an improvement of 29% while for the 2-day forecast it achieves a skill score of 31% taking in to account the reference ranges of 49–22% (1-day horizon) and 35–28% (2-day horizon).

6. Conclusions

Two different Model Output Statistics (MOS) were built to refine the day-ahead forecast of the GHI provided by the WRF/GFS model.

The first MOS, namely MOSRH, uses a physically based algorithm to correct the treatment of humidity in the WRF radiation schemes. The second MOS, called MOSNN, is based on Artificial Intelligence techniques and aims to correct the main systematic and learnable errors of the Numerical Weather Prediction output.

The 1-day and 2-day forecast accuracies reached with and without the MOS approaches were analyzed for different years and locations: years 2008–2012 in Rome and 2010–2011 in Lugano.

It appears that the direct WRF/GFS model output, for all the cases studied, exhibits no or very poor performance improvement with respect to the accuracy of the persistence model. Indeed, for Rome and Lugano, a yearly average RMSE of 138.4–199 W/m^2 (1-day forecast) and 141.4–263 W/m^2 (2-day forecast) was found. The forecast of the PM produces, for both the sites, a RMSE of 140–192 W/m^2 (1-day horizon) and 153–217 W/m^2 (2-day horizon). Similar behaviour was found in MAE.

Moreover, the use of NCEP-DOE II Reanalysis data for *WRF initialization* did not improve forecast accuracy.

On the contrary, a substantial reduction in error was achieved by using the MOS approaches. In Rome, the MOSRH seems to be enough to correct the main forecast bias errors of the WRF/GFS model, leading to a RMSE of 99 W/m^2 for the 1-day forecast and of 103 W/m^2 for the 2-day forecast. Thus, for this site the main NWP systematic error can be associated to an oversimplified treatment of the vertical distribution of humidity in the atmosphere by the radiation schemes. In Lugano, the outperforming forecast can be achieved by using the hybrid MOS, called MOS cascade, which means using the stochastic learning MOSNN along with the MOSRH output, namely MOSNN(MOSRH). In this case, the MOSNN(MOSRH) brings the RMSE to 137 W/m^2 for the 1-day forecast and of 149 W/m^2 for the 2-day forecast. Indeed, the complex orography of Lugano induces a high weather variability resulting in greater forecasting difficulties. Thus, the MOSNN corrects all the main learnable errors not directly related to the forecast of the water vapour concentration.

Finally, the results obtained were compared with the benchmark accuracy reached for the Southern Spain and the Swiss average climatic conditions (Lorenz et al., 2009b). Indeed, in Rome the performance reached by the persistence model was found to be between that of these two countries. The MOSRH provided a skill score (calculated with respect to the RMSE of PM) of 29–32% for the one and two day forecast which was perfectly inside the benchmark skill score ranges obtained for both Southern Spain (1-day forecast: between 36% and 1%; 2-day forecast: between 41% and 11%) and Switzerland (1-day forecast: between 49% and 22%; 2-day forecast: between 35% and 28%). Similar results were found for Lugano, even if the RMSE of the PM in Lugano is approximately between 20% and 25% higher than the PM RMSE averaged over the Swiss stations considered by Lorenz et al. for the benchmark calculation. In this case, the MOSNN(MOSRH) reached a skill score of 29% (one day ahead) and 31% (two days ahead) that is in the middle-upper part of the skill score ranges of the average Swiss climatic conditions reported above.

Acknowledgements

We gratefully acknowledge Raffaele Salerno of Epson Meteo Center (MOPI) for his support and encouragement in pursuing this research. Thank you to PV Performance Modeling Collaborative for providing PV_lib Matlab libraries.

The ESTER facility is part of the Centre for Hybrid and Organic Solar Energy (CHOSE), Italy, partially funded by the Region of Lazio.

References

Arakawa, A., Lamb, V.R., 1977. Computational design of the basic dynamical processes of the UCLA general circulation model. In: Meth. Comput. Phys., vol. 17. Academic Press, New York, pp. 173–265.

Beyer, H.G., Polo Martinez, J., Suri, M., Torres, J.L., Lorenz, E., Müller, S.C., Hoyer-Klick, C., Ineichen, P.D., 2009. Report on Benchmarking of Radiation Products. Report Under Contract No. 038665 of MESoR, 2009. <http://www.mesor.net/deliverables.html> (12.01.12).

Cai, T., Duan, S., Chen, C., 2010. Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement. In: Proceedings of the 2nd International Symposium on Power Electronics for Distributed Generation Systems, PEDG 2010, pp. 773–777.

Cao, J., Lin, X., 2008. Study of hourly and daily solar irradiation forecast using diagonal recurrent wavelet neural networks, 2008. Energy Convers. Manage. 49 (6), 1396–1406.

Chen, C., Duan, S., Cai, T., Liu, B., 2011. Online 24-h solar power forecasting based on weather type classification using artificial neural network. Sol. Energy 85, 2856–2870.

Cornaro, C., Pierro, M., Bucci, F., 2014. Master optimization process based on neural networks ensemble for 24-h solar irradiance forecast. Sol. Energy 111, 297–312.

Dudhia, Jimy, 1996. A multi-layer soil temperature model for MM5. The Sixth PSU/NCAR Mesoscale Model Users' Workshop.

Environmental Modeling Center, 2003. The GFS Atmospheric Model. NCEP Office Note 442, Global Climate and Weather Modeling Branch, EMC, Camp Springs, Maryland.

ERA-40, 2004. ECMWF 45-year reanalysis of the global atmosphere and surface conditions 1957–2002. ECMWF Newsletter 101, Summer/Autumn.

Fu, Q., Liou, K.N., 1992. On the correlated k-distribution method for radiative transfer in nonhomogeneous atmosphere. J. Atmos. Sci. 49, 2139–2156.

Galleano, R., Zaaiman, W., Virtuani, A., Pavanello, D., Morabito, P., Minuto, A., Spena, A., Bartocci, S., Fucci, R., Leanza, G., Fasanaro, D., Catena, M. Intercomparison campaign of spectroradiometers for a correct estimation of solar spectral irradiance: results and potential impact on photovoltaic devices calibration. Prog. Photovolt: Res. Appl. (published in early view at time of writing).

Grell, G.A., Peckham, S.E., Schmitz, R., McKeen, S.A., Frost, G., Skamarock, W.C., Eder, B., 2005. Fully coupled 'online' chemistry in the WRF model. Atmos. Environ. 39, 6957–6976.

Hong, Song-You, Noh, Yign, Dudhia, Jimy, 2006. A new vertical diffusion package with an explicit treatment of entrainment processes. Mon. Wea. Rev. 134, 2318–2341.

Houghton, J.T., 2002. The Physics of Atmospheres, third ed. Cambridge University Press, 340 pp.

Huang, Y., Lu, J., Liu, C., Xu, X., Wang, W., Zhou, X., 2010. Comparative study of power forecasting methods for PV stations. In: 2010 International Conference on Power System Technology: Technological Innovations Making Power Grid Smarter. Proceedings of POWERCON2010.

Ineichen, P., Perez, R., 2002. A New air mass independent formulation for the Linke turbidity coefficient. Sol. Energy 73, 151–157.

Kain, John S., 2004. The Kain-Fritsch convective parameterization: an update. J. Appl. Meteor. 43, 170–181.

Kleissl, J., 2013. Solar Energy Forecasting and Resource Assessment. ISBN 9780123971777, 1st Edition 2013.

Lara-Fanego, V. et al., 2011. Evaluation of the WRF model solar irradiance forecast in Andalusia (Southern Spain). Sol. Energy. http://dx.doi.org/10.1016/j.solener.2011.02.014.

Lorenz, E., Hurka, J., Heinemann, D., Beyer, H.G., 2009a. Irradiance forecasting for the power prediction of grid connected photovoltaic systems. IEEE J. Select. Topics Appl. Earth Observat. Remote Sens. 2 (1), 2–10.

Lorenz, E., Remund, J., Müller, S.C., Traunmüller, W., Steinmaurer, G., Pozo, D., Ruiz-Arias, J.A., Lara Fanego, V., Ramirez, L., Romeo, M.G., Kurz, C., Pomares, L.M., Guerrero C.G., 2009b. Benchmarking of Different Approaches to Forecast Solar Irradiance. In: Proceedings of the 24th European Photovoltaic Solar 35 Energy Conference 2009, Hamburg, Germany, pp. 4199–4208.

Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997. Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res. 102 (D14), 16663–16682.

Marquez, R., Coimbra, C.F.M., 2011. Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database. Sol. Energy 85, 746–756.

Muller, S.C., Remund, J., 2010. Advances in radiation forecast based on regional weather models MMF and WRF. In: Proceedings 25th EUPVSEC Conference 2010, pp. 4629–4632.

Mellit, A., 2008. Artificial intelligence technique for modelling and forecasting of solar radiation data: a review. Int. J. Artif. Intell. Soft Comput. 1, 52–76.

NCEP-DOE AMIP-II Reanalysis (R-2): M. Kanamitsu, W. Ebisuzaki, J. Woollen, S-K Yang, J.J. Hnilo, M. Fiorino, and G. L. Potter. 1631–1643, Nov 2002, Bulletin of the American Meteorological Society.

Oreopoulos, L., Barker, H.W., 1999. Accounting for subgrid-scale cloud variability in a multi-layer 1-D solar radiative transfer algorithm. Quart. J. Roy. Meteor. Soc. 125, 301–330.

Paulescu, M., Paulescu, E., Gravila, P., Badescu, V., 2013. Weather Modeling and Forecasting of PV Systems Operation, SBN 978-1-4471-4649-0, 1st Edition 2013.

Paulson, C.A., 1970. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. J. Appl. Meteor. 9, 857–861.

Pielke, R.A., 2002. Mesoscale Meteorological Modeling, second ed. Academic Press, 676pp.

Perez, R., Moore, K., Wilcox, S., Renné, D., Zelenka, A., 2007. Forecasting solar radiation-preliminary evaluation of an approach based upon the national forecast database. Sol. Energy 81 (6), 809–812.

Perez, R., Kivalov, S., Schlemmer, J., Hemker Jr., K, Rennè, D., Hoff, T.E., 2010. Validation of short and medium term operational solar radiation forecasts in the US. Sol. Energy 84, 2161–2172.

Perez, R., Lorenz, E., Pelland, S., Beauharnois, M., Van Knowe, G., Hemker Jr, K., Heinemann, D., Remund, J., Müller, S.C., Traunmüller, W., Steinmauerg, G., Pozo, D., Ruiz-Arias, J.A., Lara-Fanego, V., Ramirez-Santigosa, L., Gaston-Romero, M., Pomaresk, L.M., 2013. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. Sol. Energy 94, 305–326.

Photovoltaic and Solar Forecasting, State of the Art. IEA PVPS Task 14, Subtask 3.1. Report IEA- PVPS T14–01: 2013. October 2013. ISBN 978-3-906042-13-8.

Reno, M., Hansen C., Stein, J., 2012. Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis, Sandia National Laboratories, SAND2012-2389.

Rogers, E., Black, T., Ferrier, B., Lin, Y., Parrish, D., DiMego, G., 2001. Changes to the NCEP Meso Eta Analysis and Forecast System: Increase in Resolution, New Cloud Microphysics, Modified Precipitation Assimilation, Modified 3DVAR Analysis, online.

Spena, A., Cornaro, C., Serafini, S., 2008. Outdoor ESTER test facility for advanced technologies PV modules. In: Proceedings of the 33rd IEEE Photovoltaic Specialists Conference, San Diego, 2008.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W., Powers, J.G., 2008. A description of the advanced research WRF Version 3, NCAR Tech Note, NCAR/TN–475+STR (2008).

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M.A., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., Cuenca, R.H., 2004. Implementation and verification of the unified NOAH land surface model in the WRF model. In: 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction, pp. 11–15.

Traunmüller, W, Steinmaurer, G., 2010. Solar irradiance forecasting, benchmarking of different techniques and applications of energy meteorology. In: Proceedings of the EuroSun 2010 Conference. September 28–October 1, 2010, Graz, Austria.

Wang, S., Zhang, N., Zhao, Y., Zhan, J., 2011. Photovoltaic system power forecasting based on combined grey model and BP neural network. In: ICECE 2011 Proceedings, pp. 4623–4626.

Yona, A., Senjyu, T., Saber, A.Y., Funabashi, T., Sekine, H., Kim, C., 2008. Application of neural network to 24hours ahead generating power forecasting for PV system. In: Proceedings of IEEE Power and Energy Society 2008 General Meeting: Conversion and Delivery of Electrical Energy In the 21st Century, PES.

Zahumenský, I., 2004. Guidelines on Quality Control Procedures for Data from Automatic Weather Stations, World Meteorological Organization.

Zhang, G.Q., Patuwo, B.E., Hu, M.Y., 1998. Forecasting with artificial neural networks: the state of the art. Int. J. Forecast. 14, 35–62.