Solar Radiation Forecast Using Artificial Neural Networks

Fernando Ramos Martins, Enio Bueno Pereira, Ricardo André Guarnieri

Center for Earth System Science

Brazilian Institute for Space Research

São José dos Campos, Brazil. 12227-010

fernando.martins@inpe.br

Abstract

The fast increase in importance of the solar energy resource as viable and promising source of renewable energy has boosted research in methods to evaluate the short-term forecasts of the solar energy resource. There is an increase on demand from the energy sector for accurate short-term forecasts of solar energy resources in order to support the planning and management of the electricity generation and distribution systems. The Eta model is the mesoscale model running at CPTEC/INPE for weather forecasts and climate studies. It provides outputs for solar radiation flux at the surface, but these solar radiation forecasts are greatly overestimated. In order to achieve more reliable information, Artificial Neural Networks (ANN) were used to refine short-term forecast for the downward solar radiation flux at the surfaceprovided by Eta/CPTEC model. Ground measurements of downward solar radiation flux acquired in two SONDA sites located in Southern region of Brazil (Florianópolis and São Martinho da Serra) were used for ANN training and validation. The short-term forecasts produced by ANN have presented higher correlation coefficients and lower deviations. The ANN removed the bias observed in solar radiation forecasts provided by Eta/CPTEC model. The skill improvement in RMSE was higher than 30%when ANN was used to provide short-term forecasts of solar radiation at the surface in both measurement sites.

Keywords

Solar Energy Forecast; Short-Term Forecast; Artificial Neural Network;Energy Meteorology

Introduction

The scientific community points out that the fossil fuelexpenditure is the major reason of the observed growth of the greenhouse gases concentrationsin atmosphere along the last century [1]. Developed countries and advanced economies have been charged for the environmental damages due to consumptionof conventional energy sources to meet their energy demand. However, emerging economies such as Brazil, India, China, and Russia are increasingly sharing this responsibility as a result of their growing demand for energy to support their fast growing economic development

The commitment to reduce the emissions of carbon dioxide (and other greenhouse gases) established at the Kyoto Protocol and the perspectives of oil depletion in next decades are keyfactors to boostthe research and development onalternatives and renewable energy sources such as solar and wind [2, 3].

Furthermore, the search for improvement on energy security has been driving the government policies and incentive programs to stimulate the employment of alternative renewable energy sources even in countries with large share of clean energy in their electricity generation matrix. For example,in Brazil, where hydroelectric energy is responsible for more than 70% of the electricity matrix, an energy shortage happened in 2001 due to very low precipitation during the wet seasonof the previous year [4]. After this event, Brazilian government created incentive programs for renewable energy sources like wind energy.

The solar energy is one of the promising alternativesin Brazil since most of its territory is located in the inter-tropical region where solar energy resources are accessible all year round [5]. The main obstacles to the commercial exploitation of solar energy resources are the highest cost compared to the conventional electricity generation technologies, lack of information on resource assessment and variability, and thedeepdependencyon the weather and climate conditions [4]. The investment costs are expected to fall during the next decades due to technological advances and market demands [6]. The growing market for solar energy leads to an increase on the demand for more reliable information concerning to solar resources, including its spatial and temporal variability in short and long terms.In addition, the management of electricity generation and distribution systems is also asking for more accurate short-term solar energy forecasts.

Several methodologies were developed in order to provide solar radiation forecast in high temporal resolutions and short-term horizons [7, 8]. Some of them use numerical weather models (NWP). Such models have radiation parameterization codes to simulate the radiative atmospheric processes. Nevertheless, solar irradiation forecasts provided by NWP models for one or two days in advance have shown large deviations from solar irradiationdata acquired at surface [9]. The major factors responsible for such deviations are related to the solar irradiation dependence on clouds and weather conditions which intrinsically involve non-linear physical processes [10].

Absorption and scattering interactions are the atmospheric radiative processes that attenuate the solar radiation flux. Therefore, the atmospheric optical properties should be known in order to correctly evaluate the solar irradiation at any specific site and time. Clouds are the main factor that modulates the solar radiation incidence at the surface [11, 12, 13, 14, 15]. Atmospheric aerosols also have an important role in atmospheric radiative processes, mainly in some regions where anthropogenic emissions from biomass or fossil fuel burning takes place.

The Eta/CPTEC mesoscale model runs operationally in the Center of Weather Forecast and Climate Studies at Brazilian Institute for Space Research(CPTEC/INPE)and provides short-term forecasts for many meteorological variables, including surface solar irradiation. However, the references [11] and [12] showed that Eta/CPTEC model systematically overestimates the surface solar irradiation, as well as the sensible and latent heat fluxes at surface. A common issue in numerical atmospheric radiation codes is the excess of the incoming shortwave radiation at the surface as a result of the deficient parameterization of extinction interactions with water vapor,atmospheric aerosols and clouds.Several methodologies were published in order to improve solar forecasts provided by numerical weather models [9, 16, 17, 18].

This work aimsto presentamethodology to reduce deviations of solar irradiation forecasts provided by Eta/CPTEC model by using a statistical post-processing applied to the model outputs. This paper presents the results obtained when Artificial Neural Networks (ANN's) were used as statistical tool to refine the solar radiation forecast provided by Eta/CPTEC model.

Artificial neural networks (ANN) are data-driven instead of model-driven techniques once the results provided by them depend on the available data used to feed the ANN. Relationships between predictors (input data) and predictions are developed after building a system which simulates the physical processes in atmosphere. Artificial neural networks have been applied in renewable energy research for modeling and design solar systems and to provide short-term forecasts for energy resources [19]. Reference [20] indicated that the ANN systems are able to predict the solar radiation time series more effectively than the conventional procedures based on the clearness index. The authors observed that the forecasting ability can be further enhanced with the use of additional meteorological parameters like temperature and wind direction. References [21] and [22]discussed different methodologies using ANN to provide short-term forecasts for solar radiationby extracting knowledge from a long ground data series. Reference [23] compared some statistical models and ANN systems using meteorological data as input data. The authors concluded that ANN systems were a promising alternative to the traditional approaches for estimating global solar radiation, especially in cases where solar radiation measurements are not readily available.

This paper presents an attempt to get better predictability for the solar energy resources using operational Eta/CPTEC model and it constitutes an important application of the meteorology science to the energy planning and decision-making processes in energy sector. The target is to provide more precise and reliable information on future availability of solar resources in order to optimize electricity generation and distribution systems.

Methodology

Forecasting solar irradiation depends on prospecting the future atmospheric conditions. Despite the intrinsic uncertainties, NWP models provide information about many meteorological variables, including solar radiation data and atmospheric optical properties for several future timeframes. However, earlier studies demonstrated that solar radiation data provided by such models presents a large bias making its use inappropriate to electricity system management where several solar power plants are connected [10, 16, 17, 18].

This work employed the weather forecast outputs provided by the Eta/CPTEC model together withenvironmental data to feed Artificial Neural Network (ANN). The main goal was to achievea short-term forecast for solar irradiation with lower deviations than the ones provided by the Eta/CPTEC model. The solar radiation data acquired in two SONDA ground sites located in the Southern region of Brazil was used as reference for training and performance evaluation of the ANN.

Model Eta/CPTEC

The Eta/CPTEC model is used for operational weather forecasting, climate investigation, regional climate change studies and research onseveral issueslikepollutanttransport [24]. The Eta model, whichhas been running at CPTEC since 1996, was set up and optimized to the South America atmospheric conditions. The Eta/CPTEC model runs routinely for South America continent and neighboring oceans: latitudes from 50.2ºS to 12.2ºN, and longitudes from 83ºW to 25.8ºW. The horizontal resolutionequals to 40km and 38 vertical layerswere used for this study.

The Eta/CPTECmodel employsthe "finite difference" scheme to solve the equations system that describes the physical processes inatmosphere. The model uses the vertical coordinate "Eta", η, defined as:

$$\eta = \frac{p - p_t}{p_{sfc} - p_t} \frac{p_{ref}(z_{sfc}) - p_t}{p_{ref}(0) - p_t} \qquad (1)$$

where p_t is the pressure at the top of the model atmosphere, p_{ref} is the reference pressure to the vertical profile, and p_{sfc} and z_{sfc} are the pressure and height of the lower boundary surface, respectively.The Eta coordinate was adopted to reduce the large errors observed in several numeric weather forecast models that use the sigma surfaces [12]. These deviations arerelated to the determination of the horizontal pressure gradient force, as well as the advection and the horizontal diffusionon a steeply sloped coordinate surface [25, 26].

The discretization of the space domain uses the Semi-Staggered Arakawa E-grid on the horizontal and the Lorenz grid on the vertical. The radiation modeling uses the schemes described in [27] for shortwave radiation, and in [28] for long wave radiation. More detailed descriptions about the physical parameterizations adopted in Eta/CPTEC model can be found in [26, 29,30, 31].

The Eta/CPTEC model was executed using initial conditions at 00UT provided by NCEP analyses. The CPTEC Atmospheric Global Circulation Model (AGCM) provided the lateral boundary conditions.

The outputs provided by Eta/CPTEC model for 2001 till 2005 were used.The output file contains forecasts for 58 atmospheric variables at the synoptic timeframes (6, 12, 18 and 24UT) for 7 days in advance. The model provided the total amount in atmospheric column for forty-nine variables, and vertical profile values at 19 atmospheric pressure levels for the remaining nine variables.Only 33 out of the 58 atmospheric variables were used in this study. All vertical profile data were discarded together with 16 variables not representative of the atmospheric condition like topography, soil temperature and humidity for levels under surface.

Table I presents a complete list of model output data used for this work with a short description ofthem. Instantaneous values at each synoptic time were recorded for most of the data. However, average values regarding to the 6-hour period before each synoptic time were stored for some of the meteorological output variables, such as "*ocis*".

SONDA network

SONDA (Brazilian System for Environmental Data applied to the Energy Sector) is a network of ground measurement sites, operated and managed by INPE. The goal is to acquire reliable surface solar irradiation and wind data at different climate areas in Brazil in order to develop, improve and validate numerical models used for renewable energy resources assessment and environmental research. The SONDA database will provide valuable information applied to the research on the energy meteorology in Brazil.

In this work, the SONDA ground data acquired at two SONDA sites was used for the ANN training and configuration as described later in this paper. Besides that, ground data were used to evaluate the deviations presented by short-term forecast provided by both methodologies: Eta/CPTEC model and ANN. Both measurement siteswerelocated in the Brazilian Southern region:

- São Martinho da Serra (SMS) – 29.44ºS/53.82ºW.
- Florianópolis (FLN) – 27.60ºS/48.52ºW;

Fig. 1 shows the location of measurement sites of SONDA network featuring SMS and FLN sites. These both sites were chosen in order to evaluate the performance of ANN and Eta/CPTEC model in two different climate conditions. SMS is located in the continental area at 500m above the sea level. FLN is located at the coastal area of Brazilian Southern regionpresenting the largest total precipitation along the year in Brazilian territory. The SMS has been collecting data since June 2004 and FLN has been acquiring data since 1995.The other SONDA sites are more recent and have smaller databases.The SONDA website (http://sonda.ccst.inpe.br) presents all information aboutmeasurement sites and describes the data quality assurance program.

For this work, data acquired from January/2001 to October/2005 in FLN and from July/2004 to October/2005 in SMS were used. The Kipp&Zonen CM-21 pyranometers [32] were used to acquire global solar irradiation data. One-minute average solar irradiation data wasstored and its quality was checked. Both sites take part in Baseline Solar Radiation Network (BSRN) and meet all the quality criteria established by World Meteorological Organization (WMO).

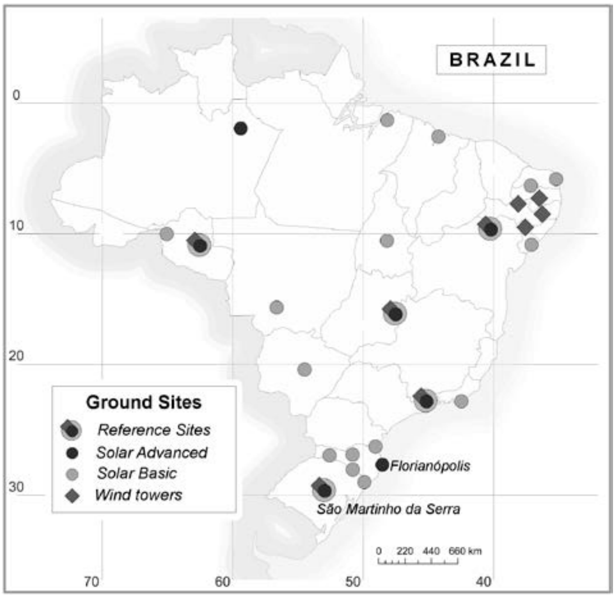

FIGURE 1 LOCATION OF GROUND SITES OF SONDA NETWORK. FLORIANÓPOLIS AND SÃO MARTINHO DA SERRA WERE USED FOREVALUATION OFSHORT-TERM FORECASTS.

After data-quality verification, 1150 days for FLN and 472 days for SMS were available for this work. The ground database was divided into 3groups as follows:

- Training group: with 575 days for FLN and 236 days for SMS;
- Validation group: with 288 days for FLN and 118 days for SMS;
- Investigation group: with 287 days for FLN and 118 days for SMS.

The training group was used for the ANN training. The validation group was employed to evaluate and establish the end of the training step. The investigation group was used to evaluate the reliability of ANN outputs. More details on each these three steps are described latter in this paper.

Data Management

As explainedearlier, the solar and meteorological database used to feed ANN comprises the output data provided by the model Eta/CPTEC (Table I). In addition, other three variables were calculated in order to supply ancillary information for the ANN: solar radiation flux at TOA (STOA), mean air mass (airm), and mean solar zenith angle (szam). Altogether, 36 variables were used as ANN predictors.

As described on Table I,the solar irradiation dataprovided by the Eta/CPTEC model, *"ocis"*, represents the 6-hour average solar irradiation. In order to achieve the same time-scale, the solar irradiation data acquired in FLN and SMS sites were averaged over the same 6-hour intervals. In summary, ground and model data of solar irradiationrepresents the total energy in the 6-hour period and they are expressed in $MJ.m^{-2}$ (mega joules per squared meter).

The 6-hour average solar radiation flux at the top of the Earth's atmosphere (STOA) was calculated taking into consideration local latitude, solar zenith angle, eccentricity and solar declination [13, 14]. As the ground solar irradiation data and *"ocis"*, the STOA solar radiation flux was also expressed in $MJ.m^{-2}$.

Relative humidity, atmospheric pressure, air temperature, wind velocities and all other instantaneous data, provided by Eta/CPTEC model for synoptic time (Table I),were averaged by taking the two consecutive values. The averages were assigned to the second synoptic timein order to set up the databasein a similar way used for ground data. This procedure aims to better represent theatmospheric and meteorological variability in the 6-hour interval.

In addition, the solar zenith angle (szam) and the air mass (airm) were obtained and stored for the same 6-hour intervals. Thus, the *"ocis"* data and all 36 variables used to feed ANN havethe same temporal resolution and represent the equivalent timeframes.

The 36 predictors and the ground data are disposed into four timeframes each day: 6:00, 12:00, 18:00 and 24:00UT. Each timeframe represents the corresponding time interval: 0-6UT (Rad06UT), 6-12UT (Rad12UT), 12-18UT (Rad18UT), and 18-24UT (Rad24UT). This paper only presents results for the Rad18UT timeframe. The Rad18UT was chosen because the highest fraction (63% – 80%) of solar radiation flux occurs during the 12-18UT intervals throughout the year at both ground sites [35].

TABLE 1 THE METEOROLOGIC DATA USED AS PREDICTORS IN ANN. ALL DATA WAS PROVIDED BY MODEL ETA/CPTEC

VARIABLE	DESCRIPTION (UNITS)	KEY FEATURES
rh2m	Relative humidity at 2m-height (0 to 1 – adimensional)	Instantaneous values
pslc	Pressure at surface (hPa)	Instantaneous values
tp2m	Temperature at 2m-height above the surface (K)	Instantaneous values
dp2m	Dew Point Temperature at 2m above the surface (K)	Instantaneous values
u10m	Zonal wind at 10m-height above the surface ($m\ s^{-1}$)	Instantaneous values
v10m	Meridional wind at 10m-height above the surface ($m\ s^{-1}$)	Instantaneous values
wnds	Wind velocity at 10m-height above the surface ($m\ s^{-1}$)	Instantaneous values
prec	Total rainfall ($kg\ m^{-2}\ dia^{-1}$)	Total in the 6h period
prcv	Convective rainfall ($kg\ m^{-2}\ dia^{-1}$)	Total in the 6h period
prge	Large scale rainfall ($kg\ m^{-2}\ dia^{-1}$)	Total in the 6h period
clsf	Latent Heat Flux at the surface ($MJ\ m^{-2}$)	Average value in the 6h period
cssf	Sensible Heat Flux at the surface ($MJ\ m^{-2}$)	Average value in the 6h period
ghfl	Heat Flux in the soil ($W\ m^{-2}$)	Average value in the 6h period
tsfc	Surface Temperature (K)	Instantaneous values
qsfc	Specific humidity at the surface ($kg(H_2O)\ kg(air)^{-1}$)	Instantaneous values
lwnv	Cloud cover Index for low clouds (0 a 1 - adimensional)	Instantaneous values
mdnv	Cloud cover Index for average clouds (0 a 1 - adimensional)	Instantaneous values
hinv	Cloud cover Index for high clouds (0 a 1 - adimensional)	Instantaneous values
cbnt	Mean Cloud cover Index (0 a 1 - adimensional)	Instantaneous values
ocis	Downward shortwave radiation flux at the surface ($MJ\ m^{-2}$)	Average value in the 6h period
olis	Downward longwave radiation flux at the surface ($MJ\ m^{-2}$)	Average value in the 6h period
oces	Upward shortwave radiation flux at the surface ($MJ\ m^{-2}$)	Average value in the 6h period
oles	Upward longwave radiation flux at the surface ($MJ\ m^{2}$)	Average value in the 6h period
roce	Upward shortwave radiation flux at the TOA ($MJ\ m^{-2}$)	Average value in the 6h period
role	Upward longwave radiation flux at the TOA ($MJ\ m^{-2}$)	Average value in the 6h period
albe	Albedo (0 a 1 - adimensional)	Instantaneous values
cape	Available potential convective energy ($m^2\ s^{-2}$)	Instantaneous values
cine	Energy to avoid convection ($m^2\ s^{-2}$)	Instantaneous values
agpl	Instantaneous precipitable water amount ($kg\ m^{-2}$)	Instantaneous values
pcbs	Pressure at the bottom of the clouds (hPa)	Instantaneous values
pctp	Pressure at the top of the clouds (hPa)	Instantaneous values
tgsc	Soil temperature at the surface layer (K)	Instantaneous values
ussl	Soil humidity at the surface (0 a 1 - adimensional)	Instantaneous values

Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANN) is computing systems, which attempt to simulate the structure and function of biological neurons. Generally,the ANNconsists of a number of interconnected processing elements, called neurons. Fig. 2 presents an artificial neuron. The ANN usually consists of an input layer, some hidden layers and an output layer. Signals flow from the input layer through to the output layer via unidirectional connections (synapses). Synapses connect neurons of neighboring layers. The input data (x_i) is weighted by values associated with each synapse (w_{ij}), called synaptic weights. Knowledge is usually stored as a set of connection weights (presumably corresponding to synapse efficacy in biological neural systems). The activity level of a neuron (v_j) is determined by summing up all its weighted values together with its bias (b_j). The neuron output is a result from an activation function ($\phi(v_j)$). Generally, the activation function is a linear or hyperbolic-tangent function. The non-linear activation functions allow ANNs to simulate non-linearity behaviors and complex patterns [19].

The ANN architecture depends on the physical process, the training method and the kind of data that the neural network will simulate. The multi-layer perceptron (or feedforward ANN) is the most widely ANN architecture used in meteorological topics [23]. A schematic diagram of typical multilayer neural network architecture is shown in Fig. 3. The input layer consists on one neuron for each input data (called predictor),and the output layer consists of one neuron for eachsimulated data (called predictant). The number of hidden layers and their total amount of neurons are not a priori established. There is no standard procedure to identify the best combination of neurons and layers.

The most widespread training algorithm used for multilayer perceptrons is the back propagation algorithm [33]. In this work, we use a modified version of back propagation, called Resilient Back propagation or Rprop [34]. The validation dataset was employed to verify the performance of the ANN with an independent data sample – data not used in training process. This procedure allowed to check the generalization capacity achieved by the ANN along the training and to find out the appropriate moment to stop the trainingstep in order to avoid overlearning. After training, the weights and bias are fixed and the ANN isready to be used in simulations.

For this study, preliminary experiments revealed that better ANN performances were achieved using two hidden layers of neurons. These experiments were developed in two different situations. First, the 36 variables described earlier were used as input to the ANN; and, in the second situation, only a set of 8 out of the 36 input variables were used. Table II shows the best neurons distributions verified for each ANN-model. On both cases, only one neuron is the output layer to provide information on solar radiation flux at surface. The number of neurons in the input layer is equal to the number of predictors used to feed ANN.

The investigation dataset was used to evaluate the performance of ANN to provide reliable solar irradiation forecast. The next topic discusses the statistical parameters used to evaluate deviations of the ANN and Eta/CPTEC outputs and the skills of each model to provide reliable forecasts.

TABLE 2NUMBER OF ARTIFICIAL NEURONS IN EACH ANN LAYER

	ANN-36p	ANN-8p
Input layer	36	8
First hidden layer	36	16
Second hidden layer	18	8
Output layer	1	1

ANN-36p – ANN using 36 variables as predictors
ANN-8p – ANN using 8 variables as predictors

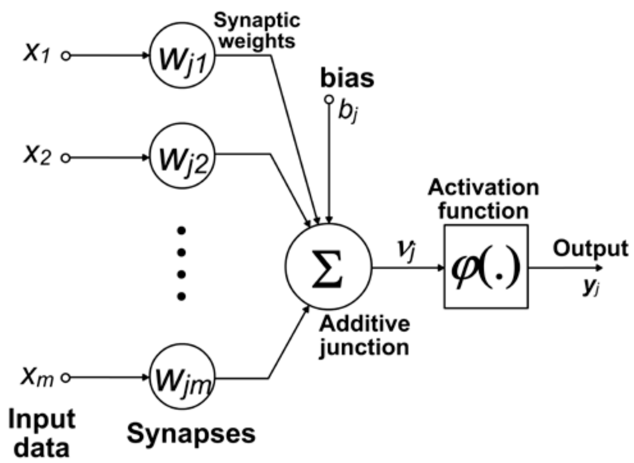

FIGURE 2 SYMBOLIC REPRESENTATION OF AN ARTIFICIAL NEURON AND ITS PARAMETERS.

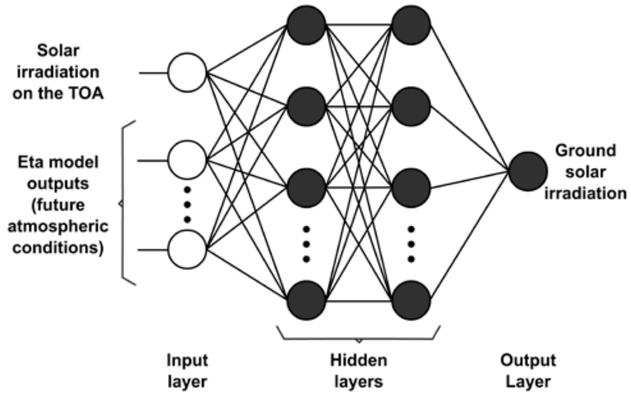

FIGURE 3SCHEMATIC DIAGRAM OF A FEEDFORWARD ANN USED IN THIS STUDY.

Statistical analysis of ANN and Eta/CPTEC outputs

The outputs (forecasts – *F*) were compared with measured values (observations – *O*), and deviations between them (*F* - *O*) were calculated. The performance of the Eta/CPTEC and ANN models was checked with two statistical indices: mean error (*ME*) or bias, and root mean squared error (*RMSE*). *ME* values provide information about the systematic deviations of the forecasts indicating if the models overestimateor underestimate the actual solar irradiation at the two measurement sites. *RMSE* is a measure of how effectively the models predict ground observations. Since the deviations are squared, large deviations have greatercontribution to*RMSE*. For this study, both *ME* and *RMSE*indices were normalized and expressed as percentage of the average solar irradiation in the two measurement sites, as shown in eq. (2) and (3).

$$ME\% = 100 \cdot \frac{\sum_{i=1}^{N}(F_i - O_i)}{\sum_{i=1}^{N}(O_i)}\% \tag{2}$$

$$RMSE\% = 100 \cdot \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2}}{\frac{1}{N}\sum_{i=1}^{N}(O_i)}\% \tag{3}$$

where *N* is the number of data pairs (forecast and observation) used in the evaluation – 287for FLN and 118 for SMS.

In addition, the Pearson's correlation coefficient (*R*) was computed as described in eq. (4):

$$R = \frac{\sum_{i=1}^{N}(F_i - \overline{F})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{N}(F_i - \overline{F})^2} \cdot \sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}} \tag{4}$$

In order to compare the performance of ANN and Eta/CPTECmodel, the skill-score index was used as defined in eq. (5):

$$Skill\ \ (Score, ref) = \frac{Score - Score_{ref}}{Score_{perf} - Score_{ref}} \tag{5}$$

where*Score* can be the *ME%* or the *RMSE%* values obtained for a particular model (Eta/CPTEC or ANN) in evaluation, $Score_{ref}$ is the score calculated for a reference method and $Score_{perf}$ is the score value expected for perfect-forecast.

Results and Discussion

Initially, the Eta/CPTEC forecast and ground data for solar radiation flux were compared. As demonstrated in previous studies [10, 11], a significant positive bias (overestimation) was observed in the solar radiation flux provided by Eta/CPTEC model. Table III shows the performance scores obtained for Eta/CPTEC estimatesusing only the investigation dataset (N = 287 for FLN; N = 118 for SMS). Similar scores were obtained when complete dataset was used for comparison between model estimates and ground data. Based on these results, it was assumed that the investigation dataset are representative of the complete dataset. Since ANN performance must be evaluated using the investigation dataset, only the Eta/CPTEC performance scoresusing this dataset were considered from this point on.

TABLE 3PERFORMANCE SCORESOBTAINED BY MODEL ETA/CPTEC

Scores	Florianópolis		São Martinho da Serra	
	N =1150	N =287*	N =472	N =118*
R	0.747	0.720	0.790	0.775
R^2	0.558	0.519	0.624	0.600
ME%	24.7%	24.6%	27.8%	28.0%
RMSE%	39.7%	40.0%	41.9%	43.2%

* - results obtained using only the investigation dataset.

As previously mentioned, various statistical analysis and simulations were performed using different subsets of the predictors listed in Table I in order to find a reduced dataset of predictors which produces a performance similar to that obtained when all 36 predictors are used. These analysis point out a set of 8 predictors: solar radiation flux at TOA (STOA), relative humidity (rh2m), surface temperature (tsfc), precipitable water amount (agpl), zonal wind speed at 10 m height (u10m), and predictors for cloud fractions (cbnt, hinv and mdnv). Hereafter, the ANNs using 36 and 8 predictors will be called ANN-36p and ANN-8p, respectively.

Table IV presents the performance scores obtained for ANN-36p and ANN-8p using the investigation dataset for both ground sites. As noticed, there is a very similar performance in terms of correlation (R) and RMSE deviations. However, the ANN-8p provided solar irradiationforecastsfor both sites with 50% less ME than the ANN-36p.

As noticed by comparing Tables III and IV, the ANN-36p and ANN-8p provided solar irradiation forecasts presenting larger correlation with ground observations in both sites. The ANN-8p outputs presented the lowest systematic deviation while Eta/CPTEC forecastsshowed the largest deviations (ME and RMSE) for both ground sites.

Fig. 4 and 5 present four scatter-plots comparing forecast values and observations. Besides the scatter-plots for Eta model, ANN-36p and ANN-8p, it is also showed a plot for a forecast method called persistence. The persistence forecast is the simplest method to predict meteorological data and it consists in taking the value observed in a previous day as the forecast for the current day. Any forecast method is useful if it can lead to better results than the persistence forecast.

According to Fig. 4 and 5, the solar radiation fluxoutputs provided by Eta/CPTEC model are better than persistence forecasts, in general. However, it can be observed the positive bias mentioned before. The Eta/CPTEC model overestimated the observations, especially for cloudy days when solar radiation flux at the surface is lower.

TABLE 4 PERFORMANCE SCORESOBTAINED BY ANN-36P AND ANN-8P

Scores	Florianópolis		São Martinho da Serra	
	ANN-36p	ANN-8p	ANN-36p	ANN-8p
R	0.804	0.790	0.839	0.848
R^2	0.646	0.625	0.704	0.720
ME%	-2.1%	-0.8%	-1.7%	-0.7%
RMSE%	26.2%	26.9%	28.8%	27.6%

All results obtained using investigation dataset.

Meanwhile, the scatter-plots for ANNs showed better agreement between forecasts and observations – most of the data points are located near the perfect-forecast line (diagonal line). Small difference was observed when ANN-8p is used instead ANN-36p, indicating that the 8 selected predictors was able to provide solar irradiation forecast as reliable as the forecast obtained by using the 36 predictors.

TABLE 5SKILL-SCORE CALCULATED WITH RMSE% VALUES FOR ANN TAKING MODEL ETA/CPTEC AND PERSISTENCE AS REFERENCE METHODS

Scores	Florianópolis		São Martinho da Serra	
	ANN-36p	ANN-8p	ANN-36p	ANN-8p
Skill(RMSE%, persistence)	0.429	0.414	0.464	0.487
Skill(RMSE%, Eta)	0.344	0.328	0.333	0.361

* - results obtained using investigation dataset.

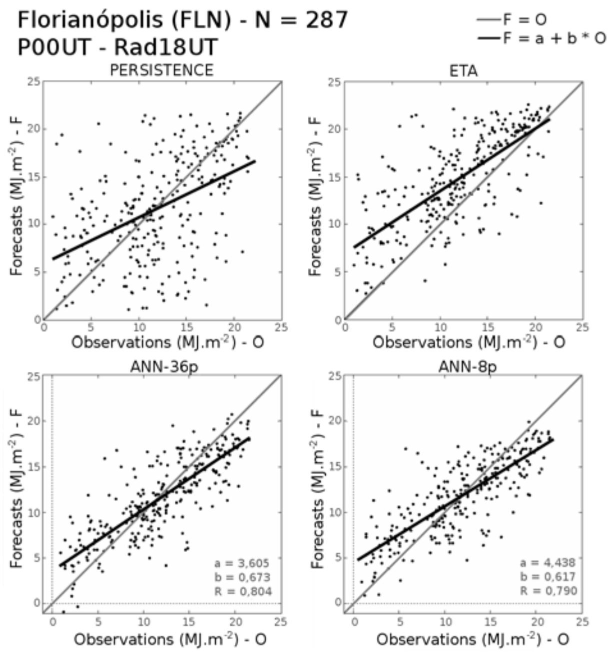

FIGURE 4 SCATTER-PLOTS OF FORECASTS VERSUS GROUND DATA FOR FLN: (A) PERSISTENCE METHOD, (B) MODEL ETA/CPTEC, (C) ANN-36P, AND (D) ANN-8P

Fig. 6 shows a short temporal series taken from the investigation dataset prepared for FLN and SMS sites. Outputs from Eta/CPTEC model and ANN were put together with observations acquired in Winter/2005 and Summer/2004-2005. Fig. 6 demonstrates the best agreement between the ANN forecasts and ground data. The deviations for each day are presented in Fig. 7. It is clear that an important improvement in short-term forecast for solar radiation fluxis achieved when ANN is used to refine solar irradiation outputs provided by model Eta/CPTEC. However, no significant differences were observed between ANN-36p and ANN-8p. Again, the analysis of Fig. 7 demonstrates that the eight selected predictors provide enough information to ANN simulate the atmospheric processes with good performance. To quantify the improvement acquired by the use of ANNs, the skill-score values were calculated using RMSE% score, and the results are presented in Table V. In general, the ANNs lead to skill-scores in RMSE% 30% higher if compared to model Eta/CPTEC.

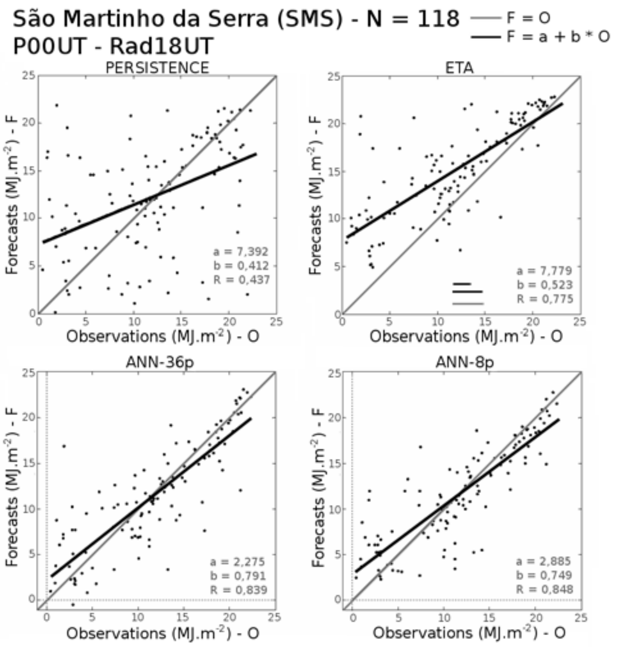

FIGURE 5SCATTER-PLOTS OF FORECASTS VERSUS GROUND DATAFORSMS: (A) PERSISTENCE METHOD, (B) MODEL ETA/CPTEC, (C) ANN-36P, AND (D) ANN-8P

Conclusions

Currently, the renewable sources of energy are getting more importance into electricity generation systems. Therefore, there is an increasing demand from the energy sector for accurate forecasts of solar energy resources in order to support and manage electricity generation and distribution systems. The forecasts provided by numerical weather models could supply this demand but, in general, these forecasts present large deviations reducing their confidence and reliability. In Brazil, the Eta/CPTEC model provided solar irradiation forecasts with bias around 25%. Lower deviations were observed when ANNwas usedto refine the forecastsprovided by the Eta/CPTEC model. The comparison between solar irradiation forecasts and ground data showed a bias reduction from 25%for Eta/CPTEC forecasts till -1% for the ANN outputs. Both ANNs, ANN-36predictors and ANN-8predictors, have presented very similar performances. The skill-score indices showed that both ANNs have improved the confidence and reliabilityonthe solar radiation forecasts in more than 30% for both sites: Florianópolis in coastal area and São Martinho da Serra in continental region. The improvements in predictability were also observed as indicated by the correlation coefficients: from 0.72 to 0.80 in FLN, and from 0.78 to 0.85 in SMS.

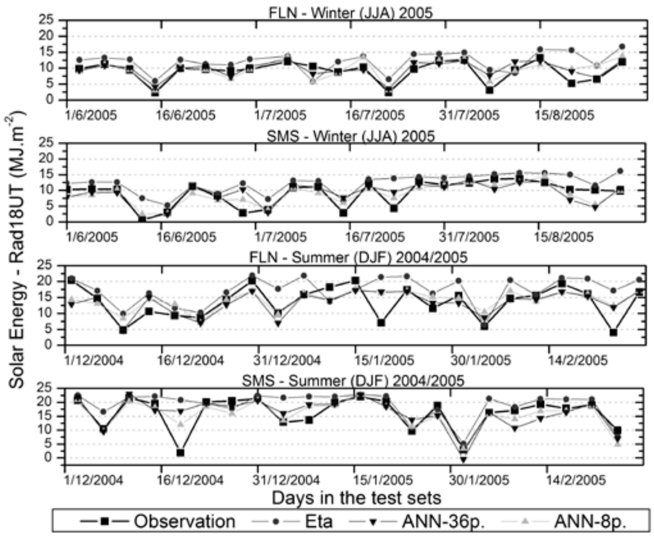

FIGURE 6 SHORT TIME SERIES COMPARING FORECASTS AND GROUND DATA FOR SOLAR RADIATION FLUX AT SURFACE IN FLN AND SMS

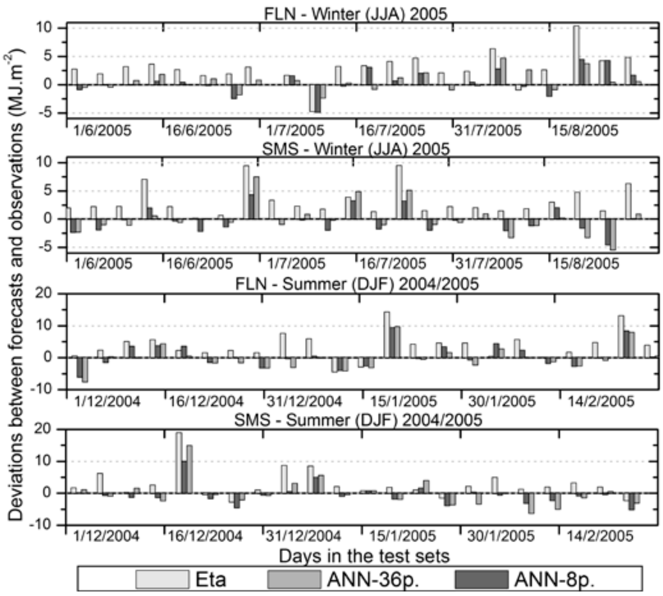

FIGURE 7 DEVIATIONS BETWEENFORECASTS AND GROUND DATA FOR SOLAR RADIATION FLUX AT SURFACE IN FLN AND SMS. THE MODEL ETA/CPTEC PROVIDED ESTIMATES WITH LARGER DEVIATIONS.

ACKNOWLEDGMENTS

The authors would like to thank FINEP (project22.01.0569.00) and PETROBRAS/CENPES (project 0050.0019.104.06.2) for their finance support to the SONDA project. Thanks are also due to the following colleagues: Silvia V. Pereira, Sheila A. B. Silva, Rafael Chagas and to the technologists of *Laboratório de Instrumentação Meteorológica* (LIM/CPTEC). In addition, acknowledgments are due to CPTEC/INPE and CNPq (grants 151700/2005-2, 141844/2006-0, 132148/2004-8, 555764/2010-9).

REFERENCES

[1] IPCC, 2007. In: Solomon, S., Quin, D., Manning, M., Chen, Z., Marquis, M., Averyt,K.B., Tignor, M., Miller, H.L. (Eds.), IPCC Climate Change 2007: The physicalscience basis. Contribution of Working Group I to the Fourth AssessmentReport of the Intergovernmental Panel on Climate Change. CambridgeUniversity Press, Cambridge, UK and New York, p. 996 Online: <htpp://www.ipcc.ch>.

[2] Bentley, R.W., 2002: Global Oil and Gas Depletion: An Overview. *Energy Policy*., 30, 189-205.

[3] Geller, H. S., 2003: Revolução Energética: políticas para um futuro sustentável. Relume Dumará, 299pp.

[4] Martins and Pereira, 2011.Enhancing information for solar and wind energy technology deploymentin Brazil. *Energy Policy*., 39, 4378-4390.

[5] Pereira, E. B., Martins, F. R., Abreu, S. L. and Ruther R., 2006. Brazilian *Atlas for Solar Energy*. Brazilian Institute for Space Research, 70pp.

[6] Viana et al, 2011. Assessing the potential of concentrating solar photovoltaic generationin Brazil with satellite-derived direct normal irradiation. *Solar Energy*, 85, 486-495.

[7] Reikard, G., 2008: Predicting solar radiationat high resolutions: a comparison of time series forecasts. *Solar Energy*, 83, 342-349.

[8] Hammer, A.; Heinemann, D.; Hoyer, C. and Lorenz, E., 2001: Satellite Based Short-Term Forecasting of Solar Irradiance - Comparison of Methods and Error Analysis, *Proceedings of the EUMETSAT Meteorological Satellite Data Users' Conference*, Antalya, Turkey, 01.-05.10.2001.

[9] Heinemann, D.;Lorenz, E.; Girodo, M., 2004: Forecasting of Solar Radiation. Proc. International Workshop on Solar Resource from the Local Level to Global Scale in Support to the Resource Management of Renewable Electricity Generation. Institute for Environment and Sustainability, Joint Research Center, Ispra, Italy.

[10] Hinkelman, L. M., Ackerman, T. P., and Marchand, R. T., 1999: An evaluation of NCEP Eta model predictions of surface energy budget and cloud properties by comparison with measured ARM data. *J. Geophys. Res.*, 104(D16), 19535-19549.

[11] Chou, S. C., Tanajura C. A. S., Xue, Y., and Nobre, C. A., 2002: Validation of the coupled Eta/SsiB model over South América. *J. Geophys. Res.*, 107(D20), LBA 56, 1-20.

[12] Mesinger, F., 2005: The Eta: brief history, summary, results, discussion. Proceedings of ICTP Regional Weather Predictabilityand Modeling Workshop. Centre for Weather Forecast and Climate Studies – Brazilian Institute for Space Research (CPTEC/INPE), São José dos Campos, Brazil. Available online at <http://etamodel.cptec.inpe.br/refer.shtml>.

[13] Kidder, S. Q., and Vonder Haar, T. H., 1995: *Satellite Meteorology: an introduction*. Academic Press, 466pp.

[14] Iqbal, M., 1983: *An Introduction to Solar Radiation*. Academic Press Canada, 390pp.

[15] Liou, K. N., 1980: *An Introduction to Atmospheric Radiation*. Academic Press Inc., 392pp.

[16] Armstrong, M. A., 2000: Comparison of MM5 Forecast Shortwave Radiation with Data Obtained from the Atmospheric Radiation Measurement Program. *Master of Science Scholarly Paper*, University of Maryland, USA.

[17] Zamora, R. J.; Solomon, S.; Dutton, E. G.; Bao, J. W.; Trainer, M.; Portmann, R. W.; White, A. B.; Nelson, D. W.; Mc Nider, R.T., 2003: Comparing MM5 Radiative Fluxes with Observations Gathered During the 1995 and 1999 Nashville Southern Oxidants Studies. *Journal of Geophysical Research*,108(D2), 4050-4060.

[18] Barbosa, H. M. J.; Tarasova, T. A.; Cavalcanti, I. F. A., 2008: Impacts of a New Solar Radiation Parameterization on the CPTEC AGCMClimatological Features. *Journal of Applied Meteorology and Climatology*, 47(5), 1377-1392.

[19] Kalogirou, S. A., 2001: Artificial neural networks in renewable energy systems applications: a review. *Renewable and Sustainable Energy Reviews*, 5, 373-401.

[20] Sfetsos, A. ; Coonick, A. H., 2000. Univariate and multivariate forecasting of hourly solarradiation with

artificial intelligence techniques. *Solar Energy*, 68(2), 169-178.

[21] Mihalakakou, G. ; Santamouris, M.; Asimakopoulos, D. N., 2000. The total solar radiation time series simulation in Athens,using neural networks. *Theoretical and Applied Climatology*, 66, 185-197.

[22] Hocaoglu, F. O.; Gerek, O. N.; Kurban, M., 2008: Hourly solar radiation forecasting using optimal coefficient 2-Dlinear filters and feed-forward neural networks. *Solar Energy*, 82, 714-726

[23] Tymvios, F.S.; Jacovides, C.P.; Michaelides,S.C.; Scouteli, C., 2005: Comparative study of Angström's and artificial neuralnetworks methodologies in estimating global solar radiation. *Solar Energy*, 78, 752-762.

[24] Chou, S. C.; Bustamante, J. F.; Gomes, J. L., 2005: Evaluation of seasonal precipitation forecasts over South America using Eta model. *Nonlin. Proc. Geophys*., 12, 537-555

[25] Staudenmaier Jr., M., 1996: A description of the Meso Eta model. *Western Region Technical Attachment* no. 96-06, 5pp. Available online at <http://www.wrh.noaa.gov/wrh/96TAs/TA9606/ta96-06.html>.

[26] Ničković, S., Mihailović, D., and Rajković, B., 1988: *Scientifc documentation of the Eta model*. Word Meteorological Organization.

[27] Lacis, A. A., and Hansen, J. E., 1974: A parameterization of the absorption of solar radiation in earth's atmosphere. *J. Atmosph. Sci*., 31, 118-133.

[28] Fels, S. B., and Schwarzkopf, M. D., 1975: The simplified exchange approximation: a new method for radiative transfer calculations. *J. Atmosph. Sci*., 32(7), 1475-1488.

[29] Black, T. L., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Weather and Forecasting*, 9(2), 265-278.

[30] Janjić, Z. L., 1994: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly Weather Review*, 122(5), 927-945.

[31] Mesinger, F.; Janjić, Z. L.; Ničković, S., and Deaven, D. G., 1988: The step-mountain coordinate: Model description and performance for cases of Alpine lee ciclogenesis and for a case of Appalachian redevelopment. *Monthly Weather Review*,116, 1493-1518.

[32] Kipp & Zonen, 2006: Kipp & Zonen Co. Webpage: manufacturer of meteorological and industrial sensors. Available online at <http://www.kippzonen.com/>.

[33] Rumelhart, D. E., Hinton G. E., and Williams, R. J., 1986: Learning representations by back-propagation errors. *Nature*, 323, 533-536.

[34] Riedmiller, M., and Braun H., 1993: A direct adaptive method for faster backpropagation learning: The RPROP algorithm. Proceedings, *IEEE International Conference on Neural Networks*.

[35] Guarnieri, R. A., 2006: Emprego de redes Neurais Artificiais e Regressão Linear Múltipla no refinamento das previsões de radiação solar do modelo Eta.Insituto Nacional de Pesquisas Espaciais, 171pp.