Conference Proceedings

EuroSun 2014
Aix-les-Bains (France), 16 – 19 September 2014

FORECAST OF SHORT-TERM SOLAR IRRADIATION IN BRASIL USING NUMERICAL MODELS AND STATISCAL POST-PROCESSING

Francisco J. L. Lima[1], Fernando R. Martins[2], Enio B. Pereira[1]

[1] INPE – Brazilian Institute for Space Research, São José dos Campos (Brazil).

[2] UNIFESP – Federal University of São Paulo, Campus Baixada Santista, Santos (Brazil)

Summary

This work aims at establishing a methodology to get reliable solar irradiation forecasts for the Brazilian Northeastern region by using WRF model together with statistical methods for post-processing data. The key issue is concerning how to deal with the diversity of typical climate features occurring in the region presenting the largest solar energy resource in Brazil. The solar irradiance forecasts for 24h in advance were obtained using the WRF model. In order to reduce uncertainties, cluster analysis technique was employed to find out areas presenting similar climate features. Comparison analysis between WRF model outputs and observational data were performed to evaluate the model skill in forecasting the surface solar irradiation. After all, post-processing of WRF outputs were performed using artificial neural networks and multiple regression methods in order to refine short-term solar irradiation forecasts.

1. Introduction

In contrast with other renewable energy sources (wind, biomass, and small hydroelectric plants) that have specific support by government incentives and polices, the penetration of the solar energy technology in Brazil is coming far behind. There were some important efforts to increase the available information on the Brazilian solar energy resources. Figure 1 presents a relative comparison between solar energy resources in Brazil and in countries where the solar energy market is far more advanced, such as Germany and countries in the Iberian Peninsula. It also shows the annual average of daily solar irradiation at the surface for each of the five Brazilian geopolitical regions. In addition to the large annual solar irradiation, its seasonal and inter-annual variability are low due to the fact that most of the Brazilian territory is located in a tropical region. Earlier studies have pointed out that the solar energy could be cost-effective all over Brazilian territory regarded particular conditions for each region. Applications like PV plants in the Amazon region to provide electricity for remote areas and small villages, concentrated solar power plants (CSP) in the arid area of Northeastern region, and water heating in South and Southeast of Brazil are examples for feasible exploitation of solar resources (Martins et al., 2008).

Despite the great solar resource and high value that can be attributed to grid-connected PV systems in commercial areas of urban centers in Brazil, the installed PV capacity is meager and restricted to universities and research institutes. The PV generation shares only 0,01% of the Brazilian electricity matrix, around 12000 kW (ANEEL, 2014).

PV applications have a promising future in commercial urban tropical and subtropical regions where high midday air-conditioning loads have typical a demand curve in a good match for the daily cicle of surface solar irradiation curve. Another important feature of the solar energy resource in Brazilian territory is regarding the peak load values in summer season. The solar resource achieves higher values along the the hotter climate conditions. This is the typical picture for most metropolitan areas and large cities in Brazil (Martins et al., 2008). Currently, the Brazilian government is discussing how to support and incentive the of solar energy market in Brazil and national policies are being implemented to regulate the smart gridding technology. Energy auctions for solar energy are being held in order to increase solar energy share and stimulate the energy sector to invest and develop solar energy applications.

Other important future prospect for PV generation is in remote areas not connected to the Brazilian Interconnected Distribution System such as Brazilian Amazon region, for example. There are currently only a few hundreds of mini-grids operated by independent power producers (IPPs) or local state utilities in the Brazilian Amazon, that cover the main share of local demand. All the electricity is generated in fossil fuel power plants but most of the sites are not easily accessible, increasing cost and decreasing reliability of supply. The potential for using PV systems, however, is huge, and can be estimated in tens to hundreds of

Selection and/or peer review under responsibility of EuroSun2014 Scientific Committee
doi:10.18086/eurosun.2014.09.01 Available at http://proceedings.ises.org

MWp, even if only a fraction of the existing Diesel oil plants would adopt some PV to an optimum Diesel / PV mix (Pereira et al., 2006).

Presently, several energy companies are evaluating the economic feasibility and planning to operate solar power plants (PV, PV concentrated and CSP) in the Brazilian Northeastern and Mid-West regions. In order to contribute to this effort, INPE and several universities are working to provide reliable scientific information on solar energy assessment and spatial and temporal variability. INPE is preparing the revised edition of Brazilian Atlas for Solar Energy in order to deliver data for solar energy resource taking into consideration typical atmospheric conditions observed in Brazil like high aerosol load during dry season due to biomass burning events. Furthermore, this is the appropriate moment to research and develop a suitable methodology to provide short-term forecast of solar irradiance taking into consideration the climate and environment characteristics of Brazilian territory. Since Brazilian energy scenarios are pointing out that the solar generation would grow in near future, the reliable short-term forecasts will be required for planning and managing the PV plants and the electricity system.

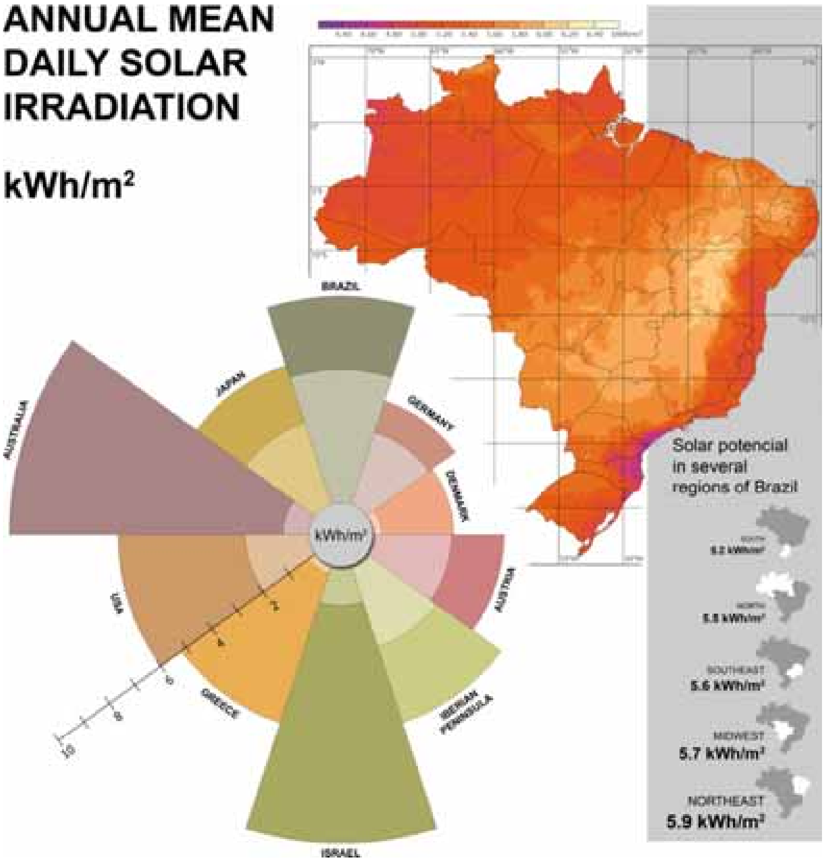

Figure 1. Mean annual range of the solar resources in Brazil compared to other countries (source, Pereira et al., 2006).

2. Methodology

The observational solar data for hourly solar irradiation used in this study were acquired in 110 Automatic Weather Stations (AWS) located in the Northeastern region of Brazil along the two typical climate seasons: rainy and dry seasons. All of ground measurement sites are managed by Brazilian Institute for Meteorology (Instituto Nacional de Meteorologia – INMET). Data for the rainy season were acquired in May/2009 Along the dry season, data was acquired in September/2009. The first step was to check the quality of observed data, this procedure was necessary to minimize the chances of a biased analysis due to the low reliability of observed data. The quality criteria used in BSRN (Baseline Surface Radiation Network) managed by Word Meteorological Organization) (Martins e Pereira, 2011) were also employed here. Figure 2 shows the location of ground sites presenting high quality data in Brazilian Northeastern region (NEB).

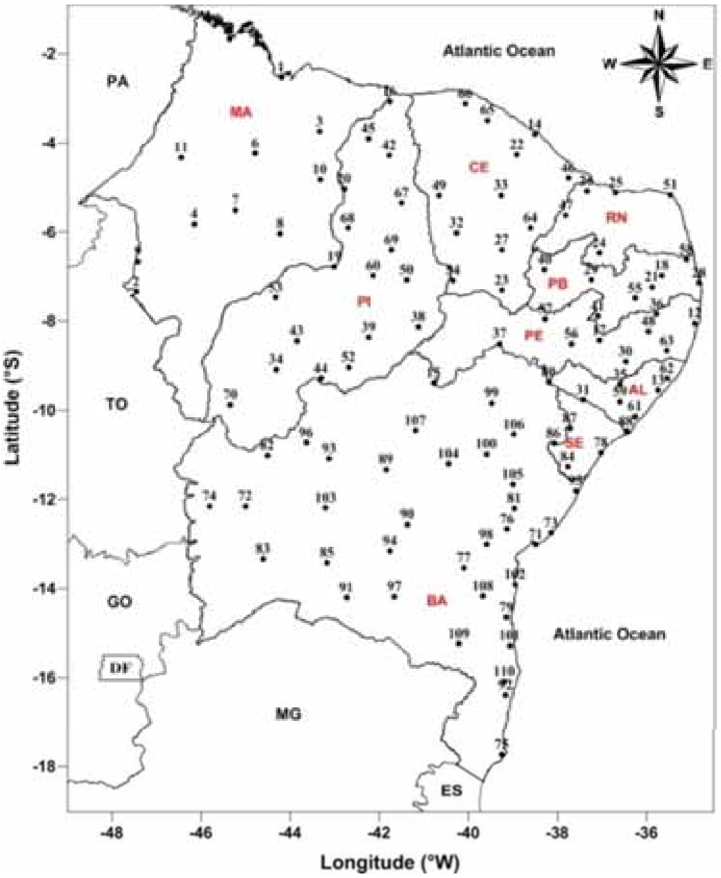

Figure 2. Location of stations across of Northeastern of Brazil.

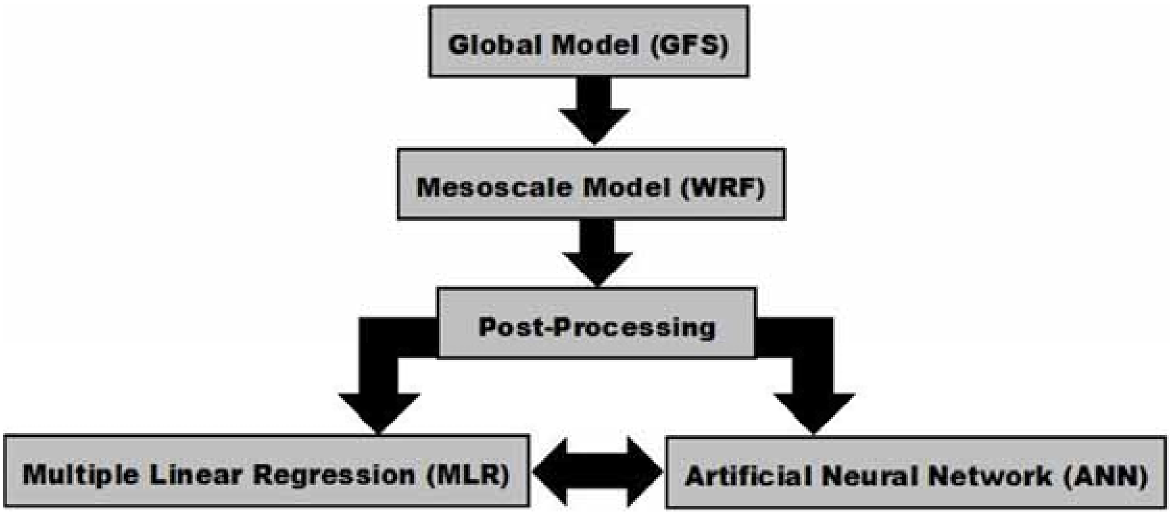

Figure 3. Illustrates basically the methodology used in this research.

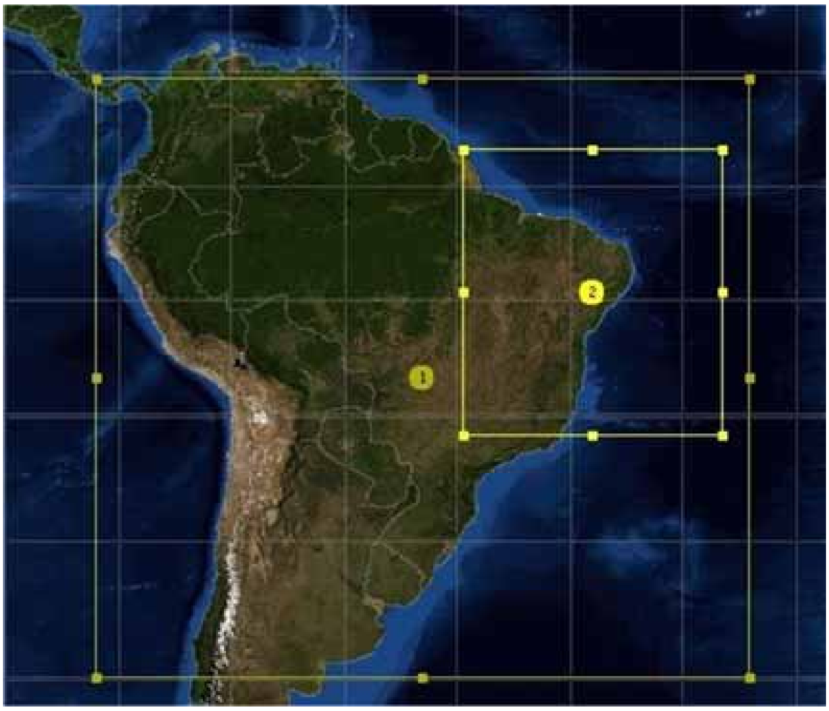

Figure 4. Resolution of WRF Model and their respective Grids.

The recognition of areas presenting similar behavior concerning surface solar irradiation in both typical climate seasons was the next step. The cluster analysis technique was applied in order to delineate areas according to similarity of meteorological and solarimetric ground data. The agglomerative hierarchical clustering method, the method was used (Ward, 1963) by using Euclidean distance as the grouping function, (Hair et al., 2005). The dendrogram analysis provided all information to define the number of groups and the ground measurement sites contained in each of them.

The solar irradiance forecast for 24h in advance were obtained following the procedure illustrated in Figure 3. In the first step, WRF model (Skamarock et al., 2005 and 2008) provides the first guess for surface solar irradiance. Model simulations extended over thirty-one days in May (rainy season) and the thirty days in September (dry season) of 2009, and forecast outputs were saved at once every hour. Global (FNL) datasets, gridded in 1° latitude by 1° longitude and originated from 6-hour archive data of NCEP provided the initial and boundary conditions for WRF runs.

The WRF model setup was prepared for two spatial domains by using two-way nesting strategy as shown in Figure 4. The large domain includes Brazil and it consists of 120 x 110 square grid points with 45 km resolution. The finer domain covers the Northeastern region of Brazil consisting of 142 x 157 square grid with 15 km horizontal resolution. Both domains have the same vertical structure comprising 50 levels. The Monin-Obukov (Eta) scheme (Janjic,2002), the BouLac scheme (Bougeault and Lacarrere, 1989) and the Purdue (Lin) scheme (Lin, Farley and Orville, 1983) are used by WRF in order to simulated and prediction surface layer physics, boundary layer processes and atmospheric microphysics' processes, respectively. The Grell-3 used to parameterize cumulus physics, and long and short wave radiation fluxes in the atmosphere were simulated by using the Rapid Radiative Transfer Model (RRTM) scheme (Mlawer et al., 1997) and the Dudhia scheme (Dudhia, 1989), respectively.

Two post-processing methods were applied to reduce uncertainties observed in WRF 24h-forecasts. The proposed MLR method is described in (eq.1). It was assumed that surface solar irradiance (I_{glob}) can be obtained from the product of the clear sky index (k*) by the surface solar irradiance at the cloudless condition (I_{clear}). The clear sky model proposed by Dumortier (1998) provided the surface solar irradiation in a cloudless condition. The MLR method provided k^* value by using the low, medium and high cloud cover index data provided by the WRF model in order to find an equation for prediction of solar irradiation, Equation 1 . Finally, a bias-correction in dependence on the predicted cloud situation is applied.

$$Irradiation = (a(clflo) + b(clfmi) + c(clfhi) + c)ClearSky \tag{eq.1}$$

The ANN is parallel and distributed systems consisting of simple units, nodes or neurons that compute (especially nonlinear) mathematical functions. The ANNs are presented as statistical tools capable of storing knowledge from examples, and has been applied to function fitting problems, pattern recognition, predictive modeling and other applications in various areas of human knowledge. Analogous to the human nervous system, the nodes are arranged in one or more layers interconnected by numerous links generally unidirectional, called synapses (Haykin, S., 2001). All these connections are associated with values, called synaptic weights, responsible for weighting the inputs of each node as a storage form of knowledge. Data for clear sky index, clflo, clfmi and clfhi were used as input data for the ANN training (60%) and validation (40%). The Levenberg-Marquardt backpropagation algorithm was choosen for ANN training.

The solar irradiation outputs provided by MLR and ANN were compared to observations collected by INMET that provides hourly averaged data. The quality and reliability of models result was evaluated by using the following statistical indexes: BIAS, RMSE and correlations.

3. Results

The cluster analysis using data acquired at 110 stations was conducted by using the Ward methodology employing the Euclidean distance as a function to evaluate the similarity or dissimilarity between ground measurement data. Other methods were tested but did not present results as good as Ward method.

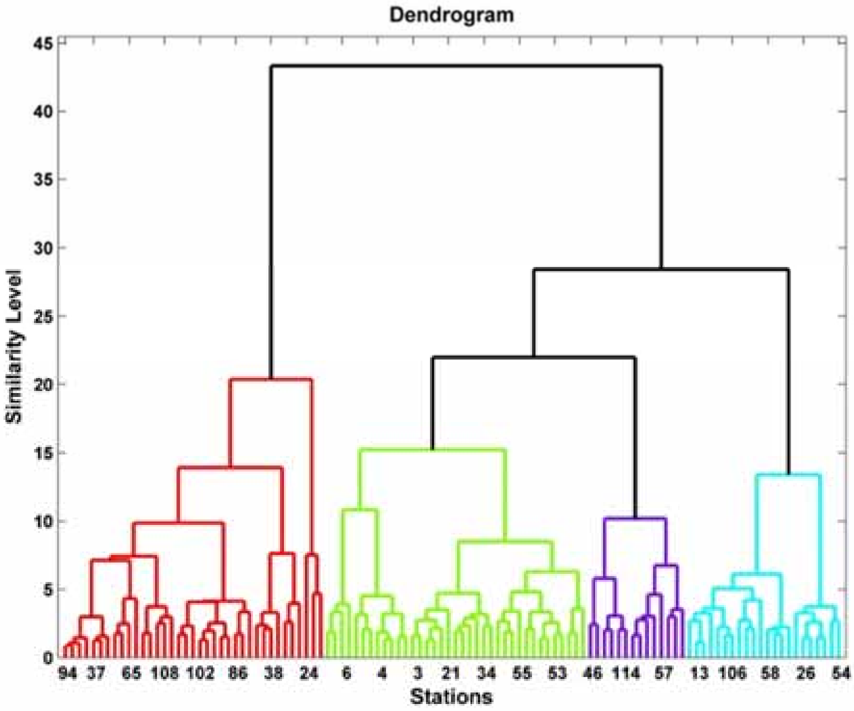

Figure 5. Dendrogram obtained from the cluster analysis of data acquired in 110 automatic weather stations using the agglomerative hierarchical Ward method.

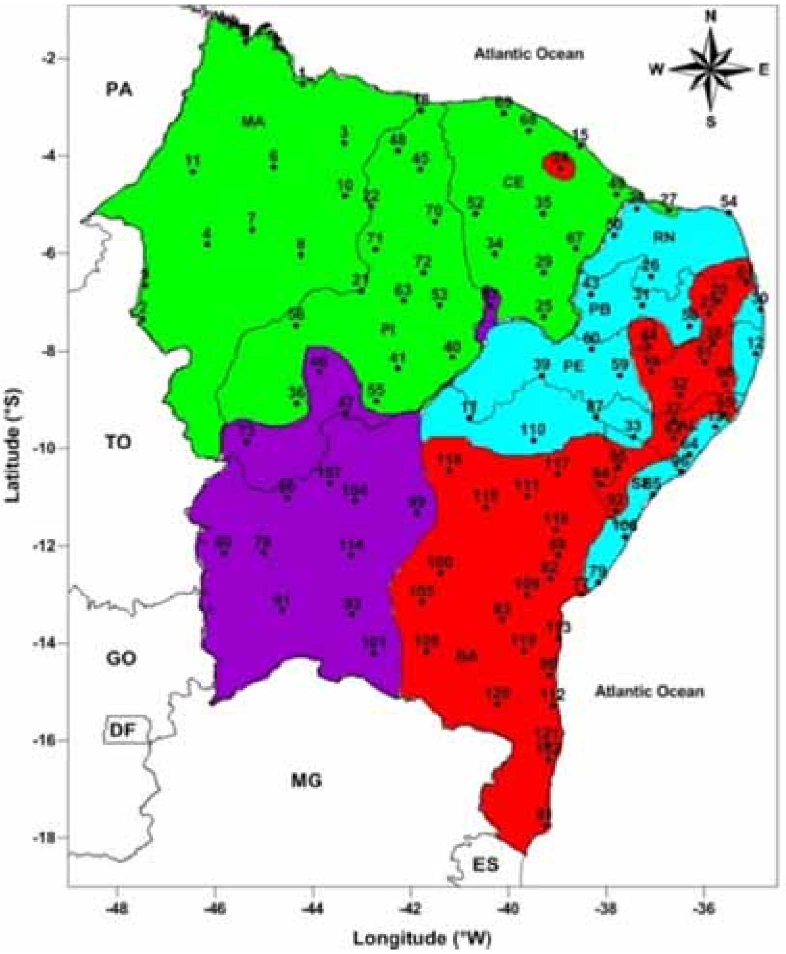

Figure 6. Distribution of the four similar areas in Brazilian Northeastern region concerning surface solar irradiation.

The choice of for the number of regions was made subjectively taking the dendrogram presented in Figure 5. Four homogeneous areas in NEB were easily identified. The number of regions was considered consistent with the climate features observed in NEB. Each of these regions showed a representative number of ground measurement site in operation. The vertical axis in Figure 5 shows the similarity level, corresponds a numeric value representing the level of grouping. The larger the index, more individuals are heterogeneous, or in this case, the individuals are meteorological stations in the horizontal axis.

Figure 6 shows the map of the grouping of the areas in the northeastern. The solar radiation variability in northeastern Brazil is significantly heterogeneous, whether by dynamic effects caused by its topography or by action of different atmospheric systems that reach this region, showing large spatial and temporal variability. In the Northeast, the high levels of solar radiation are explained by low cloud cover, especially in the backcountry. Maximum annual solar irradiation are situated in Sector North and Northeast, green and light blue areas, except in the coastal region and in the Chapada da Diamantina in Bahia, area of high altitude, Red region. The largest observed solar radiation indices in the backcountry, purple, light blue and green areas, where has the smallest occurrences of rain in the Northeast. This distribution is characterized by the performance of meteorological phenomena that operate in the Northeast throughout the year, which include: the Intertropical Convergence Zone (ITCZ), frontal systems, Upper Air Cyclonic Vortices and Eastern Disturbances (sea breeze and trade winds from the Southeast), as cited in Silva, (2003). The main meteorological phenomenon that act in the green, light blue and purple areas is the Intertropical Convergence zone (ITCZ) which is more intense in the months of March, April and May, featuring the months with lower incidence of solar irradiation and weaker in the months of August, September and October where has the highest rates of irradiation in the northeastern region. In the Red region in addition to the influences of the (ITCZ) has a strong influence of frontal systems of upper air cyclonic Vortices and Eastern disturbances which features a greater cloud cover in this region throughout the year, these phenomena also has influences in light blue and purple. Being that the Green Group is largely influenced by ITCZ and sea breeze.

Figures 7 and 8 presents the maps for daily total of solar irradiation in May and September 2009. Both Figures allow comparing the monthly average of observed and forecasted data for surface solar irradiation. It can be noted that ANN methods provide outputs, in both seasons, presenting better agreement to the observational data concerning the spatial distribution of solar energy resource.

In order to validate and compare both statistical post-processing WRF 24h-forecasts, BIAS and RMSE values of the outputs for solar irradiation provided by MLR and ANN were obtained by using data acquired in several ground sites in all four similar areas. Table 1 presents the BIAS, RMSE and correlation index obtained when forecasts were compared to ground data acquired in four sites in each similar area. The statistical indexes demonstrated that MLR and ANN were able to reduce uncertainties in forecasts provided by WRF model. However, ANN presented better performance and provided forecasts with lower BIAS and RMSE for the most of ground sites used in validation step.

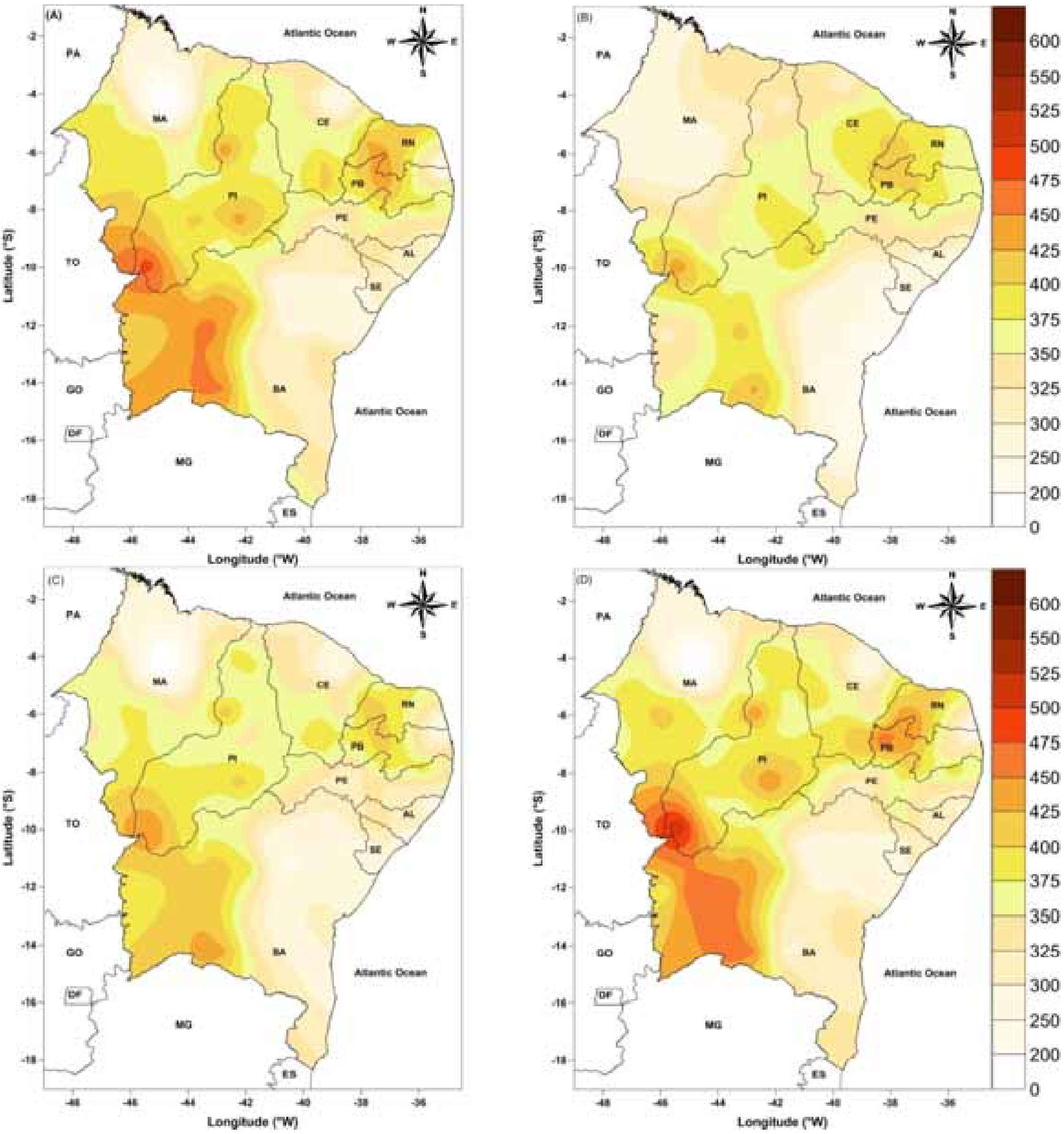

Figure 7. Map of solar irradiation (W/m²) for May/2009 (rainy season). (A) measured data, (B) Model WRF, (C) Model MLR e (D) Model ANN.

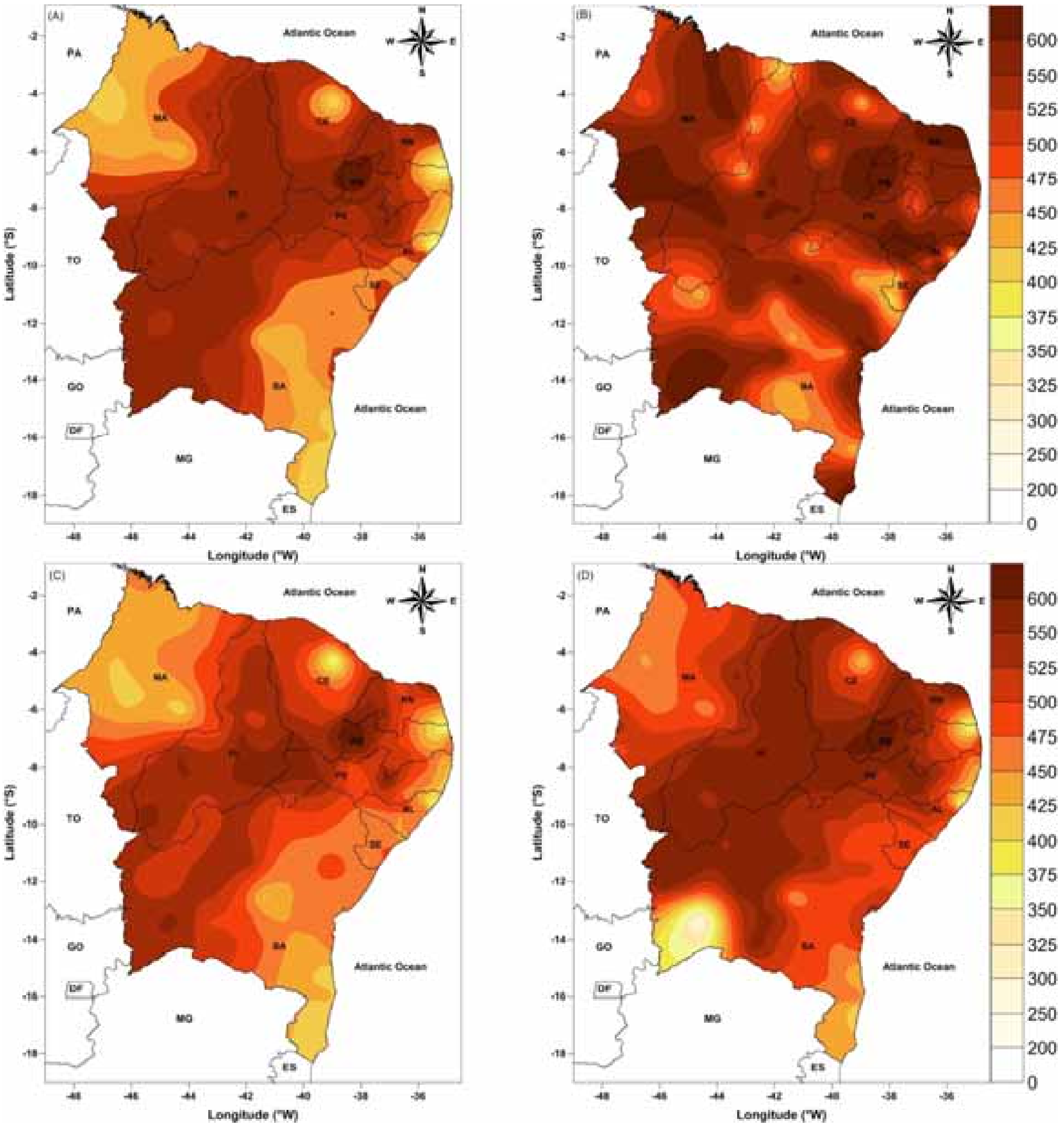

Figure 8. Map of solar irradiation (W/m²) for September/2009 (dry season). (A) Measured data, (B) Model WRF, (C) Model MLR e (D) Model ANN.

Table 01. Statistical indexes presented by 24h-forecasts provided by WRF model, MLR and ANN post processing methods for May/2009.

MODELS		WRF	MLR	ANN	WRF	MLR	ANN	WRF	MLR	ANN
ID	**INDEX**	**BIAS (W/m²)**			**RMSE (W/m²)**			**r**		
AREA 01										
46	**A DO GURGEIA**	-50.75	-14.88	-7.55	234.34	151.82	140.35	0.69	0.82	0.85
90	**S R DE CASSIA**	-56.47	-28.19	-6.22	211.83	144.10	131.11	0.74	0.84	0.86
93	**B J DA LAPA**	-67.15	-27.15	0.04	234.89	128.47	127.00	0.66	0.86	0.85
99	**IRECÊ**	-87.77	-30.22	-18.41	255.53	130.84	122.44	0.64	0.87	0.89
AREA 02										
13	**MACEIÓ**	-38.39	-9.77	-24.13	272.96	176.33	167.91	0.54	0.73	0.77
28	**MOSSORÓ**	-10.67	-24.38	-2.91	232.96	159.38	147.00	0.71	0.81	0.83
54	**CALCANHAR**	-35.58	-28.16	-15.47	237.89	183.67	164.62	0.73	0.81	0.85

85	**ARACAJÚ**	-76.88	-23.33	13.87	271.83	178.37	175.99	0.56	0.75	0.75
					AREA 03					
23	**CAMP GRANDE**	15.89	-15.30	-37.06	218.21	147.58	166.38	0.77	0.81	0.76
62	**ARAPIRACA**	-47.77	-2.41	-6.07	237.71	138.71	140.11	0.63	0.80	0.80
81	**CARAVELAS**	-29.64	-25.30	-9.22	227.56	138.07	135.26	0.63	0.79	0.79
112	**UNA**	-89.01	-24.69	-5.02	232.41	129.30	119.39	0.61	0.80	0.83
					AREA 04					
9	**ESTREITO**	-80.28	-33.98	-10.30	279.83	182.62	172.37	0.54	0.73	0.75
18	**PARNAÍBA**	-2.30	-28.61	-5.44	236.40	148.97	133.27	0.68	0.82	0.85
34	**TAUÁ**	5.82	-1.44	6.96	228.89	135.29	120.43	0.74	0.84	0.87
68	**ITAPIPOCA**	-51.45	-25.28	0.20	285.19	177.30	169.09	0.55	0.78	0.80

Table 02. Statistical indexes presented by 24h-forecasts provided by WRF model, MLR and ANN post processing methods for September/2009.

	MODELS	**WRF**	**MLR**	**ANN**	**WRF**	**MLR**	**ANN**	**WRF**	**MLR**	**ANN**
ID	**INDEX**	**BIAS (W/m²)**			**RMSE (W/m²)**			**r**		
					GROUP01					
46	**A DO GURGEIA**	47.67	-16.14	6.47	136.68	103.56	86.65	0.94	0.95	0.96
90	**S R DE CASSIA**	-117.39	-37.99	9.85	308.64	115.23	101.88	0.66	0.93	0.94
93	**B J DA LAPA**	71.41	-33.90	-9.76	193.87	127.40	111.55	0.86	0.90	0.92
99	**IRECÊ**	-125.38	-43.63	14.17	273.93	127.86	105.34	0.78	0.93	0.95
					GROUP02					
13	**MACEIÓ**	-85.72	-36.54	-6.74	263.26	162.00	148.46	0.72	0.85	0.87
28	**MOSSORÓ**	0.48	-28.00	-23.19	167.80	115.92	91.66	0.90	0.93	0.96
54	**CALCANHAR**	18.33	-33.77	15.60	133.37	119.08	84.78	0.93	0.94	0.97
85	**ARACAJÚ**	121.52	-30.12	-12.75	201.76	164.81	143.36	0.89	0.86	0.89
					GROUP03					
23	**CAMP GRANDE**	79.80	-47.81	-3.82	205.54	174.86	168.23	0.86	0.83	0.83
62	**ARAPIRACA**	6.90	-39.67	36.67	220.60	147.48	138.72	0.81	0.87	0.89
81	**CARAVELAS**	195.62	-28.61	-0.58	299.32	177.78	158.85	0.77	0.79	0.83
112	**UNA**	103.75	-22.80	-12.69	244.65	151.68	132.68	0.83	0.84	0.87
					GROUP04					
9	**ESTREITO**	124.20	-29.71	39.94	243.12	152.70	159.87	0.77	0.83	0.82
18	**PARNAÍBA**	-138.75	-36.63	0.46	251.79	110.49	90.09	0.80	0.94	0.95
34	**TAUÁ**	-54.22	-21.29	-9.25	255.59	146.67	139.04	0.76	0.90	0.91
68	**ITAPIPOCA**	27.78	-14.34	4.66	204.76	128.44	90.40	0.85	0.92	0.95

4. Conclusions

The aim of this study was to investigate the main methodology for establishing a system of simulation and forecasting using the WRF and the post-processing models ANN and MLR. It was concluded that the numerical prediction of solar irradiation with the use of WRF, in general, has a satisfactory performance with good relationship between the forecast data, compared with the observed data and has average correlations for the month of May, rainy season and significant correlations in September, dry season, the period of greatest intensity irradiation. However the model has a high systematic error with high values of BIAS and RMSE for the four areas and two months under study. You can establish an efficient methodology for the prediction and simulation of solar irradiation in the Northeast using the WRF coupled with a statistical refinement with the use of artificial neural networks (ANN) or multiple linear regressions (MLR). Was observed with the use of ANN and MLR a significant decrease in systematic error of solar irradiation and reduction of BIAS and RMSE in almost all the stations that form the areas at northeastern Brazil. There is an improvement of RNA compared the RLM, but both methods were efficient in improvements of the forecasts of the WRF. In short it is possible to use WRF model for predicting solar irradiation in Brazil to the use of a due statistical refinement to reduce the systematic error of the model.

5. Acknowledgments

Authors thank the National Institute of Meteorology (INMET), for the date used in this research. Also, thanks the Conselho Nacional de Desenvolvimento Cientifico e Technologic (CNPq), for the financial support.

6. References

ANEEL, 2014. Banco de informações de geração. Agência Nacional de Energia Elétrica. Available for public access at http://www.aneel.gov.br/aplicacoes/capacidadebrasil/capacidadebrasil.cfm

Bougeault, P., and P. Lacarrere., 1989. Parametrization of orography-induced turbulence in a mesoscale model. Mon. Weather Rev., 117, 1872-1890.

Dudhia, J., 1989. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. Journal of Atmospheric Science, v. 46, p. 3077-3107.

Dumortier D., 1998. The Satellight model of turbidity variations in Europe. Technical Report.

Hair, Jr. J. F., et al., 2005. Multivariate Data Analysis. Bookman. Inc., Porto Alegre, ISBN 85-363-0482-0.

Haykin, S., 2001.Neural Networks: Principles and Practice, Bookman. Inc., Porto Alegre. 900p.

Janjic, Z. I., 2002. Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso model, NCEP Office Note, No. 437, 61 p.

Lin, Y uh-Lang, Richard D. Farley , Harold D. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model. J. Appl. Meteor., 22, 1065-1092.

Martins, F. R., Pereira, E. B., 2008. Estudo comparativo da confiabilidade de estimativas de irradiação solar para o sudeste brasileiras obtidas a partir de dados de satélite e por interpolação/extrapolação de dados de superfície. Revista Brasileira de Geofísica. Vol. 29(2), p. 265-276, 2011.

Martins, F. R., et al., 2008. Solar energy scenarios in Brasil, Parto one: Resource assessment. Energy Policy. Vol. 36, p. 2843-2854, 2008.

Mlawer, E. J., and S. A. Clough, 1997. On the extension of RRTM to the shortwave region. In Proceedings of the Sixth Atmospheric Measurement (ARM) Science Team Meeting, CONF-9603149, pp. 223-226. U.S. Department of Energy, Washington, D.C

Pereira, E. B., et al., 2006. "Atlas brasileiro de energia solar." São José dos Campos: INPE, 2006. 60p.

Silva, G. R. Características de Vento da Região Nordeste – Análise, Modelagem e Aplicações para Projetos de Centrais Eólicas. Recife, 2003. Dissertação (Mestrado) Universidade Federal de Pernambuco.

Skamarock, W. C., et al., 2008. Description of the Advanced Research WRF Version 3 National Center for Atmospheric Research Boulder.

Skamarock, W. C., et al., 2005. A Description of the Advanced Research WRF Version 2, 2005.

Ward, J. H., 1963. Hierarchical grouping of optimized objective function. Journal of American Statistical Association, 58:236-244.