# Forecasting Solar Irradiance in Andalusia (Southern Spain) based on the WRF model: the role of the spatial averaging of the model outputs.

**V Lara-Fanego, J.A. Ruiz-Arias, F.J. Santos-Alamillos, A.D. Pozo-Vázquez* and J. Tovar-Pescador**

Department of Physics, University of Jaén, Campus Lagunillas Edif. A3, 23071, Jaén (Spain)

* Corresponding Author, dpozo@ujaen.es

**Abstract**

In this work, results of a study about the role of the spatial averaging of the WRF model DNI and GHI forecasting errors are presented. The study was conducted for two stations located in the region of Andalusia (southern Spain) along the period 2007 and 2008. Firstly, WRF three days ahead GHI and DNI forecasts, with 3 km of spatial resolution, were analyzed. Secondly, these forecasts were spatially averaged, in a post processing procedure, in increasing areas of up to 24 km x 24 km. Results showed that the spatial averaging increased notably the reliability of the WRF estimates in terms of the RMSE. For instance, for 24-hours-ahead forecasts and during summer, relative RMSE improved from 22%, when the evaluation station location grid point is considered (3 km x 3 km), to 15%, when averaging over the nearest 9 by 9 grid points (24 x 24 km).

## 1. Introduction

Much of the focus on sustainable energy is aimed at different ways of tapping the most abundant renewable resource: solar energy [1]. There are mainly two different ways for solar electricity production: by Solar Thermal Power Plants (STPP) and Photovoltaic plants (PV). The STPP use the direct normal solar irradiance (DNI) to convert solar energy through focusing receivers into heat, which is then used to drive a thermodynamic cycle and thereby produce electricity. PV systems enable direct conversion of global hemispheric irradiance (GHI) into electricity through semi-conductor devices.

A major challenge for the future will be the integration large scale solar yields into existing energy supply structures. The problem arises from the fluctuating character of the solar resource, as compared to conventionally generated electricity, and its dependence on non-deterministic weather patterns. Electricity obtained based on solar energy sources private costs are not competitive in the power generation systems because of its intermittent production creates negative externalities as those associated grid integration costs. The experience with wind energy shows that accurate wind energy forecasts can substantially reduce grid integration costs. Similarly, accurate information on expected solar irradiance can be used for the management of the electricity grids, for scheduling conventional power plants and also for decision making on the energy market. This will reduce the integration costs. But, unlike the wind power, solar yield forecast is still on an early state and very few works have dealt with the forecasting of the solar resources and its application for management of solar-based electricity power plant and grid integration strategies. Numerical Weather Prediction (NWP) models are the basis of solar yield forecasts with up 48 hours time horizon, the time range useful for grid integration and decision making in the energy market.

In this work, we present some preliminary results from a comprehensive evaluation study of the reliability of the Weather Research and Forecasting (WRF) [2] NWP model in forecasting the GHI and DNI in Andalusia (southern Spain) with a 72 hours time horizon. Particularly, in this work, we evaluate the role of the spatial interpolation of the WRF outputs in the reliability of the forecasts. Note that cloud are by large the most important issue regarding a reliable forecast of the GHI and DNI based on NWP models. This, mainly, because the difficulty of reproducing the clouds in space and time with precision. Post processing of the WRF output based on spatial filters may increase the reliability of the estimates. This because spatial aggregation of the solar radiation may reduce the forecast errors during episodes when clouds are presents but where not forecasted and vice versa. GHI forecasts are provide by the WRF model directly. On the other hand, in this work, the DNI forecasts were obtained based on a physical post processing [3], which uses WRF outputs and satellite retrievals. The study has been conducted at two validation locations around Andalusia along the years 2007 and 2008.

The work is organized as follows: in Section 2 the experiment design is presented. Particularly, this section includes the description of the study area and the observational data, the WRF model setup, the post-processing methodology to derive the DNI and, finally, the spatial averaging post-processing procedure. The evaluation of the model forecasts based on the observation is presented in Section 3. Finally, a summary of the results and some conclusions are provided in Section 4.

## 2. Methodology

### 2.1. Study area and ground data

The study was conducted for the region of Andalusia, southern Iberian Peninsula (Figure 1). Particularly, two validation ground stations located in Andalusia were analyzed: Cordoba and Andasol. The station of Córdoba (4.8506 º W, 37.8444 º N; 91 m.a.s.l.) is situated 6 km outside the city in an almost homogeneous area. Station of Andasol (3.1680 º W, 37.2180 º N; 1150 m.a.s.l) is located near Guadix (Granada). This late station is on the middle of one of the highest plateaus of Spain, surrounded by mountains reaching elevations of 3000 m.a.s.l. Both stations areas present a continental climate, with a wide range of annual temperatures from cold winters to dry and hot summers. Precipitations occur mainly in autumn and winter due to humid fronts travelling from west to east, ruled by Azores High; although convection clouds and thunderstorms are common in summer.
The present evaluation was carried out based on hourly GHI and DNI values collected in these stations along the years 2007 and 2008. Particularly, the periods August 2007 (representative of summer), September 15th to October 15th 2007 (autumn), 15th February to 15th March 2008 (winter) and April 2008 (spring) were analyzed.

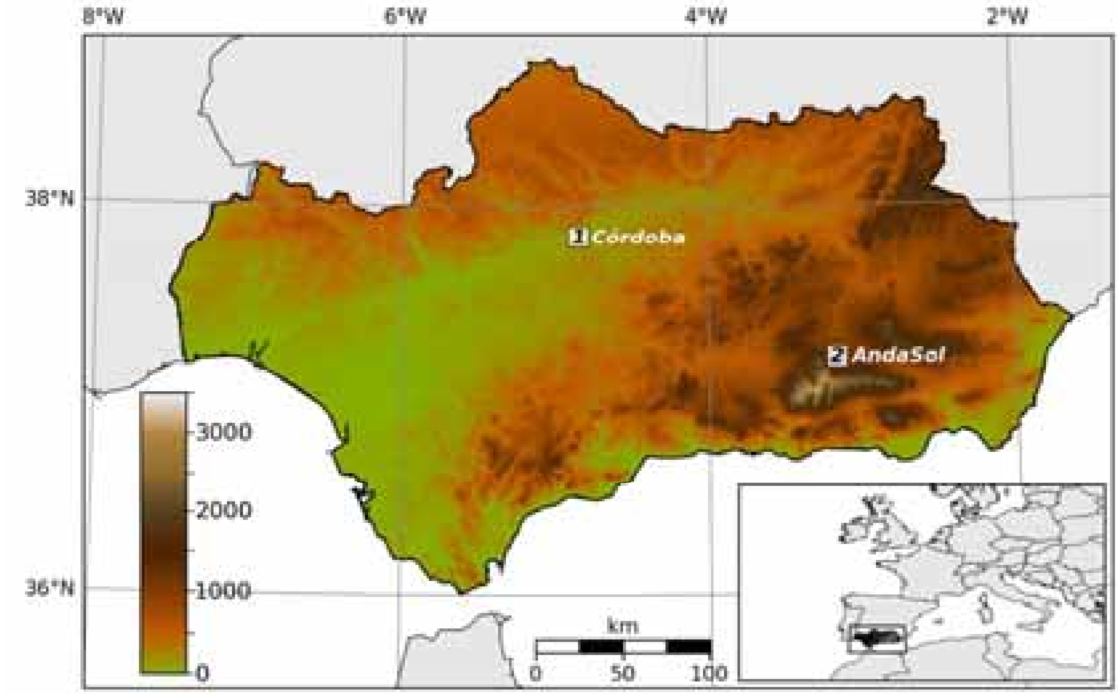


Figure 1. Study region and radiometric station locations.

**2.2. WRF setup**

Forecasts were obtained from a set of WRF (dynamic solver ARW, version 3) simulations from 0h to 72h UTC along the experiment periods. Global forecast system (GFS) from NOAA National Climatic Data Center (NCDC; http://nomads.ncdc.noaa.gov) with temporal resolution of 6 hours and spatial resolution of 1ºx1º were used as initial and boundary conditions. Three domains were used for temporal and spatial downscalling form GFS to a final grid covering the complete experimental area with 3km of horizontal grid resolution and 27 unevenly spaced vertical η-levels. Parameterizations were selected following Ruiz-Arias [4]. Particularly the scheme of Dudhia [5] was used as the shortwave radiation parameterization.

**2.3. Derivation of the DNI**

The DNI is not usually provided by NWP models, e.g. the WRF, but they do provide a comprehensive description of the state of atmosphere. This information, along with other information may be used to derive the DNI. In this work, the DNI forecasts are obtained based on a physical post processing (Ruiz-Arias et al, 2010) which uses WRF outputs and satellite retrievals readily available at the forecasting time. Particularly, column amount of ozone, from the EOS Aura OMI version 3 daily level 3 global 0.25º x 0.25º gridded data, and Aerosol Optical Depth (AOD) at 550 nm, from the MODIS/Terra Aerosol Daily L3 Global 1Deg CMG dataset, are used in the model. These aerosol and ozone values were maintained constant along the 72h forecasting time for each simulation.

### 2.4. Spatial interpolation post processing

As highlighted above, clouds are by large the most important issue regarding a reliable forecast of the GHI and DNI based on NWP models, mainly because of the difficulty of reproducing them in space and time with precision. Post processing of the WRF output based on spatial filters may increase the reliability of the estimates. This because spatial aggregation of the solar radiation may reduce the forecast errors during episodes when clouds are presents but where not forecasted and vice versa. To investigate further this issue, the reliability of the WRF forecasts at the validation station grid (same cell, sc, hereinafter), compared to the reliability of forecasts obtained by spatial interpolation of different grids nearby to the validation station, were analyzed. Particularly, in this study four different spatial averages were tested (Figure 1): x3 denote the three dots length square surrounding sc (equivalent to a 6km x 6km square), x5 represents 5 dots length square (12km x12 km), x7 (18 km x 18 km) and x9 (24 km x 24 km). In addition, a bilinear (bl) interpolation was also evaluated.

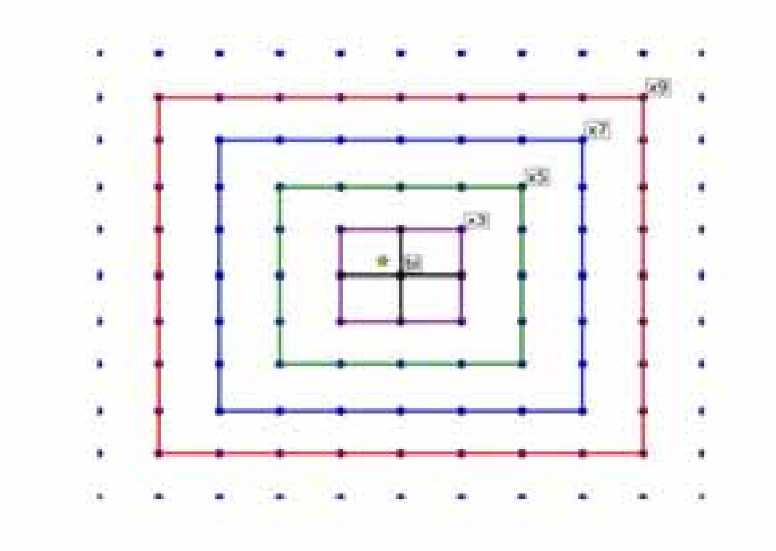


Fig.2. Configurations for the spatial interpolation of the WRF outputs. The star symbol represents the station location. Blue dots are the model grid points.

The evaluation of the forecasts was based on the mean bias error (MBE) and the root mean squared error (RMSE). The evaluation procedure was carried out for 24 hours (D1 hereinafter), 48 hours (D2) and 72 hours (D3) forecast horizon, for the different seasons of the year. An independent validation was also carried out as a function of the sky conditions. Particularly, three different sky conditions were considered: clear sky, cloudy and complete overcast conditions. These sky conditions were classified based on a daily clearness index $k_t$ [6]:

$$k_t = \frac{\sum_{i=1h}^{24h} I_i}{\sum_{i=1h}^{24h} I_{0i} \cdot e_{0i} \cdot \cos\theta_{zi}}$$

where I is the hourly global horizontal irradiance at the earth's surface, $I_0$ = 1367 $Wm^{-2}$ is the solar constant, $e_0 = (r / r_0)^2$ is the eccentricity correction factor at each hour and $\theta_z$ is the hourly average solar zenith angle. Values of daily kt greater than 0.65 were considered indicative of clear sky conditions, values between 0.65 and 0.4 were considered as broken clouds situations (cloudy conditions) whereas kt bellow 0.4 were proposed as complete overcast situations.

Finally forecasts were tested against persistence (a trivial reference model), in terms of the RMSE. Persistence forecasts were computed considering that the forecast values for a certain day were the observed values of the days before according to the forecast horizon (24h, 48h and 72h respectively).

**3. Results**

Overall, forecasting errors showed a strong dependence on the station (Córdoba presents the best results in almost all the analyzed aspects) and on the seasonal period (summer is the season with lowest errors, less cloudiness). Particularly, for GHI and for D1 forecasts, the relative RMSE values for the Andasol ranges from 33% in summer to 43% in winter, whereas for Córdoba ranges from 19% in summer to 35% in autumn. Regarding the sky conditions, the best results, as expected, are obtained for clear sky. For instance, for the Andasol station, for the GHI and the D1 forecasts, relative RMSE values ranges from 10% in autumn to 20% in summer. Similarly, for the DNI, RMSE values ranges from 12% in autumn to 24% in spring (in this station, aerosol summer loads play an important role). Both RMSE and MBE values considerable increases for cloudy and overcast conditions.

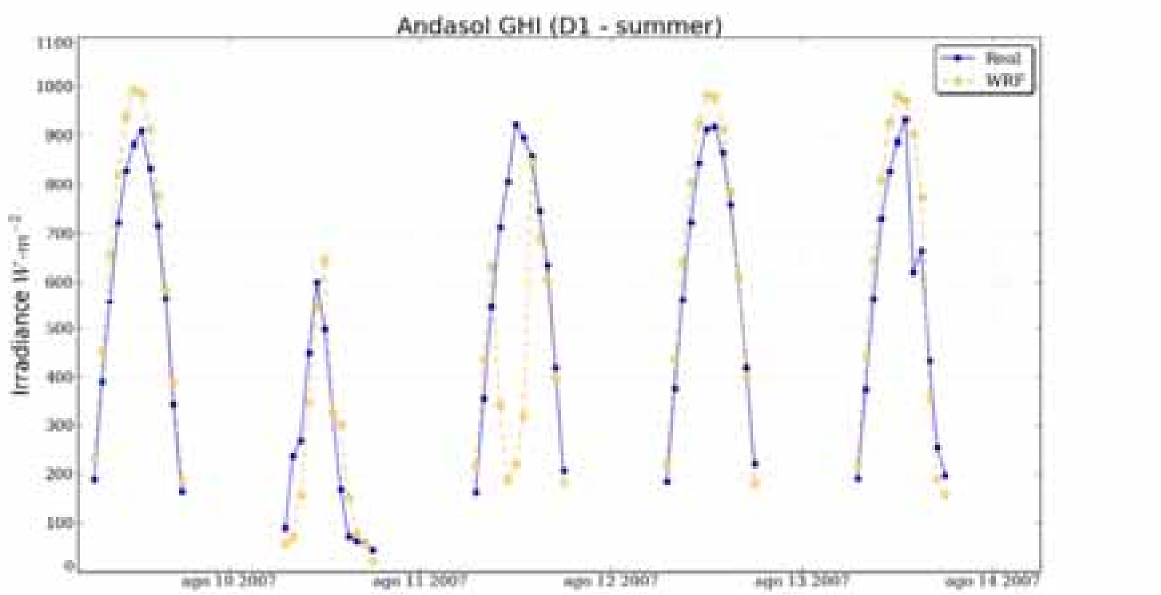


Fig. 3. WRF 24-hours-ahead GHI forecasts, for the days 9-13 Aug. 2007, for the Andasol station.

The most important factors in the accurate forecasting of GHI and DNI are aerosol and clouds. For instance, the Duhdia's shortwave scheme uses climatic values for aerosol and ozone. When the value of this parameter is lower than the real one, less radiation extinction is considered by the model and then WRF overestimates GHI [7]. This can be observed in Figure 2 for the Andasol station. Particularly, the figure represents the GHI 24 hours-ahead-forecasts (D1) during several days of the summer. Note that for the 1st, 4th and 5th days represented, the GHI is overestimated. On the other hand, note in Figure 2 that during 3th day, a non-existent cloud was forecasted by the model. However WRF is also able to forecast properly the presence of clouds, as can be seen in 2th day represented.

The former result refers to the grid where the validation station is located. Figure 4 present the results based on the spatial averaging of the WRF outputs. Particularly, the Figure 4 presents, for clear-sky conditions and for 24 hours ahead forecasts (D1), the DNI relative MBE and RMSE values for the different seasons of the year and for the different spatial interpolations. The general conclusion is that the proposed post processing methodology improves the results in terms of RMSE, in agreement with [8] and [9], whereas it tends to increase slightly the MBE. The analysis of the time series revealed that the periods where the estimations are improved, in terms of RMSE, are those which contain days with isolated clouds or trace clouds, not filtered by the $k_t$ criterion employed in this study. The most important improvement is found for the summer: from a relative RMSE value of 22% for sc to 15% for x9. In autumn, all days analyzed in this study were completely cloudless. As a consequence, the averaging post processing do not improved the estimates obtained for sc. Regarding cloudy conditions, the results revealed a lower improvement compared to clear sky conditions. For instance for the same case, Andasol D1 in summer, the relative RMSE values was 69% for the sc and 66% for x9. Finally it should be notice that in most of the cases the values of x5, x7 and x9 are quite similar in terms of RMSE.

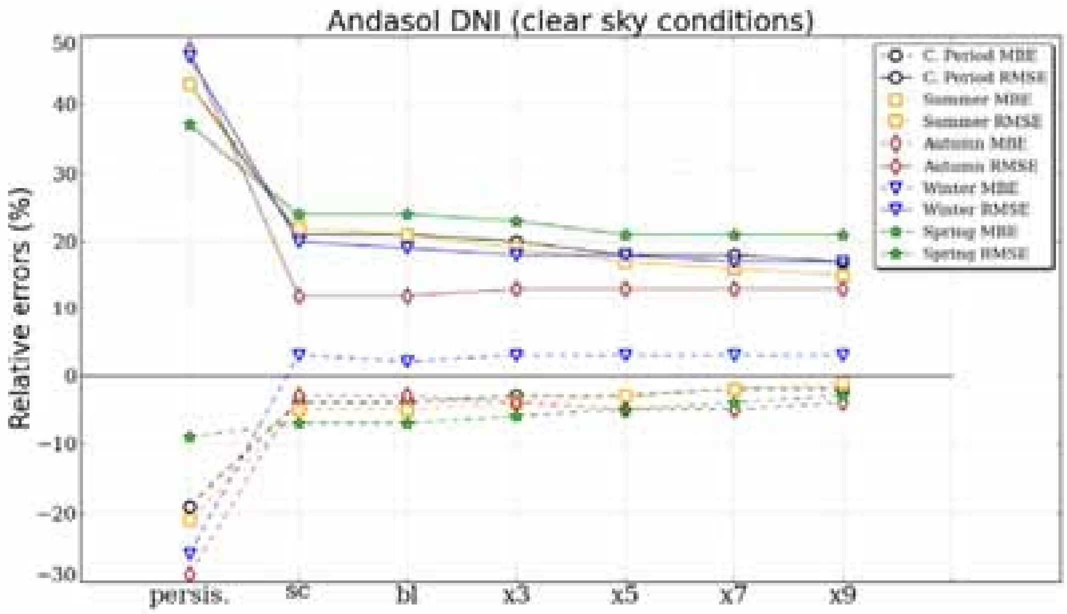


Fig.3. Andasol station clear sky conditions DNI forecasting evaluation results, based on the spatial interpolation, for 24 hours ahead forecasts (D1). The seasonal relative RMSE and MBE values are represented for the different spatial interpolations. Persistence model is also shown.

**4. Conclusions**

The role of the spatial averaging in the improvement of the WRF DNI and GHI forecasting errors were evaluated. The study was conducted for two stations located in the region of Andalusia (southern Spain) along the period 2007 and 2008. Firstly, three days ahead GHI and DNI forecasts, with 3 km of spatial resolution were analyzed. Secondly, these forecasts were spatially averaged, in a post processing procedure; in increasing areas of up to 24 km x 24 km. Results showed that the spatial averaging increases notably the reliability of the WRF estimates in terms of the RMSE. For instance, for 24 hours ahead forecasts and during summer, relative RMSE can be improved from 22%, when the evaluation station location grid point is considered (3 km x 3 km), to 15%, when averaging over the nearest 9 by 9 grid points (24 x 24 km). Therefore, this statistical post-processing has shown to be a valuable tool to improve the models forecasts. A slightly increment of the MBE was also obtained.

The ongoing work is focussed on further analyzing these preliminary results. Particularly, the analysis of the dependence of the best spatial averaging on the season of the year is now being undertaken. This because the study region presents different cloudiness characteristics depending on the seasons: from clouds associated to Atlantic frontal systems during the winter, to convective clouds in summer and clouds associated to Mediterranean origin frontal activity during spring and autumn. In addition, a correction of the systematic errors (bias correction) is now under study.

## References

[1] Szuromi, P., Jasny, B., Clery, D., Austin, J. and B. Hanson (2007). Energy for the long haul. *Science, 315*(5813), 781.
[2] Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.Y., Wang, W. and J.G. Powers (2008) A description of the Advanced Research WRF Version 3, NCAR/TN-475+STR, NCAR Technical Note, Boulder, Colorado, USA.
[3] Ruiz-Arias, J.A., Pozo-Vázquez, D., Lara-Fanego, V. and J. Tovar-Pescador (2010) A topographic downscaling prodedure for the solar irradiance in mountainous regions: application to the Weather Research and Forecast model. Submitted to the Journal of Applied Meterolo.
[4] Ruiz-Arias. J.A., Pozo-Vázquez, D., Sánchez-Sánchez, N., Montávez, J.P., Hayas-Barrú, A., Tovar-Pescador, J., 2008. Evaluation of two MM5-PBL parameterizations for solar radiation and temperature estimation in the South-Eastern area of the Iberian Peninsula, Il Nuovo Cimento, 31(5-6), 825-842
[5] Dudhia, J., 1989. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, J. Atmos. Sci. 46, 3077-3107.
[6] Perez, R., Ineichen, P., Seals, R., Zelenka, A., 1990. Making full use of the clearness index for parameterizing hourly insolation conditions, Sol. Energy, 45(2), 111-124
[7] Zamora, R.J., Solomon, S., Dutton, E.G., Bao, J.W., Trainer, M., Portmann, R.W., White, A.B., Nelson, D.W., McNider, R.T., 2003. Comparing MM5 radiative fluxes with observations gathered during the 1995 and 1999 Nashville southern oxidant studies, J. Geophys. Res. 108(D2), 4050.
[8] M. Girodo, Solarstrahlungsvorhersage auf der Basis numerischer Wettermodelle. Ph.D. dissertation, Oldenburg Univ., Germany, 2006.
[9] E. Lorenz, J. Hurka, D. Heinemann, and H.G. Beyer. Irradiance Forecasting for the Power Prediction of Grid-connected Photovoltaic Systems. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol.2, No.1, March 2009.