Rodríguez-Benítez, F.J., Arbizu-Barrena, C., Huertas-Tato, J., Aler, R., Galván-León, I.M., Pozo-Vázquez, D. (2020). A short-term solar radiation forecasting system for the Iberian Peninsula. Part 1: Models description and performance assessment. *Solar energy*, 195, pp. 396-412.

DOI: 10.1016/j.solener.2019.11.028

A short-term solar radiation forecasting system for the Iberian Peninsula. Part 1: Models description and performance assessment

Francisco J. Rodríguez-Benítez[a], Clara Arbizu-Barrena[a], Javier Huertas-Tato[b], Ricardo Aler-Mur[b], Inés Galván-León[b], David Pozo-Vázquez[a,*]

[a] MATRAS Res. Group, Department of Physic, University of Jaén, 23071 Jaén, Spain
[b] EVANNAI Res. Group, Department of Computing Science, Univ. Carlos III, 28911 Madrid, Spain

ARTICLE INFO

Keywords:
Short-term forecasting
GHI
DNI
MSG
NWP
Weather patterns

ABSTRACT

The ability of four models to provide short-term (up to 6 h ahead) GHI and DNI forecasts in the Iberian Peninsula is assessed based on two years of data collected at four stations. The models follow (mostly) independent ap-proaches: one pure statistical model (Smart Persistence), one model based on CMV derived from satellite images (Satellite), one NWP model (WRF-Solar) and a hybrid satellite-NWP model (CIADCast). Overall, results show Smart Persistence to be the best at the first lead steps, advective models (Satellite and CIADCast) at intermediate ones and the WRF-Solar at the end of the forecasting period. The break-even point between the advective models and WRF-Solar varies between 1 and 3 h for GHI and 3 and 5 h for DNI. Nevertheless, a detailed analysis shows enormous differences between models performance related to 1) the local geographic and topographic condi-tions of the evaluation stations; 2) the evaluated variable (GHI vs. DNI); and 3) the sky and synoptic weather conditions over the study area. Depending on the station and lead time, rRMSE values range from 25% to 70%for GHI and from 35% to 100% for DNI. For the same stations and leading time, rRMSE values for DNI are between 50% and 100% higher than the corresponding GHI counterparts. Depending on the synoptic pattern, rRMSE values are about 10/20% for GHI/DNI (3 h lead time, during high pressure conditions) to about 80/180% for GHI/DNI (during low pressure conditions). All models show a poor performance at a coastal station, at-tributed to a lack of ability to forecast clouds associated with sea-land breezes. To conclude, no single model proves to be the best performing model and, therefore, results show that the four models are, somehow, com-plementary. The advantages attained by this complementarity are further explored in a companion paper (Part II).

1. Introduction

Throughout the last years an enormous effort has been made in order to include solar energy as a feasible alternative to conventional energy sources. Many countries have already reached a notable solar energy share in their energy mixes and an important growth is expected over the next decades (International Energy Agency, 2018). Despite the fact that the Iberian Peninsula is one of the areas in Europe with the highest potential for the deployment of solar energy (Santos-Alamillos et al., 2017), nowadays the installed solar power is still modest. There are about 6.7 GW installed that only provide about 4.6% of the electricity demand in Spain (Red Eléctrica de España, 2017) and about 0.5 GW in Portugal (Redes Energéticas Nacionais, 2018). Nevertheless, an important growth is expected in the next years.

The use of some renewable energies to generate electricity, particularly solar energy, implies a great drawback: its integration into the electricity grid. Contrary to conventional generation, solar production is conditioned by weather, thus being highly intermittent. The problems related with solar energy grid integration have been the subject of several studies in recent years (Ela et al., 2011; Brouwer et al., 2014; Zhang et al., 2015a). One particular problem is that photovoltaic solar energy has lower spatial variability than, for instance, wind energy (Santos-Alamillos et al., 2014). This issue constrains the reduction of the intermittency by spatial aggregation. As a consequence, the power variability generated in systems with a high solar share is higher than in the case of system with high wind energy share (Lew et al., 2012). The importance of this solar integration issue is expected to increase as the participation of solar energy rises.

Currently, in addition to expensive storage-based solutions, solar radiation forecasts are a plausible way to mitigate the intermittency and

* Corresponding author.
E-mail address: dpozo@ujaen.es (D. Pozo-Vázquez).

facilitate solar energy grid integration. Solar power forecasts are used by grid operators for the scheduling of solar power plants and by their owners in order to participate in the electricity market and also for plant management (Antonanzas et al., 2017). For these purposes, solar forecasts are needed in time horizons ranging from minutes and hours to several days ahead. Short-term forecasts, i.e., up to 6 h ahead are particularly important for the management of concentrating solar power plants (Dersch et al., 2019), to participate in the intra-day electricity markets or forecasting ramp events (Zhang et al., 2015b). Therefore, the development of accurate solar radiation forecasting methods has become a key element for the increase of solar energy deployment and its grid integration (Renné, 2014; Brancucci Martinez-Anido et al., 2016; Haupt, 2018).

For short-term forecasting, reference methods are based on satellite imagery processing (Lorenz and Heinemann, 2012). Consecutive cloud index (CI) maps are used to estimate the cloud motion vectors (CMVs). Then the underlying atmospheric flow is estimated to finally provide the cloudiness and solar radiation forecasts. This methodology was firstly proposed in the context of the Meteosat Second Generation (MSG) images by Beyer et al. (1994), but there have been several improvements since then (see for a review (Diagne et al., 2013; Inman et al., 2013)). Recently, improved methods for CMV estimation have been proposed (Schroedter-Homscheidt et al., 2018; Wang et al., 2019). These approaches rely on the use of cloud physical properties derived from the MSG data instead of the (CI) images. Sirch et al. (2017) have proposed a cloud-specific method for direct normal irradiance (DNI) nowcasting.

On the other hand, numerical weather prediction (NWP) models have also been used to provide short-term solar radiation forecasts (Mathiesen and Kleissl, 2011; Lara-Fanego et al., 2012; Wolff et al., 2016). Although NWP models are able to forecast the arrival of frontal systems and their associated cloudiness, they still lack the ability to reproduce cloud amount and location accurately (Deng et al., 2014; Arbizu-Barrena et al., 2015). Improvements in subgrid-scale cloud-radiation feedback, for instance shallow cumulus, yield notable enhancements in global horizontal irradiance (GHI) forecasting derived from NWP models (Lee et al., 2017).

Overall results from the comparison between satellite and NWP based short-term GHI forecast show that satellite tends to provide better forecasts in the first hours, after which NWP models show lower errors. Nevertheless, the break-even point varies among studies. Some studies (Kühnert et al., 2013) showed that this point is at a horizon of 3 h, approximately. However, Wolff et al. (2016) and Perez et al. (2010), found this break-even point to be around 5 h. Lee et al. (2017) evaluated several short-term GHI forecasting techniques (ranging from pure statistical, to CMV based and NWP models). Evaluation was conducted under different sky conditions, finding that the performance of the different models depended on the weather conditions. Even in some evaluations over complex terrain, satellite approaches have been shown to be unsuitable (Guillot et al., 2012; Arbizu-Barrena et al., 2017). These discrepancies arise because of differences between the evaluated models, the study area or the period of analysis. It should be highlighted that, unlike GHI, very few studies have compared different approaches for short-term DNI forecasting. Only recently Dersch et al. (2019) compared several satellite and NWP based DNI nowcasts for two Mediterranean locations. Results for one of the stations showed that the NWP-based method outperforms the satellite-based nowcast for horizons greater than 2 h, approximately.

Recently, several models have been proposed with the aim of combining the capabilities of satellites and NWP models. For instance, the CIRACast forecast algorithm (Miller et al., 2018) uses winds derived from a NWP to advect cloud fields estimated based on satellite images, providing in this way short-term solar radiation forecasts. NWP models have the ability not only to advect but also to diffuse cloudiness information. This ability is used by the MADCast model (Descombes et al., 2014), where clouds are transported and diffused in three dimensions within the Weather Research and Forecasting (WRF) NWP model (Skamarock et al., 2008). Lee et al. (2017) showed that this method outperforms other models at short lead times except on rapidly cloud forming, growing or dissipating situations. Arbizu-Barrena et al. (2017) presented a similar hybrid approach (CIADCast) which also advects and diffuses the MSG CI images within the WRF model. Other authors have proposed probabilistic approaches to improve short-term solar radiation forecasting (Boland and Grantham, 2018).

Finally, other approaches aimed at improving short-term solar forecasts rely on the development of model blending approaches. Typically, multiple models are blended together using some statistical technique. This may provide a synthetic model with enhanced forecasting capability compared to the individual models. Linear combination of satellite and NWP models were evaluated in Lorenz et al. (2012). In Wolff et al. (2016) irradiance measurements, satellite and NWP derived short-term solar forecasts were used as input of a support vector regression machine learning algorithm in order to derive a blended forecasting model. In that case, results showed that the blending approach outperforms the individual models for all the forecasting horizons.

The purpose of this work (Part I) and the work presented in the companion paper (Part II) is twofold: a) to provide a comprehensive assessment of the performance of four short-term GHI and DNI state-of-the-art forecasting models, based on different approaches, in the Iberian Peninsula; b) to explore the benefits obtained by developing optimal blending of these models using machine learning approaches.

This first part includes the description of the four models and the evaluation of their ability to provide single forecasts. The four models rely on different approaches: a pure statistical model (Smart Persistence), a model based on satellite images (Satellite), a NWP model (WRF-Solar) and a hybrid satellite-NWP model (CIADCast). Specific analyses are conducted to evaluate their performance dependence on the sky and synoptic weather conditions.

In the companion paper (Part II), different model blending approaches are presented and the improvement attained in the forecasting accuracy is discussed. Additionally, in this second part, regional forecasting models are specifically developed and evaluated.

This paper is organized as follows: Section 2 presents the study area and describes the ground and remote sensing dataset used. Section 3 introduces the four forecasting models. Section 4 describes the metrics and procedure involved in the models assessment. In Section 5 the results are presented and discussed and, finally, in Section 6 a summary of the main results and some conclusions are provided. Additional analyses are discussed in Appendix A and Appendix B.

2. Datasets

2.1. Evaluation sites and measurements

One-minute time resolution solar irradiance data collected at four radiometric stations representative of the central (Madrid station), southern/southwestern (Seville and Jaen) and western (Lisbon) areas of the Iberian Peninsula (Fig. 1) are used. Except in Madrid station, where only GHI and DNI are available, the three components of the solar radiation, GHI, diffuse horizontal irradiance and DNI, are gathered. Given the characteristics of the database, an uncertainty of about $\pm 2.5\%$ for the DNI and about $\pm 3.0\%$ for the GHI data can be assumed following Sengupta et al. (2015). These uncertainties accounts for that of calibration and other sources of measurement errors. The observation covers the period from March 2015 to February 2017, (i.e., two years). Radiometers description and data quality procedure were fully detailed in Rodríguez-Benítez et al. (2018). Data associated with solar zenith angle >75° are discarded. Only the records available simultaneously at the four stations are considered.

The semi-permanent anticyclone of the Azores Island rules the atmospheric circulation over the study area (Trigo et al., 2002). Changes

Fig. 1. Shaded relief map of the study area and location of the evaluation stations (garnet dots).

in the location and intensity of the Azores anticyclone cause a marked seasonal climate variability in the study area. During winter, the Azores anticyclone is at lower latitudes, allowing the study area to be influenced by westerly winds. Thus, transient synoptic perturbation reaches the study area usually from October to March. During summer, this anticyclone migrates northward, blocking the westerly circulations. This marked seasonality is clearly observed also in the solar radiation variability (Santos-Alamillos et al., 2012). In a recent paper (Rodríguez-Benítez et al., 2018), the weather patterns associated with different solar radiation modes of variability in the study area were analyzed.

From a topographic point of view, the study area may be split into two different parts. The western area is a homogeneous region, with wide valleys open to the Atlantic Ocean. The Lisbon and Seville stations are located in this part (Fig. 1). On the other hand, the central and southern area shows a more complex topography, with several mountain ranges in the south (where the Jaen station is located) and a plateau in the center (where the Madrid station is located). Although local cloudiness may be found in mountain areas (orographic clouds) or in coastal areas as Lisbon station (convective clouds), most of the cloudiness over the study area is related with transient weather patterns that affect uniformly the entire study area (Trigo et al., 2002; Rodríguez-Benítez et al., 2018).

2.2. Satellite retrievals

CI maps, obtained based on the Heliosat-2 method (Rigollier et al., 2004), are used for the satellite-based and CIADCast (see Section 3.4) forecasting models. MSG satellite images (Schmetz et al., 2002) have 15 min temporal resolution and the projection applied produces a spatial resolution of about 5 km. According to the Heliosat-2 method, GHI is computed multiplying the clear-sky index (K_c) value by the clear-sky irradiance. The European Solar Radiation Atlas (ESRA) clear-sky solar irradiance model (Rigollier et al., 2000) is used, with the worldwide monthly climatology Linke turbidity parameter proposed by Remund et al. (2003). DNI estimates are obtained from GHI values using the DirIndex method (Perez et al., 2002), using the ESRA model with the same parameterization.

Cloud top height (CTH) estimates from the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Satellite Application Facility on Climate Monitoring (CMSAF) have been used in order to compute the CIADCast model forecasts. This CTH database, which is available for the whole study area, provides estimates of the height at the top of the highest cloud. A root mean square error (RMSE) of about 4 km, and a bias (BIAS) of about 2 km, has been reported for this dataset in Hamann et al. (2014).

3. Forecast models description

Six-hours ahead GHI and DNI forecasts provided by four different models are assessed in this work. The models are described below.

3.1. Smart Persistence

Smart Persistence model is used as the reference.

$$I(t) = \frac{I_0}{I_{clear-sky}(t_0)} \cdot I_{clear-sky}(t), \quad (1)$$

where I_0 is the 1-min time resolution measured irradiance at the initial time (t_0), $I_{clear-sky}(t_0)$ is the corresponding clear sky irradiance and $I_{clear-sky}(t)$ is the clear sky irradiance at a future time t. I stands for both GHI and DNI. The ESRA (Rigollier et al., 2000) clear-sky model is used to compute $I_{clear-sky}$.

3.2. Satellite

The CI images are used to provide satellite-based GHI and DNI estimates, as described in Section 2.2. Firstly, the open source Particle Image Velocimetry (OpenPIV) (Mori and Chang, 2003) algorithm, with a square window of 41 px, is used to estimate the CMVs field. In a second step, and similarly to Nonnenmacher and Coimbra (2014), the streamlines based on the CMVs are derived. The CI forecasts are then obtained following a streamline backwards the corresponding time steps. Irradiances are finally obtained as explained in Section 2.2. Fig. 2 represents the area of the MSG images analyzed to derive the CMV fields.

3.3. WRF-Solar

GHI and DNI forecasts are obtained from the WRF-Solar (Jimenez et al., 2016), a particular physical configuration of the WRF model version 3.6 devised for solar energy applications. Some of the parameterizations used are the Thompson aerosol-aware microphysics scheme (Thompson and Eidhammer, 2014) with monthly climatological aerosol which accounts for that emitted by natural and anthropogenic sources. It is also used the Rapid Radiative Transfer Model for General Circulation Models (RRTMG) short- and long-wave radiation parameterizations (Iacono et al., 2008), called every 5 min, which are coupled with cloud physics parameterization. Direct effect of rural aerosol is taken into account in order to fully connect the cloud-aerosol-radiation system. A new option for shallow convection (Deng et al., 2014) is included in WRF-Solar, thus the cumulus parameterization is switched off.

The model is configured with one domain of 5 km spatial resolution (30 s of model time-step) and 50 vertical levels. The domain is drawn in

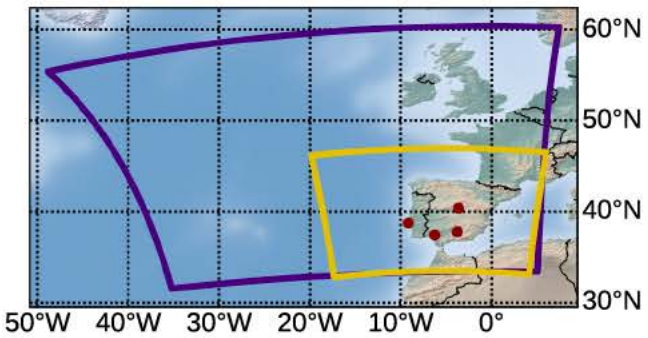

Fig. 2. Domains of the CIADCast and WRF-Solar (yellow line), and Satellite-based model (purple line). The garnet dots indicate the location of the evaluation stations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. Initial and boundary conditions are taken from the National Centers for Environmental Prediction (NCEP) GFS [0.5 Deg] dataset (NCEP, 2006). For each day, three 18-h simulations are run starting at 00, 06 and 12 UTC, discarding the first 6 simulated hours as spin-up. Outputs are saved every 15 min, at the moment when the satellite passed over the study area.

3.4. CIADCast

The CIADCast model (Arbizu-Barrena et al., 2017) is here also assessed. CIADCast is a hybrid model which aims to combine the accuracy of the cloud representation in satellite images with the dynamical capabilities of the WRF NWP model. In this work, CIADCast forecasts are run simultaneously with the WRF-Solar model, as explained in Section 3.3. In this model, CI retrievals from satellite are advected and diffused to obtain cloudiness forecasts and, therefore, radiation forecasts. To this end, firstly, the CI maps are interpolated, using nearest neighbors criteria, in order to adapt the initial satellite domain to that of the WRF-Solar runs. Secondly, the interpolated CI maps are ingested in the model for every satellite gathering time at a particular vertical layer of each column of the model. In this work, unlike in Arbizu-Barrena et al. (2017), the MSG CTH information is used to select this vertical layer. Thus, the CTH images are also interpolated to the WRF-Solar grid. Empty pixels are filled with nearest values so all of the columns in the WRF-Solar model have CTH value. The CI maps are ingested in WRF-solar as mass mixing ratios (each grid-cell value is divided by the model dry air mass of the grid-cell), and thus the total amount of CI is conserved during the simulation. The WRF-Solar model advects and diffuses the CI values as dynamical tracers mainly horizontally, but also vertically. Then, the CI forecasts are de-normalized with the dry air mass of the corresponding grid-cell and the sum of each column values is computed to obtain again two-dimensional CI maps. Finally, the GHI and DNI forecasts at the stations are obtained as explained in Section 2.2.

During the simulation run, the model is stopped at the satellite retrieval time, then the CI image is inserted, and the simulation is restarted for the next 6 h. Outputs are saved every 15 min, at the moment when the satellite passed over each evaluation station.

4. Evaluation procedure

Evaluation of the forecasts with ground data is carried out at 15 min time steps, according to the time when the MSG images were obtained for each station, and at 1-min time resolution. Table 1 shows the number of forecasts available (for each model) as a function of the forecasting horizon. Forecasts are evaluated only when they are available from all the models and for all stations at a given forecasting horizon. A total of 116458 forecasts were obtained, ranging from about 7800 for the first horizon (15 min) to 1500 at the 6 h forecasting horizon.

Table 1
Number of predictions obtained at each forecasting horizon.

Horizon (minutes)	Number of forecasts	Horizon (minutes)	Number of forecasts
15	7779	195	4145
30	7594	210	3785
45	7427	225	3451
60	7248	240	3140
75	7019	255	2856
90	6754	270	2607
105	6433	285	2364
120	6072	300	2094
135	5722	315	1837
150	5306	330	1592
165	4884	345	1384
180	4484	360	1175
		Total	107152

No post-processing procedure has been applied to any of the models, i.e., the outputs of the models are assessed directly. However, in the companion paper (Part II), machine learning algorithms are applied to the models retrievals, allowing the benefits obtained by using advanced post-processing procedures to be quantified.

The performance of the models is assessed based on several metrics: the RMSE, the Forecast Skill (FS) based on the comparison between the RMSE values produced by the forecasting models and those from Smart Persistence model (Coimbra et al., 2013), the relative RMSE (rRMSE), the relative mean absolute error (rMAE) and BIAS error defined as:

$$\mathrm{RMSE}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right)^2} \tag{2}$$

$$\mathrm{FS}(t) = 1 - \frac{\mathrm{RMSE}_{\mathrm{forecast}(t)}}{\mathrm{RMSE}_{\mathrm{Smart\ Persistence}(t)}} \tag{3}$$

$$\mathrm{rRMSE}(t) = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right)^2}}{\frac{1}{N}\sum_{i=1}^{N} I_{\mathrm{measured}(t),i}} \cdot 100 \tag{4}$$

$$\mathrm{rMAE}(t) = \frac{\frac{1}{N}\sum_{i=1}^{N}\left|I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right|}{\frac{1}{N}\sum_{i=1}^{N} I_{\mathrm{measured}(t),i}} \cdot 100 \tag{5}$$

$$\mathrm{BIAS}(t) = \frac{1}{N}\sum_{i=1}^{N}\left(I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right), \tag{6}$$

where, N is the number of data involved in the calculation, I is the irradiance (GHI or DNI) and t the time step. Independent yearly and seasonal analyses are conducted.

In addition, and in order to gain insight into the models performance dependence, additional analyses are conducted. Firstly (Section 5.2), the ability of the models to forecast GHI and DNI variability is assessed. The direct deterministic forecast accuracy of the solar has important limitations, caused by the double penalty problem of the cloud forecasting. Accurate forecasts of the solar variability, instead of deterministic forecasts, can be used to mitigate solar power intermittency, facilitating the solar energy grid integration (Perez and Hoff, 2013). As opposed to deterministic forecasts (assessed in Section 5.1), solar variability predictions involve the use of specific scores that account for this variability along a certain reference period. To this end, the nominal variability is used (Perez et al., 2016), defined as:

$$\sigma_{\Delta K_c} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\Delta K_{c,i} - \overline{\Delta K_c}\right)^2}, \tag{7}$$

where $\Delta K_{c,i}$ is the difference between two consecutive values of the clear-sky index and $\overline{\Delta K_c}$ is the mean over the considered period. The formula stands for both GHI and DNI variability. For each forecasting horizon and starting every 15 min, the observed and forecasted nominal variability is computed, when at least one sample is available, using a reference period of 3 h. Then, the correlation between the observed and forecasted values are used to assess the skill of the different models to forecast GHI and DNI variability.

Finally, in Section 5.4, the models' performance dependence on the weather conditions is assessed. The type of analysis is usually addressed using the clearness index (Lorenz et al., 2016; Wolff et al., 2016; Wang et al., 2019). Other authors used specific sky conditions in the analysis (Lee et al., 2017) or cloud regimes (McCandless et al., 2016). Here we follow a different approach based on the use of weather types. To this

end, the four weather types (WTs) of the annual analysis described in Rodríguez-Benítez et al. (2018) are used. Notably, the first type, "WT 1", accounts for presence of a synoptic perturbation over the study area. In terms of cloudiness, this means overcast conditions. The second type, "WT 2", represents transient weather conditions, i.e., those usually observed some days before or after a synoptic perturbation that passes over the study area. Cloudiness shows intermediate values for this WT. The third type, "WT 3", accounts for the presence of moderate high pressure anomalies over the study area, which allows the development of local weather features as, for instance, convection. Cloudiness is highly variable among the stations for this WT. Finally, "WT 4" accounts for the presence of a high pressure system over the whole study area. As a consequence, clear sky is mostly observed at the four stations. Each forecast is assigned to one of these four categories, according to the WT observed at the time when the prediction was issued.

For the sake of completeness, the models' performance dependence on the sky conditions are also evaluated.

5. Results

5.1. General evaluation

Fig. 3 shows the rRMSE and rMAE values for the four prediction models and the four stations, depending on the forecasting horizon. Values are shown for both GHI and DNI.

GHI prediction errors (upper row Fig. 3) show a notable variability among stations. Seville station (fourth column) shows the lowest errors (about 30% rRMSE value at 2 h-ahead forecasting horizon) followed by Jaen and Madrid. The Lisbon station (second column) shows considerably higher errors (about 45% rRMSE at 2 h-ahead). Similarly, the lowest rMAE values are found at the Seville station and the highest at Lisbon. Regarding rRMSE, Smart Persistence is the best performing model at the first horizon (i.e., 15 min ahead) for all stations, except at Seville. At this station, the Satellite model performs as the best, up to 2 h and 30 min, after which WRF-Solar provides the best forecasts. For the Jaen station, Smart Persistence provides the most accurate forecast in the first 2 h and 30 min, after which, as in the case of Seville, WRF-Solar outperforms the rest of the models. For the Madrid station, the break-even point between Satellite and WRF-Solar forecasts is observed at the first leading hour. Between the 15 and 60 min forecasting horizons, the Satellite model provides the best forecasts. Finally, for the Lisbon station, CIADCast and Smart Persistence show similar performance in the first 2 h. From 2 to 5 h, the Satellite model shows slightly better forecasts, while the WRF-Solar performs as the best after 5 h, approximately. Regarding the rMAE metric, and except in Jaen, the Smart Persistence model outperforms the rest of the models in the first 2–3 h, after which WRF-Solar shows a superior performance.

DNI forecasts errors (lower row Fig. 3) are considerably larger than GHI ones. Notably, both rRMSE and rMAE are about 60% higher than their GHI counterparts for the same leading time at Seville and Jaen stations and about 75% higher for Madrid and Lisbon stations. At the Seville station, the performance of the different models, in terms of the rRMSE, is similar to the case of GHI. Also, at the Jaen station, the models show a performance similar than to the GHI case, although the break-even point between Smart Persistence and WRF-Solar is located at about 3 h (half an hour later than in the GHI case). The performance of the different models differs from the GHI cases at the Madrid and Lisbon stations. At Madrid, Smart Persistence outperforms the rest in the first 4 h and a half, after which WRF-Solar provides slightly better results. At the Lisbon station, Smart Persistence is the best performing model in the first 3 h. From this time, Satellite provides the most accurate forecasts. Only around the end of the forecasting period, WRF-Solar provides a similar performance to Satellite. Regarding the rMAE, the performance of the four models is similar to the GHI case for all stations except Madrid, where Smart Persistence performs best for all the forecasting horizons.

Fig. 4 shows the RMSE values produced by the four models at the four stations, for both GHI and DNI predictions. The performance of the different models, in terms of RMSE, is similar than in case of rRMSE (Fig. 3). For GHI predictions, RMSE values show a low dependence on the forecasting horizon. Values keep below 230 $W \cdot m^{-2}$, except in Lisbon station. On the other hand, the DNI RMSE values show a notable upward trend with the forecasting horizon, reaching values between 330 and 400 $W \cdot m^{-2}$, approximately, depending on the station, at the final of the forecasting window.

The CIADCast, Satellite, and WRF-Solar model forecasts are compared with those from the Smart Persistence model using the FS score. Fig. 5 shows the results for both GHI and DNI forecasts. As may be expected based on the analysis showed at Figs. 3 and 4, notable differences in the skill are observed depending of the evaluation station and the variable of interest. Overall, the FS values for GHI forecasts are slightly larger than those for DNI. At the Jaen station and both for GHI and DNI, the WRF-Solar shows positive skill after 3 h. Satellite and CIADCast models show positive skill only at the end of the forecasting period. At Lisbon station, Satellite model shows positive skill for both the GHI and DNI after 3 h approximately, and WRF-Solar after 4 h. At Madrid station, WRF-Solar shows positive skill for GHI predictions from the first lead hour, while Satellite shows positive skill below one h.

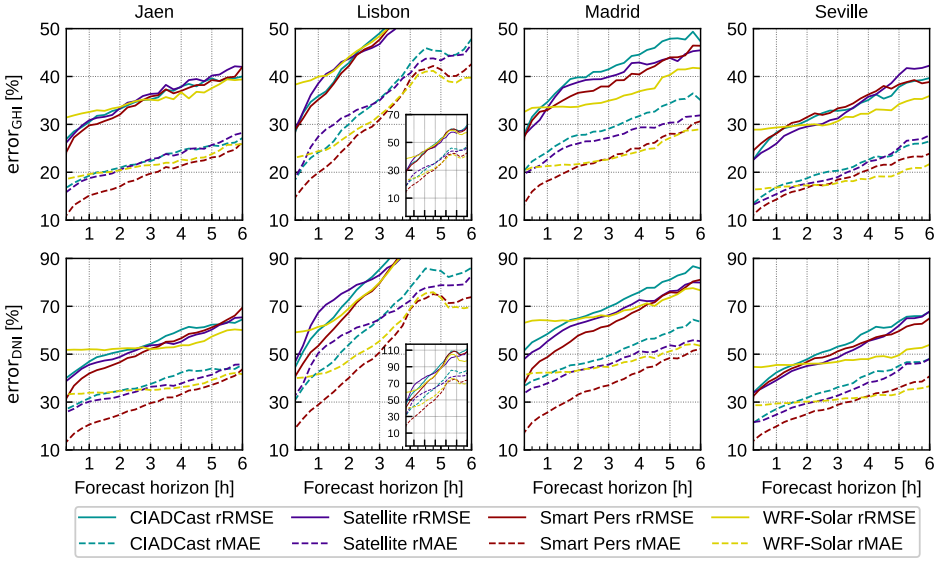

Fig. 3. rRMSE (solid lines) and rMAE (dashed lines), depending on the forecasting horizon, for predictions of GHI (upper row) and DNI (bottom row). Values (in %) are displayed for CIADCast, Satellite, Smart Persistence, and WRF-Solar model with a turquoise, purple, garnet, and yellow line, respectively. The first, second, third, and fourth columns show the results of Jaen, Lisbon, Madrid, and Seville station, respectively. The small figures for the Lisbon stations (low right hand corner) represent the same values but using a different vertical scale. This allows representing the whole range of values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

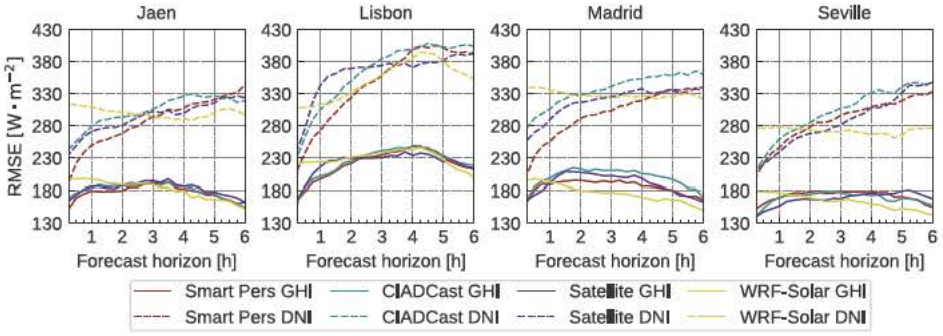

Fig. 4. RMSE values for GHI (solid lines) and DNI (dashed lines) forecasts, depending on the forecasting horizon. The color code is the same used in Fig. 3.

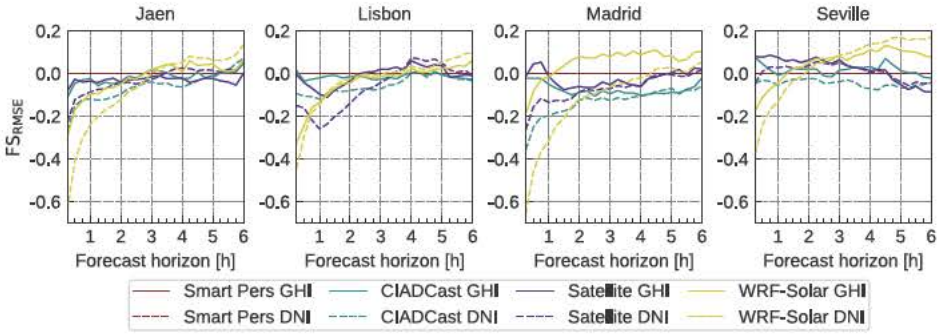

Fig. 5. Forecast skills (FS) of the GHI (solid lines) and DNI (dashed lines), depending on the forecasting horizon. The color code is the same used in Fig. 3.

Finally, Seville station shows the best performance for the different models. In case of GHI predictions, both Satellite and CIADCast models show positive skill up to 5 h ahead (approximately). WRF-Solar shows positive skill at lead times higher than 1 h and 2 h in case of GHI and DNI predictions, respectively, reaching, with the latter, values close to 0.2 at the end of the forecasting window.

To sum up, it should be highlighted that Smart Persistence proves to be a very competitive model regarding all the metrics here evaluated, particularly for DNI predictions.

To conclude this first evaluation analysis, Fig. 6 shows the BIAS error values for both GHI and DNI. GHI BIAS values are relatively low and negative (between 20 and -80 W·m^{-2}) and with almost no dependence on the forecasting horizon, except at Lisbon. At this station, a considerable increase of the BIAS values with the forecasting horizon is observed, reaching maximum positive values at about 4 h and 30 min ahead. Results for DNI are qualitatively similar. The main difference is observed for WRF-Solar, since the BIAS errors are considerably higher and positive for all stations. This feature is particularly outstanding at the Madrid and Lisbon stations. For the latter, all the models except Satellite show positive BIAS values at longer forecasting horizons, indicating a lack of ability to predict cloudy conditions at these lead times.

The differences in model performance among stations can be explained according to topographic and geographic conditions, described in Section 2.1. Firstly, models performance is particularly poor at the only coastal analyzed station: Lisbon. This station also shows the steepest increment of the errors with the forecasting horizon. The rapid development of convective clouds in the area of the station, that are particularly difficult to forecast, may explain this result. The role of this type of cloud is further discussed in Section 5.4.

After Lisbon, Madrid shows the greatest forecasting errors. The geographic location of this station can explain this result. Most synoptic perturbations enter in the study area from the Atlantic Ocean. In their way to the central Iberian Peninsula, they interact notably with mountain ranges. Pure advective models, such as Satellite and CIADCast, had been reported to be unable to account for this interaction properly (Arbizu-Barrena et al., 2017), providing a poor performance at longer forecasting horizons. Only the WRF-Solar model is able to account for this interaction, providing the best results. On the other hand, the lowest errors are found at Seville station, where advective models (Satellite and CIADCast) show a good performance. These results can be explained based on the topographic features of this station, at the entrance of a wide valley (Guadalquivir) open to the Atlantic Ocean. There are no relevant mountain ranges between the station and

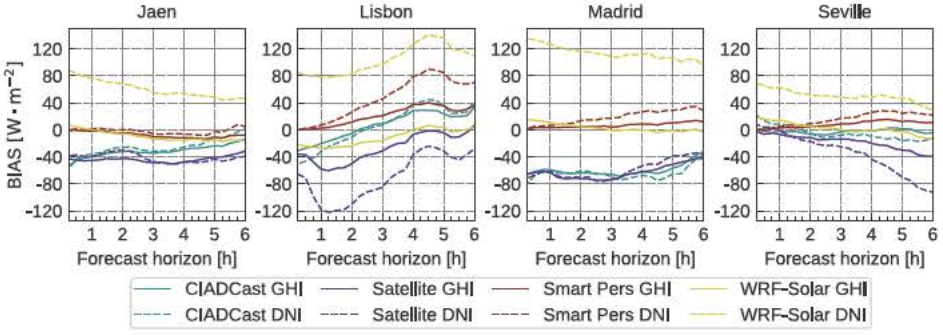

Fig. 6. As in Fig. 4 but for. the BIAS score.

the usual storms track direction in the study area. Finally, the Jaen station is located up in the Guadalquivir Valley, sharing some features with the Seville station. Nevertheless, it is surrounded by notable mountain ranges, which reduce the performance of the advective models and increase the models' forecasting errors in general.

In addition to the annual analysis, a study of the seasonal dependence on the models performance is conducted Appendix A. For the GHI (Fig. A.12) and for all the stations except Lisbon, the highest forecasting errors are observed for winter, followed by spring and autumn, being summer the season with the lowest errors. Similar results are observed for the DNI (Fig. A.13). Lisbon station results differ considerably. At this station, summer forecasting error values are of similar magnitude than winter ones (even higher for the DNI), showing the steepest increase with the forecasting horizon. Also for spring and autumn, forecasting errors at Lisbon are slightly higher than those for the other stations. These results can be associated with presence of land-sea breeze, since this local phenomenon is more frequent in summer, followed by spring and autumn. This issue is discussed in Section 5.3. In general, differences between models are low, except in winter, when WRF-Solar beats Smart Persistence at lead times higher than 1.5 or 3 h, depending of the station. These conclusions are also supported by the analysis of the Forecast Skill scores (Fig. A.14). Finally, for spring and winter, the WRF-Solar shows a superior performance, especially for the DNI (Fig. A.14).

Comparison of the results here presented with those obtained in similar work is difficult due to, particularly, the differences in the datasets (study region climatology, length of the datasets). Nevertheless, results are in close agreement with those reported in Wang et al. (2019) and Dersch et al. (2019). The main difference is a higher dependence of the forecasting error with the lead time.

5.2. Variability forecasts evaluation

Fig. 7 shows the correlation coefficient, as a function of the forecasting horizon, between the observed and forecasted nominal variability index. Therefore, this figure indicates the skill of the different models predicting GHI and DNI nominal variability. Correlation values decrease with the forecasting horizon for all stations and models, as may be expected. Note also that correlation values tend to be higher for GHI compared to DNI. Up to 4 h lead time, Smart Persistence clearly outperforms the rest of the models, showing correlations higher than 0.4. For forecasting horizons longer than 4 h, results greatly vary between stations. For instance, at Seville and particularly at Lisbon, WRF-Solar is shown to be the best performing model both for GHI and DNI. On the contrary, for the Jaen station and in terms of GHI, Satellite provides the highest correlations. For the rest of the cases (i.e., Jaen DNI and Madrid GHI and DNI) WRF-Solar, Satellite and Smart Persistence models show similar results.

5.3. Forecasts dependence on the time of day

Fig. 8 shows the rRMSE values depending on the time of the day for two representative forecasting horizons: 1 and 4 h. Fig. 9 shows the BIAS error for the 4 h horizon.

According to Figs. 8 and 9, no significant dependence on the time of the day is observed, except at the Lisbon station (Fig. 8). For this station, error dramatically increases during evening hours (approximately from 16:00 onward).

The time-of-day dependence forecating errors at Lisbon may be attributed to a daily cloudiness cycle. This cycle is caused by availability of moisture, along with the frequent occurrence of sea-land breezes at this coastal station. This local circulation reaches its maximum in the afternoon hours. As observed in Fig. 9, all models, and for both GHI and DNI, show positive BIAS in the evening, when convective clouds development is more frequent. This indicates a lack of ability of all the models to forecasts convective cloudy conditions, even at 1 h forecasting horizon (Fig. 8). The lack of ability of the statistical model (Smart Persistence) or advective models (Satellite and CIADCast) to accurately forecast cloudiness may have been anticipated (Sirch et al., 2017). But this is not the case of the WRF-Solar model, which also shows scant skills in this regard. The accurate representation of local thermally-driven circulation still presents a challenge for numerical weather models. For instance, Avolio et al. (2017) evaluated the performance of the WRF model when the local weather in a coastal area in southern Italy was simulated. Results showed that the WRF model had a considerable better performance for synoptic scale forcing than for local scale (i.e., breeze) conditions. In addtion to this issue, the WRF-Solar domain set-up and shallow cumulus parameteriztion here used may also account for these results.

5.4. Forecasts dependence on the synoptic weather conditions

Figs. 10 and 11 show the models performance, depending on the weather type observed when the forecasts were issued, for GHI and DNI, respectively.

As may be expected, the performance of the different models has a

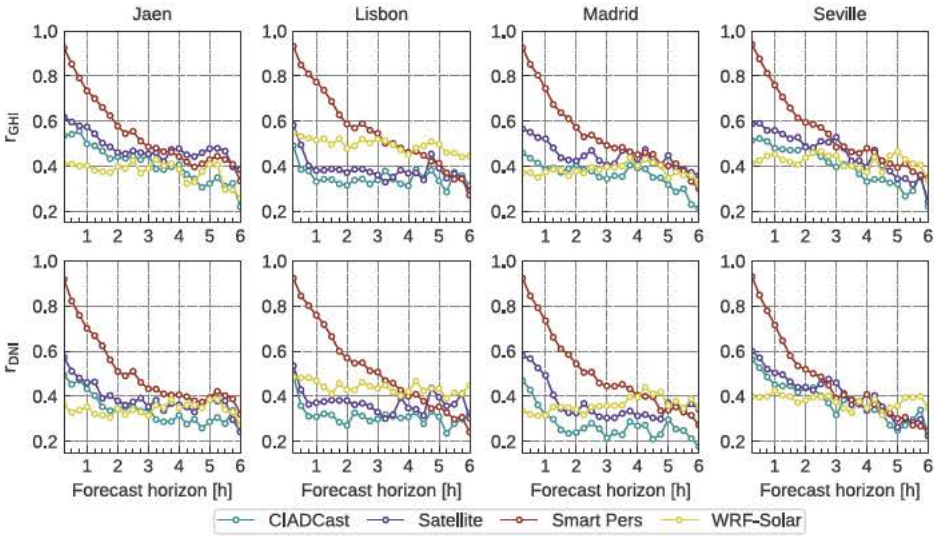

Fig. 7. Correlation coefficient, *r*, between the forecasted and observed solar radiation nominal variability index, as function of the forecasting horizon. Color code, rows, and columns structure as in Fig. 3.

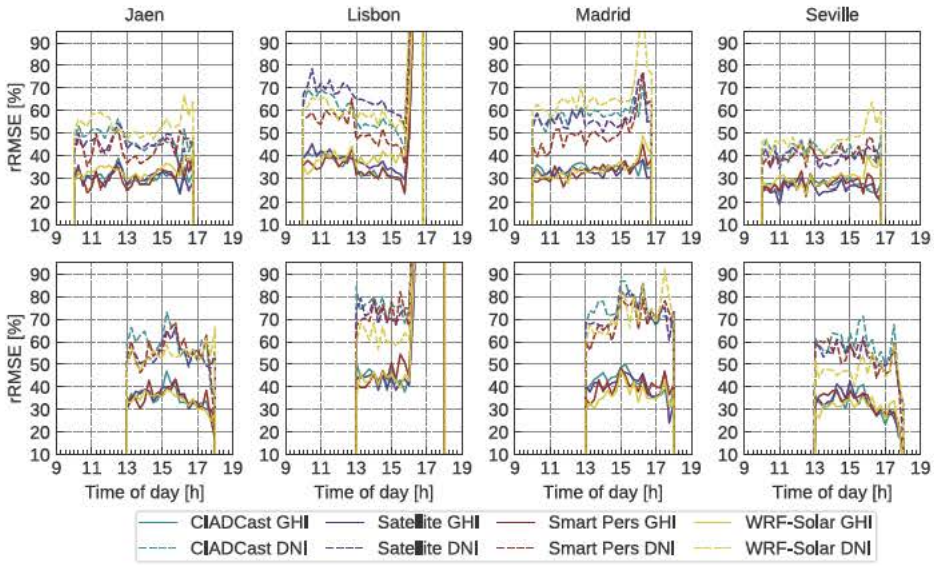

Fig. 8. Upper row, rRMSE values (GHI/DNI solid/dashed lines) for 1 h ahead forecasts as function of the time of the day. Lower row, as upper row but for 4 h ahead forecasts. Color code as in Fig. 3.

marked dependence on the WT for both GHI and DNI. Overall, WT 1 (low pressure system) shows the highest errors for all stations, followed by WT 2 (transition) and WT 3 (relatively high pressures/local conditions). On the other hand, WT 4 (high pressures system) shows the lowest errors. In addition, for the same WT and station, DNI rRMSE error values tend to be, approximately, double the corresponding GHI values.

For WT 1 and GHI (Fig. 8 first row), the Satellite model presents the lowest rRMSE values for almost all the forecasting horizons and for all stations except Jaen. Performance is particularly outstanding at Seville and Lisbon stations, which are located most westerly. At the Lisbon station, and only at the very end of the forecasting period, WRF-Solar shows a superior performance. The performance of the different models for the DNI forecasts for this WT (Fig. 9, first row) are qualitatively similar than that for the GHI case, although forecasting errors are considerably higher (about 2–3 times for the case of the rRMSE). The Satellite model performs the best (or is competitive) for all the forecasting horizons and stations, except Lisbon. At this station, and as in the case of GHI, at the last forecasting lead times, WRF-Solar provides lower errors. In general, and for both GHI and DNI, for this WT, WRF-Solar performance is particularly poor at the beginning of the forecasting period. This feature is particularly outstanding for DNI. Except in Lisbon, forecasting errors show little dependence on the forecasting horizon. It makes sense, since low pressure systems tend to be associated with overcast conditions covering thousands of square kilometers and lasting for more than 6 h (the forecasting horizon).

Overall, results show that Satellite tends to outperform the other models under these overcast conditions. This model performance is specially outstanding at stations located at the west of the Iberian Peninsula (Seville and Lisbon). This can be explained based on the fact that low pressure systems tend to enter into the Iberian Peninsula from the west and the Satellite model seems to be able to track these systems better than the other models. Nevertheless, forecasting errors are relatively high, particularly for the DNI.

Forecasts associated with WT 2 show the second highest forecasting errors (Figs. 10 and 11, second row). This WT is associated with changing weather conditions and intermediate cloud covers. The performance of the models varies greatly among stations. For the case of rRMSE and GHI, Satellite model performs the best at Lisbon and Seville stations in the first 3 h, after which the WRF-Solar is shown to be superior. On the other hand, at Madrid station, WRF-Solar is the best performing model from 1 to 6 h for both rRMSE and rMAE metrics. Regarding DNI and rRMSE, the performance of the models at Lisbon and Seville stations is similar in the case of GHI. The main difference is that the break-even point between the Satellite and the WRF-Solar model is now located at just 2 h. At Madrid station, all the models perform similarly except WRF-Solar for the first 2 h, which has the highest errors. At Jaen, WRF-Solar performs the best from 3 h onward.

The third row at Figs. 10 and 11 shows the results for GHI and DNI, respectively, for WT 3 (relative high pressures). Forecast error values

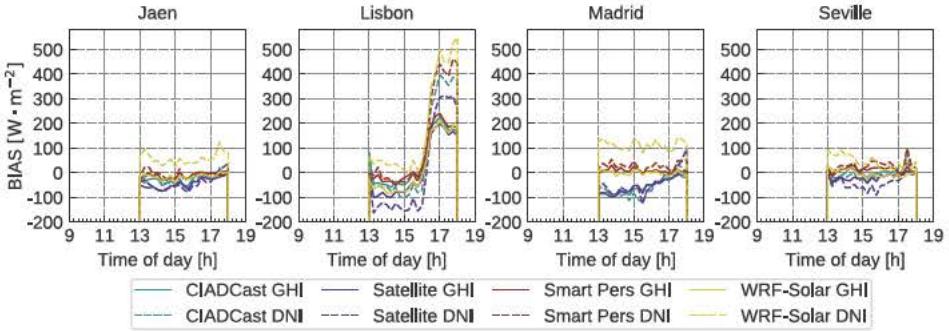

Fig. 9. BIAS values (GHI/DNI solid/dashed lines) for 4 h ahead forecasts as a function of the time of the day. Color code as in Fig. 3.

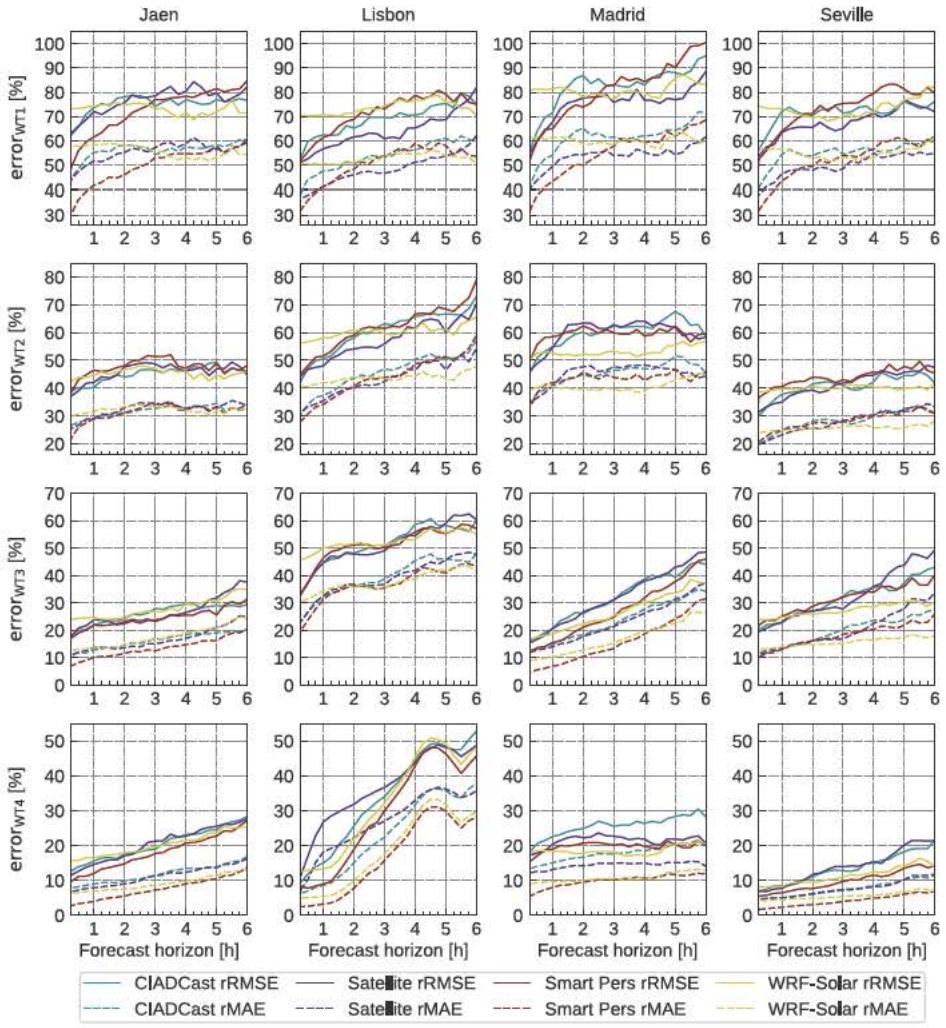

Fig. 10. rRMSE (solid lines) and rMAE (dashed lines) depending on the forecasting horizon, for predictions of GHI classified by the weather type observed at the time when the forecasts were issued. Each row represents a weather type. Results for Jaen, Lisbon, Madrid and Seville are displayed at the first, second, third and fourth column, respectively. Colors code as in Fig. 3. The vertical range of values (%) varies with the WTs, but is kept the same at the four stations, for the sak.e of comparison.

are the second lowest, but show a notable upward trend with the forecasting horizon for all stations and for both GHI and DNI. This indicates the difficulties to predict local cloudiness at large forecasting horizons, often originated by convection. The performance of the different models varies among the stations. In the case of rRMSE and GHI (Fig. 10, third row), at Jaen station, Smart Persistence performs the best for all the lead times. At Seville WRF-Solar provides lower errors from 2 h onward (4 h for Madrid). Lisbon station shows the highest errors and the performance of all the models is similar. None of the models show special skills at this station under these weather conditions. Regarding the DNI (Fig. 11, third row), the performance of the different models at the different stations is qualitatively similar in the case of GHI. The main differences are found at Lisbon station, where WRF-Solar outperforms the other models after 4 h. Interestingly, for this WT, Smart Persistence performs the best regarding the rMAE for all stations up to, approximately, the first four leading hours.

Finally, the last row in Figs. 10 and 11 show the results for WT 4 (high pressure conditions). As may be expected, errors associated with this WT are the lowest ones. For instance, at Seville, rRMSE values are below 25% for GHI and 40% for DNI. The Lisbon station shows the largest errors and a remarkable increment with lead time. The special conditions at the location of this station, once again, explain these results. Given the availability of moisture, high pressure systems enhance the development of convective clouds under clear sky conditions. For GHI and rRMSE, (Fig. 10 last row), Smart Persistence is the best model for all stations and forecasting horizons, except at Madrid, where WRF-Solar outperforms the rest from 1 h ahead horizons. At the other stations, WRF-Solar proves to be competitive at the end of the forecasting period. Results for the DNI (Fig. 11, last row) are qualitatively similar. Smart Persistence is the best performing model for all the lead times and stations except at Lisbon. At this station, Satellite model provides lower rRMSE errors after 4 h. A special feature observed for this WT is the poor performance of WRF-Solar for DNI, as opposed to GHI, where this model was competitive. This suggests a misrepresentation of the aerosol load in the WRF-Solar model, that in this study is set climatologically. Finally, a distinctive feature of this WT is that Smart Persistence is the best model for DNI and rMAE for all stations and leading times. For GHI, WRF-Solar proves to be competitive at the end of the

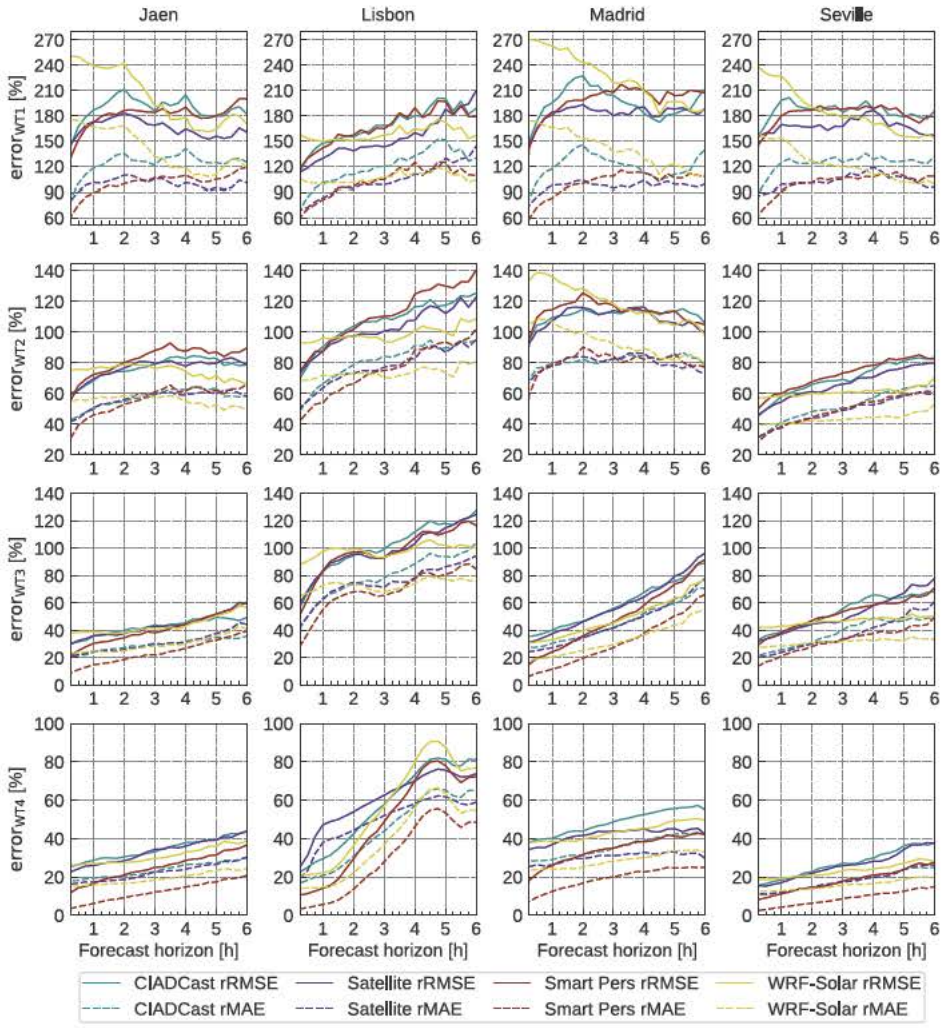

Fig. 11. As in Fig. 10 but for the DNI.

forecasting horizon. Again, this seems to be related with a misrepresentation of the aerosol in the WRF-Solar model, but also in the Satellite and CIADCast models. Generally, the use of enhanced sources of aerosol data would improve the results of forecasting models.

In addition to the former analysis based on weather types, a study on the models dependence on the sky conditions (using the clearness index, k_t) is conducted Appendix B. General results show that Smart Persistence is difficult to beat in clear ($k_t > 0.65$) and cloudy ($k_t < 0.4$) skies, but performs the worst in days of high intermittency. In the latter, other models show a superior performance for lead times higher than around 1 h. Under clear sky conditions, WRF-Solar and Smart Persistence are the best performing models, while CIADCast and Satellite models show considerable higher erros. In case of overcast conditions, overall errors increase considerably, especially for DNI.

6. Summary and conclusions

In this work four short-term solar radiation forecasting models are evaluated in the centre and south of the Iberian Peninsula, a region particularly suitable for the development of solar energy. The four models follow (mostly) independent approaches: one pure statistical model (Smart Persistence), one model based on CMV derived from satellite images (Satellite), one NWP model (WRF-Solar), and a hybrid satellite-NWP model (CIADCast). GHI and DNI forecasts are obtained for the four models, for two years of up to 6 h ahead and 15 min time step. The models' performance is assessed using radiometric data collected at four stations, representative of different topographic and geographic conditions within the study area. Performance of the models is analyzed in the light of these features. The models' ability to forecast solar radiation variability and the dependence of the models performance on the synoptic weather conditions over the study area are also analyzed. For this latter analysis, four different synoptic patterns are considered.

Overall, the four models' performance shows a marked dependence on the evaluation stations and strong differences are found between GHI and DNI. For the same stations and forecasting horizons, DNI rRMSE values are between 30% and 90% higher than the corresponding GHI counterparts. The highest errors are observed at the coastal station of Lisbon (GHI rRMSE ranging from 30% to 70% and DNI 40% to 130%) and the lowest errors are observed at the Seville station (GHI rRMSE 25% to 40% and DNI 35% to 70%), which is located at the entrance of a valley open to the Atlantic Ocean.

Notable differences among models' performance are observed. In general, and as may be expected, Smart Persistence is the best

performing model at the first lead times, then advective models (Satellite and CIADCast) show a good performance and the WRF-Solar tends to provide the most accurate forecast at the end of the forecasting period. Nevertheless, notable differences are observed depending on the evaluation station and variable (GHI vs. DNI). Advective models are the most accurate models at the stations located most westerly (Lisbon and Seville), while they show poor results for stations located at the centre and east of the study region. At these stations WRF-Solar shows a superior performance. The break-even point between Satellite and WRF-Solar models varies between stations, but it is located between 1 and 3 h for GHI and 3 and 5 h for the DNI. Finally, Smart Persistence is shown to be the best performing model for DNI up to, approximately, the third lead hour for the rRMSE and for almost the entire forecasting window for the rMAE. Smart Persistence also is shown to be the best performing model regarding the forecasts of GHI and DNI variability.

The results of the Lisbon station show some different features. First of all, all the models show poor performances and a marked increment of the errors with the forecasting horizon. Secondly, all the models show a marked dependence of the forecast errors on the time of day. Notably, a dramatically error increment during evening hours (approximately from 16:00 onward) is observed. This dependence can be attributed to a daily cloudiness cycle, associated with the development of sea-land breezes. All models show a lack of ability to forecast convective cloudy conditions, even at 1 h forecasting horizon. This result may be anticipated for all the models, except WRF-Solar that also shows limited skills in this regard. Results concerning the WRF-Solar model may be influenced by the nesting configuration and shallow cumulus parameterization used in this study. Notably, the Deng et al. (2014) parameterization has not yet tested at the here used 5 km spatial resolution. This may end in spurious cloudiness forecasting that would have implications on the solar radiation forecasts reliability.

The seasonal analysis reveals significant differences between the models performance among seasons. Notably, the WRF-Solar model shows a notable skill for winter and spring for almost all the stations and both for the GHI and DNI. On the other hand, for the other seasons, differences among models performance are, in general, lower.

Lastly, models' performance showed to be highly dependent on the synoptic weather conditions. For low pressure conditions over the study area (overcast conditions), rRMSE values of the Satellite model are shown to be the lowest for all the forecasting window. Nevertheless, error values for these synoptic conditions are shown to be the highest ones (about 80% for GHI and 180% for DNI). For transient weather conditions (intermediate cloudiness) results are more variable, but Satellite tends to be the best performing model up to 3 h ahead, after which WRF-Solar provides the most accurate forecasts. For the two former synoptic conditions, little dependence of the errors on the forecasting horizon is observed. For moderate high pressure conditions, the third analyzed synoptic pattern, overall forecasting errors are small, but show a notable increase with the forecasting horizon. This pattern is associated with mostly clear sky conditions but also with the development of local cloudiness, associated with local convection and mountain breezes. The increase of the error with horizon indicates the difficulties to forecast this cloudiness. For this weather pattern, all models tend to provide a similar performance, except at the end of the lead time when the WRF-Solar is shown to be superior. Finally, the last synoptic pattern analyzed is the presence of high pressures over the study area. This pattern is associated with clear skies and models show the lowest forecasting errors (for instance, at Seville, rRMSE is below 25% for GHI and below 40% for DNI) except at the Lisbon station. At this station, high errors are observed and also an increase of these errors with the forecasting horizon. The availability of moisture, along with the frequent occurrence of sea-land breezes at this station, may explain this result.

Results of the models dependence on the daily clearness index (k_t) show that, in partly cloudy conditions, Smart Persistence is overcome by either WRF-Solar or advective models in all stations, for rMAE and rRMSE and for GHI and DNI. Therefore, for intermittent sky conditions, models have demonstrated their usefulness in solar radiation forecasting.

To conclude, even though the four evaluation stations belong to the same climatic region, the performance of the four models shows enormous differences. These differences are associated with: 1) the local geographic and topographic conditions of the evaluation stations; 2) the evaluated variable (GHI vs. DNI) and 3) the sky and synoptic weather conditions over the study area. No single model proves to be always the best performing model and, therefore, results show that the four models evaluated are, somehow, complementary. Unlike some other previous studies, results here presented show the Smart Persistence model to have a remarkable performance, particularly for low lead-times. This result, however is constrained by the limited number of stations evaluated, and may change for a regional analysis (which itself is limited due to data availability).

Based on this fact, in the companion paper the benefits obtained by developing optimal blending of the four models are explored.

Declaration of Competing Interest

The authors declared that there is no conflict of interest.

Acknowledgements

The authors are supported by the Spanish Ministry of Economy and Competitiveness, project ENE2014-56126-C2-1-R and ENE2014-56126-C2-2-R (http://prosol.uc3m.es). The team from the University of Jaen is also supported by FEDER funds and by the Junta de Andalucía (Research group TEP-220). The authors thank all the provided support. The authors are in debt with the National Centers for Environmental Prediction (NCEP), EUMETSAT, Faculdade de Ciencias da Universidade de Lisboa, Grupo de Energía Solar of the Universidad Politécnica de Madrid and Abengoa Solar for providing the data used in this work. Finally, the authors also thank the reviewers for improving the article with their comments.

Appendix A. Seasonal analyses

Figs. A.12 and A.13 show the rRMSE and rMAE values for each season of the year. Seasons are composed by three months, being winter from December to February, spring from March to May, summer from June to August and autumn from September to November.

Overall, summer shows the lowest errors, as well as the lowest differences between models performance. Spring and autumn show intermediate errors, while the highest errors are observed during winter (up to 70% for GHI rRMSE and 120% for DNI rRMSE, almost double than those of summer). Winter, in addition, shows the greatest differences among models performance. The exception is found for the Lisbon station, where summer errors are similar to winter ones for the GHI and higher for DNI. The seasonal differences of the models performance can be explained based

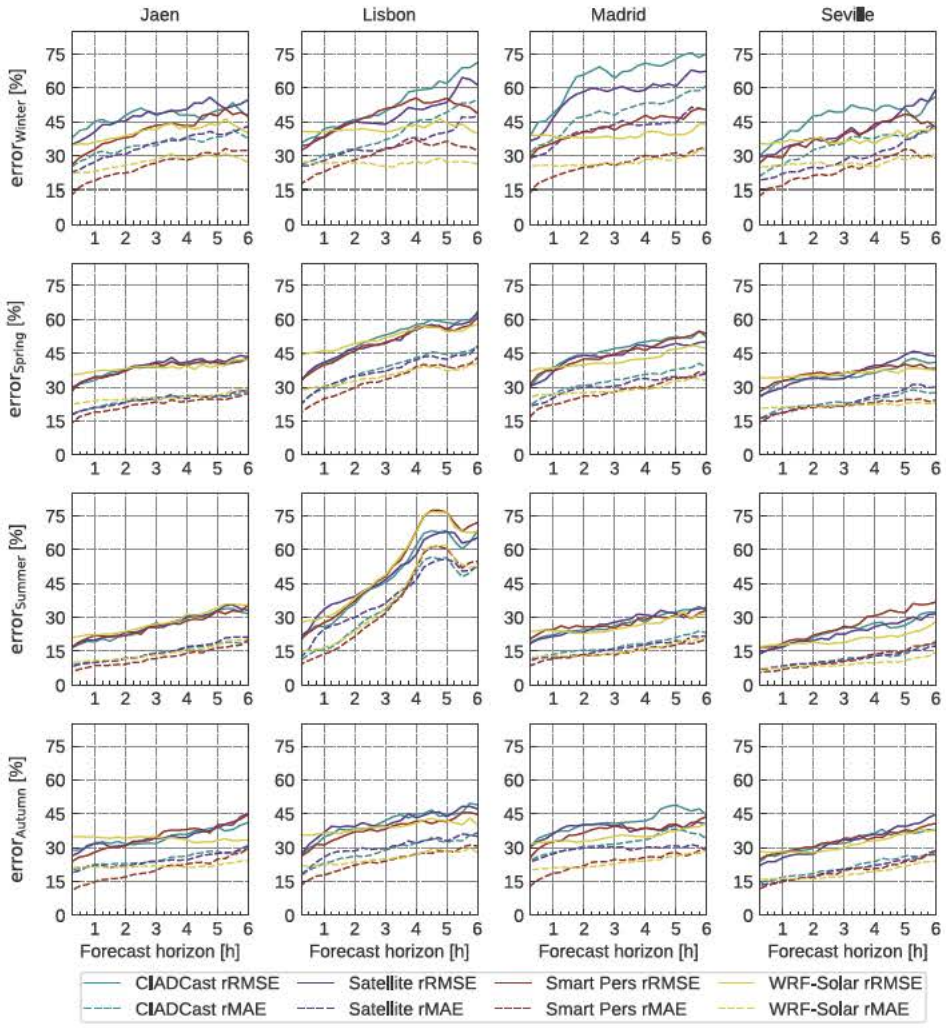

Fig. A.12. rRMSE (solid lines) and rMAE (dashed lines) depending on the forecasting horizon, for predictions of GHI classified according to the season. Results for Jaen, Lisbon, Madrid and Seville are displayed at the first, second, third, and fourth column, respectively. Colors code as in Fig. 3.

on the climatic characteristics of the study area. In particular, the higher errors during winter compared to summer (except at the Lisbon station) are related with the seasonal cycle of cloudiness and precipitation in the study area.

Regarding the relative performance of the models, for the GHI (Fig. A.12) and during winter, WRF-Solar shows a notable performance. Notably, WRF-Solar presents the lowest rRMSE values at lead time higher than 1.5 h in Lisbon and Madrid and higher than 3 h at the other two stations. At lower lead times, Smart Persistence is the best performing model. For the other seasons, differences among models performance are, in general, low. The most notable feature is observed at Madrid station during spring and autumn. In these cases, WRF-Solar clearly shows the lowest rRMSE values at lead times higher than 1 h. Also at Seville station and during summer, WRF-Solar shows a superior performance. On the other hand, performance of the WRF-Solar at Lisbon during summer is poor, probably due to the lack of ability to account for the land-see breeze circulation. The use of an alternative cumulus parameterization could have depict this behavior. The relative performance of the models for the DNI (Fig. A.13) is qualitatively similar than in the GHI counterpart. Main differences are observed at Madrid station during winter and autumn, where Smart Persistence clearly outperforms the rest of the models.

Fig. A.14 shows the seasonal Forecast Skill (FS) score, i.e., the relative performance of the models with respect to Smart Persistence model in terms of RMSE, computed by season. Overall, the lowest skills are observed during summer, followed by autumn. But notable differences are oberved among evaluation stations. Notably, the Jaen station shows the lowest skill for all the seasons. During winter, WRF-Solar shows a notable performance for all the stations and for the GHI, specially for lead times greater than 3 h. During summer, CIADCast and Satellite models show significant FS scores at the Lisbon station and the WRF-Solar at the Seville station. For DNI, the WRF-Solar shows a notable performance during spring.

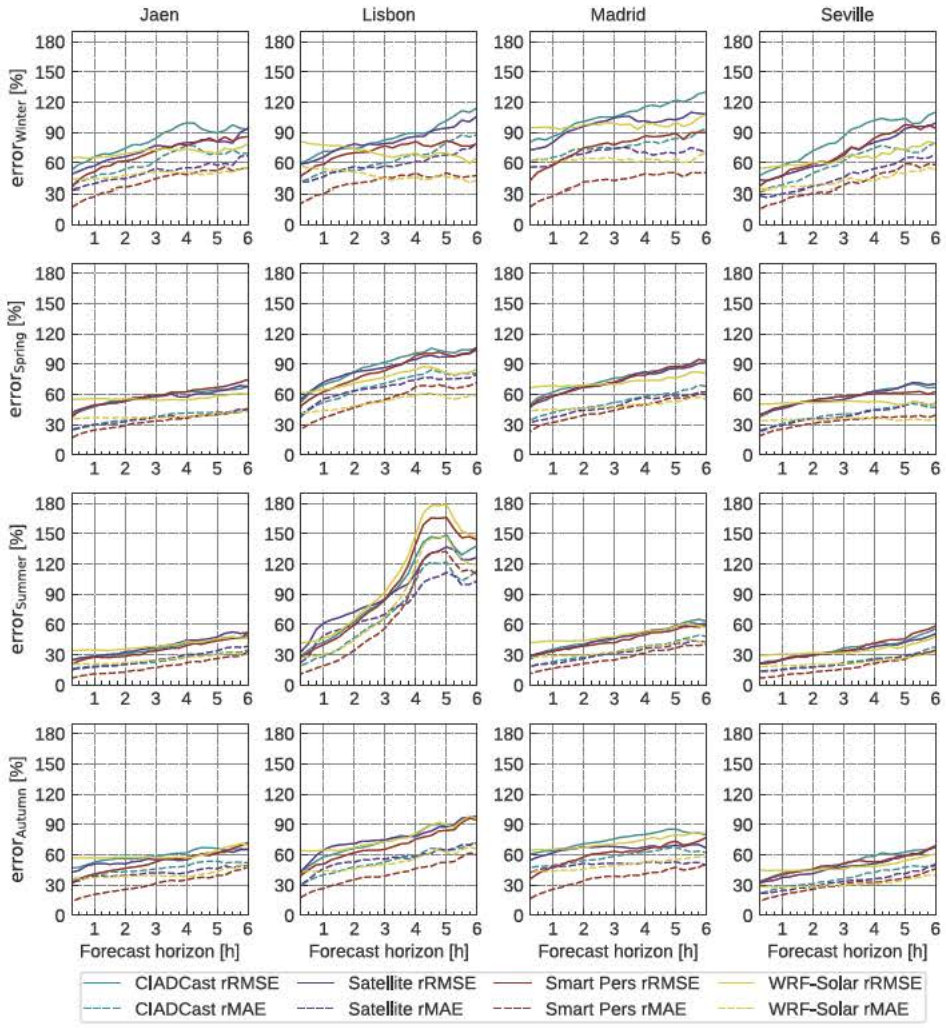

Fig. A.13. As in Fig. A.12 but for DNI.

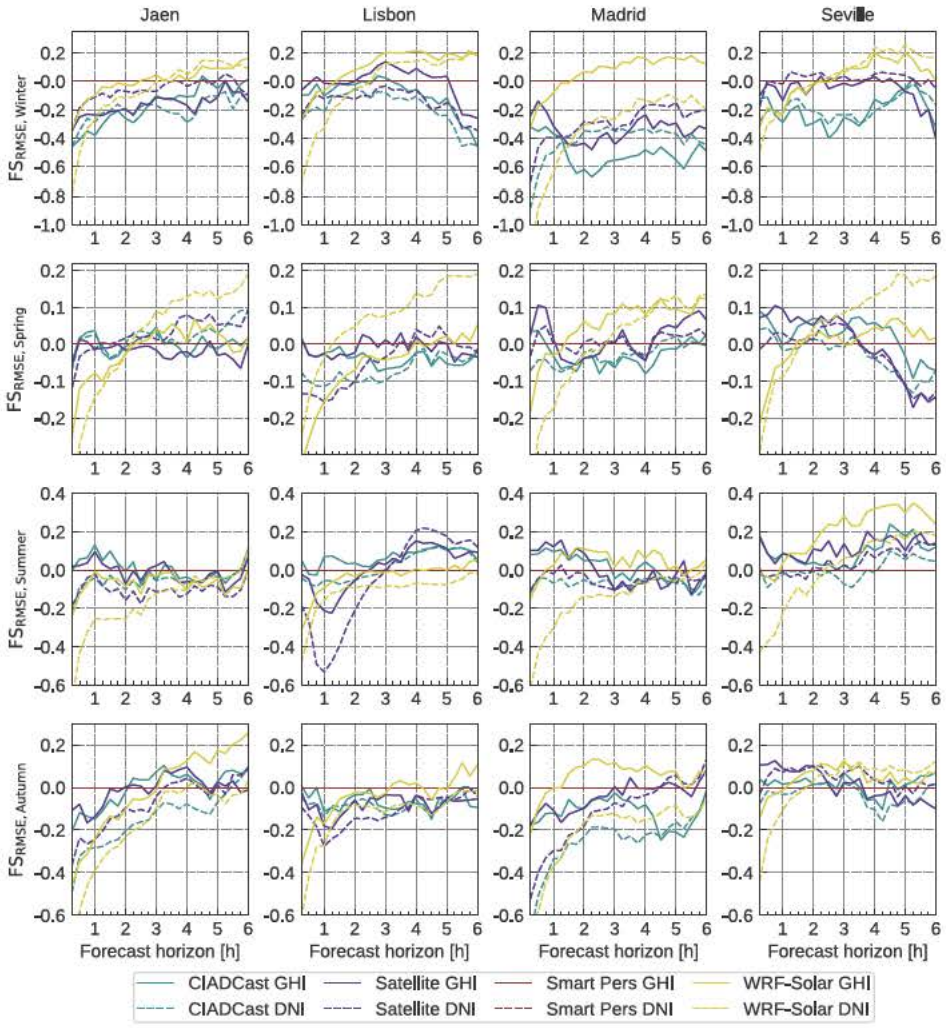

Fig. A.14. Forecast skills (FS) of the GHI (solid lines) and DNI (dashed lines), depending on the forecasting horizon, classified according to the season. The color code is the same used in Fig. 3.

Appendix B. Daily k_t distribution analysis

Figs. B.15 and B.16 show the models performance, according to the daily mean clearness index (k_t) observed when the forecasts were issued, for GHI and DNI, respectively. Following Lara-Fanego et al. (2012), days with mean (k_t) greater than 0.65 are considered as clear sky conditions, between 0.4 and 0.65 partly cloudy, and below 0.4 as overcast conditions.

For clear-sky conditions and GHI (Fig. B.15 first row), errors reach around 20% for rMAE and 30% for rRMSE, except in Lisbon where errors of CIADCast and Satellite models reach 50% of rRMSE at 6 h-ahead. These models might wrongly advect satellite-derived cloudiness, commonly present at coastline, into the station, resulting in higher errors, in agreement with their negative bias (not shown). This performance is observed to a less extent at the rest of the stations. Values for DNI (Fig. B.16 first row) are qualitatively similar but 50% higher than those of the GHI counterpart. In general, for clear conditions, Smart Persistence is the best performing model.

The performance of the models in partly cloudy (Figs. B.15 and B.16, second row) and clear sky conditions considerably differs. During partly cloudy conditions, Smart Persistence is the worst performing model, as may be expected. For the GHI, rMAE values range from 25% to 55% and rRMSE from 35% to 75% (Fig. B.15). For the DNI case (Fig. B.16) rMAE values range from 40% to 140% and rRMSE from 60% to 170%. For these partly cloudy conditions, the WRF-Solar shows to be the best performing model in Madrid and Jaen for the GHI (Fig. B.15). At the Seville station for both GHI and DNI, and at the Jaen station just for the DNI, CIADCast and Satellite models show lower errors for lead times lower than 3 h, when WRF-Solar performs similarly.

For overcast conditions (Figs. B.15 and B.16 third row), as in clear sky conditions, Smart Persistence shows a superior performance. Only in Lisbon station and for the GHI, the WRF-Solar provides better forecasts at the end of the forecasting horizon (Fig. B.15). For the GHI, rMAE values in the range 50% to 90% and rRMSE values between 80% and 120% are observed (Fig. B.15). For DNI (Fig. B.16), error values are considerable higher.

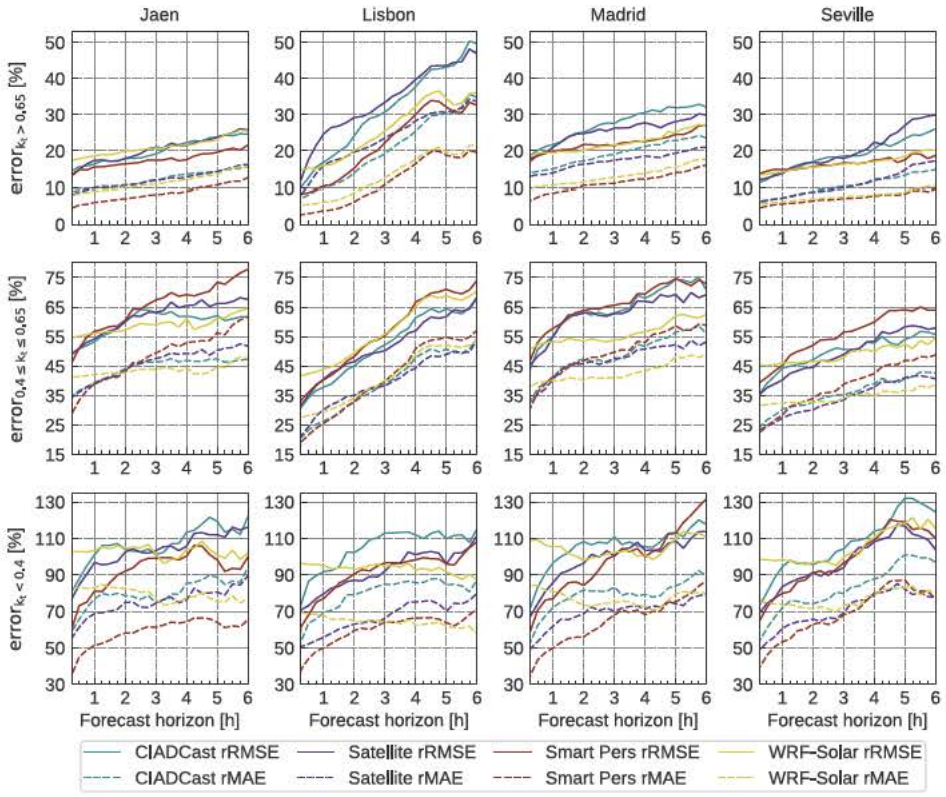

Fig. B.15. rRMSE (solid lines) and rMAE (dashed lines) depending on the forecasting horizon, for predictions of GHI classified according to the k_t distribution observed at the day when the forecasts were issued. First row represents $k_t > 0.65$ (clear sky), second row shows $0.4 \leqslant k_t \leqslant 0.65$ (partly cloudy), and third row represents $k_t < 0.4$ (overcast conditions). Results for Jaen, Lisbon, Madrid and Seville are displayed at the first, second, third, and fourth column, respectively. Colors code as in Fig. 3. The vertical range of values (%) varies with the k_t distribution, but is kept the same at the four stations, for the sake of comparison.

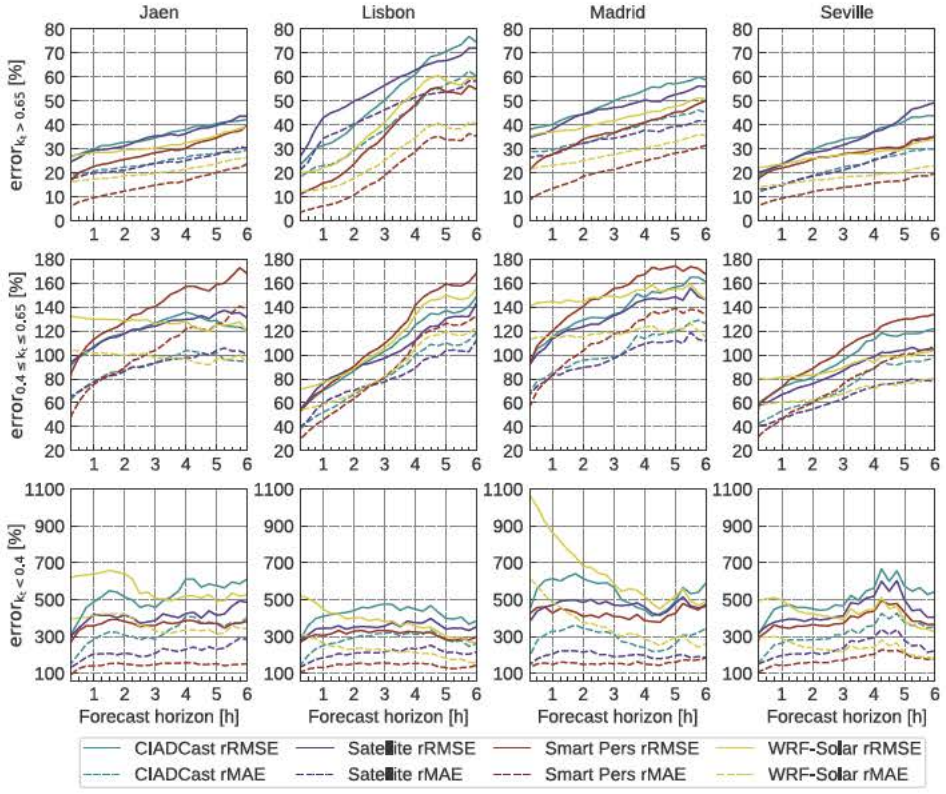

Fig. B.16. As in Fig. B.15 but for the DNI.

References

Antonanzas, J., Pozo-Vázquez, D., Fernandez-Jimenez, L., de Pison, F.M., 2017. The value of day-ahead forecasting for photovoltaics in the Spanish electricity market. Sol. Energy 158, 140–146. https://doi.org/10.1016/j.solener.2017.09.043. <http://www.sciencedirect.com/science/article/pii/S0038092X17308307>.

Arbizu-Barrena, C., Pozo-Vázquez, D., Ruiz-Arias, J.A., Tovar-Pescador, J., 2015. Macroscopic cloud properties in the WRF NWP model: An assessment using sky camera and ceilometer data. J. Geophys. Res.: Atmospheres 120, 10297–10312. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023502, https://doi.org/10.1002/2015JD023502, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015JD023502.

Arbizu-Barrena, C., Ruiz-Arias, J.A., Rodríguez-Benítez, F.J., Pozo-Vázquez, D., Tovar-Pescador, J., 2017. Short-term solar radiation forecasting by advecting and diffusing MSG cloud index. Sol. Energy 155, 1092–1103. https://doi.org/10.1016/j.solener.2017.07.045. <http://www.sciencedirect.com/science/article/pii/S0038092X17306308>.

Avolio, E., Federico, S., Miglietta, M., Feudo, T.L., Calidonna, C., Sempreviva, A., 2017. Sensitivity analysis of WRF model PBL schemes in simulating boundary-layer variables in southern Italy: An experimental campaign. Atmos. Res. 192, 58–71. https://doi.org/10.1016/j.atmosres.2017.04.003. <http://www.sciencedirect.com/science/article/pii/S0169809516307189>.

Beyer, H., Costanzo, C., Heinemann, D., Reise, C., 1994. Short range forecast of PV energy production using satellite image analysis. In: Proc. 12th European Photovoltaic Solar Energy Conference, Amsterdam. p. 15.

Boland, J., Grantham, A., 2018. Nonparametric conditional heteroscedastic hourly probabilistic forecasting of solar radiation. J - Multidiscip. Sci. J. 1, 174–191. https://doi.org/10.3390/j1010016. <https://www.mdpi.com/2571-8800/1/1/16>.

Brancucci Martinez-Anido, C., Botor, B., Florita, A.R., Draxl, C., Lu, S., Hamann, H.F., Hodge, B.M., 2016. The value of day-ahead solar power forecasting improvement. Sol. Energy 129, 192–203. https://doi.org/10.1016/j.solener.2016.01.049. <http://www.sciencedirect.com/science/article/pii/S0038092X16000736>.

Brouwer, A.S., van den Broek, M., Seebregts, A., Faaij, A., 2014. Impacts of large-scale Intermittent Renewable Energy Sources on electricity systems, and how these can be modeled. Renew. Sustain. Energy Rev. 33, 443–466. https://doi.org/10.1016/j.rser.2014.01.076. <http://www.sciencedirect.com/science/article/pii/S1364032114000987>.

Coimbra, C.F., Kleissl, J., Marquez, R., 2013. Chapter 8 - Overview of solar-forecasting methods and a metric for accuracy evaluation. In: Kleissl, J. (Ed.), Solar Energy Forecasting and Resource Assessment. Academic Press, Boston, pp. 171–194. http://www.sciencedirect.com/science/article/pii/B9780123971777000085, https://doi.org/10.1016/B978-0-12-397177-7.00008-5.

Deng, A., Gaudet, B., Dudhia, J., Alapaty, K., 2014. Implementation and evaluation of a new shallow convection scheme in WRF. In: 26th Conf. on Weather Analysis and Forecasting/22nd Conf. on Numerical Weather Prediction. https://ams.confex.com/ams/94Annual/webprogram/Paper236925.html.

Dersch, J., Schroedter-Homscheidt, M., Gairaa, K., Hanrieder, N., Landelius, T., Lindskog, M., Müller, S., Ramirez Santigosa, L., Sirch, T., Wilbert, S., 2019. Impact of DNI nowcasting on annual revenues of CSP plants for a time of delivery based feed in tariff. Meteorol. Z. https://doi.org/10.1127/metz/2019/0925.

Descombes, G., Auligne, T.D., Lin, H.C., Xu, D., Schwartz, C.S., Vandenberghe, F., 2014. Multi-sensor Advection Diffusion nowCast (MADCast) for cloud analysis and short-term prediction. Technical Report. NCAR Technical Note NCAR/TN-509+STR. URL http://n2t.net/ark:/85065/d7gq6x6v, https://doi.org/10.5065/D62V2D37.

Diagne, M., David, M., Lauret, P., Boland, J., Schmutz, N., 2013. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renew. Sustain. Energy Rev. 27, 65–76. https://doi.org/10.1016/j.rser.2013.06.042. <http://www.sciencedirect.com/science/article/pii/S1364032113004334>.

Ela, E., Milligan, M., Kirby, B., 2011. Operating Reserves and Variable Generation. https://doi.org/10.2172/1023095.

Guillot, E.M., Vonder Haar, T.H., Forsythe, J.M., Fletcher, S.J., 2012. Evaluating Satellite-based cloud persistence and displacement nowcasting techniques over complex terrain. Weather Forecast. 27, 502–514. https://doi.org/10.1175/WAF-D-11-00037.1, arXiv:https://doi.org/10.1175/WAF-D-11-00037.1.

Hamann, U., Walther, A., Baum, B., Bennartz, R., Bugliaro, L., Derrien, M., Francis, P.N., Heidinger, A., Joro, S., Kniffka, A., Le Gléau, H., Lockhoff, M., Lutz, H.J., Meirink, J.F., Minnis, P., Palikonda, R., Roebeling, R., Thoss, A., Platnick, S., Watts, P., Wind, G., 2014. Remote sensing of cloud top pressure/height from SEVIRI: analysis of ten current retrieval algorithms. Atmospheric. Meas. Tech. 7, 2839–2867. https://doi.org/10.5194/amt-7-2839-2014. <https://www.atmos-meas-tech.net/7/2839/2014/>.

Haupt, S.E., 2018. Short-Range Forecasting for Energy. Springer International Publishing, Cham, pp. 97–107. https://doi.org/10.1007/978-3-319-68418-5_7.

Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A., Collins, W.D., 2008. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. J. Geophys. Res.: Atmospheres 113. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD009944, https://doi.org/10.1029/2008JD009944, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2008JD009944.

Inman, R.H., Pedro, H.T., Coimbra, C.F., 2013. Solar forecasting methods for renewable energy integration. Prog. Energy Combust. Sci. 39, 535–576. https://doi.org/10.1016/j.pecs.2013.06.002. <http://www.sciencedirect.com/science/article/pii/S0360128513000294>.

International Energy Agency, 2018. Market Report Series: Renewables 2018. Analysis and Forecasts to 2023. https://webstore.iea.org/market-report-series-renewables-2018.

Jimenez, P.A., Hacker, J.P., Dudhia, J., Haupt, S.E., Ruiz-Arias, J.A., Gueymard, C.A., Thompson, G., Eidhammer, T., Deng, A., 2016. WRF-solar: description and clear-sky assessment of an augmented NWP model for solar power prediction. Bull. Am. Meteorol. Soc. 97, 1249–1264. https://doi.org/10.1175/BAMS-D-14-00279.1.

Kühnert, J., Lorenz, E., Heinemann, D., 2013. Chapter 11 - Satellite-based irradiance and power forecasting for the german energy market. In: Kleissl, J. (Ed.), Solar Energy Forecasting and Resource Assessment. Academic Press, Boston, pp. 267–297. https://doi.org/10.1016/B978-0-12-397177-7.00011-5. <http://www.sciencedirect.com/science/article/pii/B9780123971777000115>.

Lara-Fanego, V., Ruiz-Arias, J., Pozo-Vázquez, D., Santos-Alamillos, F., Tovar-Pescador, J., 2012. Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain). Sol. Energy 86, 2200–2217. http://www.sciencedirect.com/science/article/pii/S0038092X11000582, https://doi.org/10.1016/j.solener.2011.02.014. progress in Solar Energy 3.

Lee, J.A., Haupt, S.E., Jiménez, P.A., Rogers, M.A., Miller, S.D., McCandless, T.C., 2017. Solar Irradiance nowcasting case studies near sacramento. J. Appl. Meteorol. Climatol. 56, 85–108. https://doi.org/10.1175/JAMC-D-16-0183.1.

Lew, D., Brinkman, G., Ibanez, E., Hummon, M., Hodge, B.M., Heaney, M., King, J., 2012. Sub-hourly impacts of high solar penetrations in the Western United States: Preprint, United States. URL https://www.osti.gov/servlets/purl/1051154.

Lorenz, E., Heinemann, D., 2012. 1.13 - Prediction of solar irradiance and photovoltaic power. In: Sayigh, A. (Ed.), Comprehensive Renewable Energy. Elsevier, Oxford, pp. 239–292. http://www.sciencedirect.com/science/article/pii/B9780080878720001141, https://doi.org/10.1016/B978-0-08-087872-0.00114-1.

Lorenz, E., Kühnert, J., Heinemann, D., 2012. Short term forecasting of solar irradiance by combining satellite data and numerical weather predictions. In: Proceedings of the 27th European PV Solar Energy Conference (EU PVSEC), Frankfurt, Germany, pp. 44014405.

Lorenz, E., Kühnert, J., Heinemann, D., Nielsen, K.P., Remund, J., Müller, S.C., 2016. Comparison of global horizontal irradiance forecasts based on numerical weather prediction models with different spatio-temporal resolutions. Progress Photovolt.: Res. Appl. 24, 1626–1640. https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.2799, https://doi.org/10.1002/pip.2799, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.2799.

Mathiesen, P., Kleissl, J., 2011. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States. Sol. Energy 85, 967–977. https://doi.org/10.1016/j.solener.2011.02.013. <http://www.sciencedirect.com/science/article/pii/S0038092X11000570>.

McCandless, T., Haupt, S., Young, G., 2016. A regime-dependent artificial neural network technique for short-range solar irradiance forecasting. Renew. Energy 89, 351–359. https://doi.org/10.1016/j.renene.2015.12.030. <http://www.sciencedirect.com/science/article/pii/S0960148115305346>.

Miller, S.D., Rogers, M.A., Haynes, J.M., Sengupta, M., Heidinger, A.K., 2018. Short-term solar irradiance forecasting via satellite/model coupling. Sol. Energy 168, 102–117. http://www.sciencedirect.com/science/article/pii/S0038092X17310435, https://doi.org/10.1016/j.solener.2017.11.049. advances in Solar Resource Assessment and Forecasting.

Mori, N., Chang, K.A., 2003. Introduction to MPIV. user reference manual 14.

National Centers for Environmental Prediction, NCEP, 2006. [dataset] global forecast system (gfs) [0.5 deg.]. NCEI DSI 6182; gov.noaa.ncdc:C00634.

Nonnenmacher, L., Coimbra, C.F., 2014. Streamline-based method for intra-day solar forecasting through remote sensing. Sol. Energy 108, 447–459. https://doi.org/10.1016/j.solener.2014.07.026. <http://www.sciencedirect.com/science/article/pii/S0038092X14003752>.

Perez, R., David, M., Hoff, T.E., Jamaly, M., Kivalov, S., Kleissl, J., Lauret, P., Perez, M., 2016. Spatial and Temporal Variability of Solar Energy. Found. Trends Renew. Energy 1, 1–44. https://doi.org/10.1561/2700000006.

Perez, R., Hoff, T.E., 2013. Chapter 6 - solar resource variability. In: Kleissl, J. (Ed.), Solar Energy Forecasting and Resource Assessment. Academic Press, Boston, pp. 133–148. https://doi.org/10.1016/B978-0-12-397177-7.00006-1. <http://www.sciencedirect.com/science/article/pii/B9780123971777000061>.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., Vignola, F., 2002. A new operational model for satellite-derived irradiances: description and validation. Sol. Energy 73, 307–317. https://doi.org/10.1016/S0038-092X(02)00122-6. <http://www.sciencedirect.com/science/article/pii/S0038092X02001226>.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., Hoff, T.E., 2010. Validation of short and medium term operational solar radiation forecasts in the US. Sol. Energy 84, 2161–2172. https://doi.org/10.1016/j.solener.2010.08.014.

Red Eléctrica de España, 2017. The Spanish Electricity System. Preliminary Report 2017. Technical Report. Red Eléctrica de España. http://www.ree.es/sites/default/files/downloadable/avance_informe_sistema_electrico_2017_eng.pdf.

Redes Energéticas Nacionais, 2018. Technical Data 2018. techreport. Redes Energéticas Nacionais. https://www.ren.pt/en-GB/media/publications/.

Remund, J., Wald, L., Lefèvre, M., Ranchin, T., Page, J.H., 2003. Worldwide Linke turbidity information. In: ISES Solar World Congress 2003, International Solar Energy Society (ISES), Göteborg, Sweden. p. 13. https://hal.archives-ouvertes.fr/hal-00465791.

Renné, D.S., 2014. Emerging Meteorological Requirements to Support High Penetrations of Variable Renewable Energy Sources: Solar Energy. Springer, New York, New York, NY., pp. 257–273. https://doi.org/10.1007/978-1-4614-9221-4_12.

Rigollier, C., Bauer, O., Wald, L., 2000. On the clear sky model of the ESRA – European Solar Radiation Atlas – with respect to the heliosat method. Sol. Energy 68, 33–48. https://doi.org/10.1016/S0038-092X(99)00055-9. <http://www.sciencedirect.com/science/article/pii/S0038092X99000559>.

Rigollier, C., Lefèvre, M., Wald, L., 2004. The method Heliosat-2 for deriving shortwave solar radiation from satellite images. Sol. Energy 77, 159–169. https://doi.org/10.

1016/j.solener.2004.04.017. <http://www.sciencedirect.com/science/article/pii/S0038092X04001082>.

Rodríguez-Benítez, F.J., Arbizu-Barrena, C., Santos-Alamillos, F.J., Tovar-Pescador, J., Pozo-Vázquez, D., 2018. Analysis of the intra-day solar resource variability in the Iberian Peninsula. Sol. Energy 171, 374–387. https://doi.org/10.1016/j.solener.2018.06.060. <http://www.sciencedirect.com/science/article/pii/S0038092X18306108>.

Santos-Alamillos, F., Pozo-Vázquez, D., Ruiz-Arias, J., Lara-Fanego, V., Tovar-Pescador, J., 2014. A methodology for evaluating the spatial variability of wind energy resources: Application to assess the potential contribution of wind energy to baseload power. Renew. Energy 69, 147–156. https://doi.org/10.1016/j.renene.2014.03.006. <http://www.sciencedirect.com/science/article/pii/S0960148114001499>.

Santos-Alamillos, F.J., Brayshaw, D.J., Methven, J., Thomaidis, N.S., Ruiz-Arias, J.A., Pozo-Vázquez, D., 2017. Exploring the meteorological potential for planning a high performance European electricity super-grid: optimal power capacity distribution among countries. Environ. Res. Lett. 12, 114030. https://doi.org/10.1088/1748-9326/aa8f18. https://doi.org/10.1088/1748-9326/aa8f18.

Santos-Alamillos, F.J., Pozo-Vázquez, D., Ruiz-Arias, J.A., Lara-Fanego, V., Tovar-Pescador, J., 2012. Analysis of spatiotemporal balancing between wind and solar energy resources in the Southern Iberian Peninsula. J. Appl. Meteorol. Climatol. 51, 2005–2024. https://doi.org/10.1175/JAMC-D-11-0189.1, arXiv:https://doi.org/10.1175/JAMC-D-11-0189.1.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., Ratier, A., 2002. An introduction to meteosat second generation (MSG). Bull. Am. Meteorol. Soc. 83, 977—992. https://doi.org/10.1175/1520-0477(2002)083 < 0977:AITMSG > 2.3.CO;2, arXiv:https://doi.org/10.1175/1520-0477(2002)083 < 0977:AITMSG > 2.3.CO;2.

Schroedter-Homscheidt, M., Killius, N., Guevara, D.M., Sirch, T., Hanrieder, N., Wilbert, S., Yasser, Z., 2018. Satellite-based DNI nowcasting based on a sectoral atmospheric motion approach. In: AIP Conference Proceedings 2033, 190015. https://aip.scitation.org/doi/abs/10.1063/1.5067200, https://doi.org/10.1063/1.5067200, arXiv:https://aip.scitation.org/doi/pdf/10.1063/1.5067200.

Sengupta, M., Habte, A., Kurtz, S., Dobos, A., Wilbert, S., Lorenz, E., Stoffel, T., Renné, D., Gueymard, C.A., Myers, D., Wilcox, S., Blanc, P., Perez, R., 2015. Best practices handbook for the collection and use of solar resource data for solar energy applications. research report technical report NREL/TP-5D00-63112. National Renewable Energy Laboratory. URL https://hal-mines-paristech.archives-ouvertes.fr/hal-01184753. this report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Sirch, T., Bugliaro, L., Zinner, T., Möhrlein, M., Vazquez-Navarro, M., 2017. Cloud and DNI nowcasting with MSG/SEVIRI for the optimized operation of concentrating solar power plants. Atmospheric. Meas. Tech. 10, 409–429. https://doi.org/10.5194/amt-10-409-2017. <https://www.atmos-meas-tech.net/10/409/2017/>.

Skamarock, W.C., Coauthors, 2008. A Description of the Advanced Research WRF Version 3. Technical Report. NCAR Technical Note NCAR/TN-475+STR. doi: 10.5065/D68S4MVH.

Thompson, G., Eidhammer, T., 2014. A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. J. Atmospheric Sci. 71, 3636–3658. https://doi.org/10.1175/JAS-D-13-0305.1, arXiv:https://doi.org/10.1175/JAS-D-13-0305.1.

Trigo, R.M., Osborn, T.J., Corte-Real, J.M., 2002. The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms. Clim. Res. 20, 9–17. <https://www.int-res.com/abstracts/cr/v20/n1/p9-17/>.

Wang, P., van Westrhenen, R., Meirink, J.F., van der Veen, S., Knap, W., 2019. Surface solar radiation forecasts by advecting cloud physical properties derived from Meteosat Second Generation observations. Sol. Energy 177, 47–58. https://doi.org/10.1016/j.solener.2018.10.073. <http://www.sciencedirect.com/science/article/pii/S0038092X18310521>.

Wolff, B., Kühnert, J., Lorenz, E., Kramer, O., Heinemann, D., 2016. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. Sol. Energy 135, 197–208. https://doi.org/10.1016/j.solener.2016.05.051. <http://www.sciencedirect.com/science/article/pii/S0038092X16301682>.

Zhang, J., Florita, A., Hodge, B.M., Lu, S., Hamann, H.F., Banunarayanan, V., Brockway, A.M., 2015a. A suite of metrics for assessing the performance of solar power forecasting. Sol. Energy 111, 157–175. https://doi.org/10.1016/j.solener.2014.10.016. <http://www.sciencedirect.com/science/article/pii/S0038092X14005027>.

Zhang, J., Hodge, B.M., Lu, S., Hamann, H.F., Lehman, B., Simmons, J., Campos, E., Banunarayanan, V., Black, J., Tedesco, J., 2015b. Baseline and target values for regional and point PV power forecasts: Toward improved solar forecasting. Sol. Energy 122, 804–819. https://doi.org/10.1016/j.solener.2015.09.047. <http://www.sciencedirect.com/science/article/pii/S0038092X1500540X>.