Universidad de Jaén

Escuela de Doctorado

PHD THESIS

•

IMPROVEMENTS OF SHORT-TERM SOLAR RADIATION FORECASTING TECHNIQUES BASED ON SATELLITE AND SKY-CAMERA IMAGERY

PRESENTED BY:

FRANCISCO JAVIER RODRÍGUEZ BENÍTEZ

DIRECTED BY:

ANTONIO DAVID POZO VÁZQUEZ

JAÉN, OCTOBER 2019

ISBN

Universidad de Jaén

Escuela de Doctorado

TESIS DOCTORAL

•

MEJORAS DE LAS TÉCNICAS DE PREDICCIÓN DE LA RADIACIÓN SOLAR A CORTO PLAZO BASADAS EN IMÁGENES DE SATÉLITE Y CÁMARA DE CIELO

PRESENTADA POR:

FRANCISCO JAVIER RODRÍGUEZ BENÍTEZ

DIRIGIDA POR:

ANTONIO DAVID POZO VÁZQUEZ

JAÉN, OCTUBRE 2019

ISBN

El **Dr. Antonio David Pozo Vázquez**, Profesor Titular del Departamento de Física de la Universidad de Jaén, España.

CERTIFICA:

Que la presente memoria, titulada "**Improvements of short-term solar radiation forecasting techniques based on satellite and sky-camera imagery**", ha sido realizada bajo su dirección. Y considerando que representa trabajo de Tesis, autoriza su presentación y defensa para optar al grado de Doctor con mención de Doctor Internacional.

Dr. Antonio David Pozo Vázquez

Departamento de Física

Universidad de Jaén

Memoria presentada para optar al grado de:

Doctor, con mención de Doctor Internacional

Francisco Javier Rodríguez Benítez

Ingeniero de Telecomunicación

Jaén, octubre de 2019

Mi madre me ha enseñado a disfrutar de la vida,
mi padre, a esforzarme para conseguir mis objetivos,
y mis hermanos, a tomar buena perspectiva de ella.

Acknowledgements

The studies here addressed have been founded by the Spanish Ministry of Economy and Competitiveness, by grating the projects ENE2014-56126-C2-1-R and ENE2014-56126-C2-2-R (http://prosol.uc3m.es). In addition, the MATRAS reseach group of the Universidad de Jaén is supported by the Junta de Andalucía (Reseach group TEP-220) and also by FEDER founds. The authors thank all the support provided. Moreover, the authors are in debt with EUMETSAT, National Centers for Environmental Prediction (NCEP), Abengoa Solar, Faculdade de Ciências da Universidade de Lisboa, and Grupo de Energía Solar of the Universidad Politécnica de Madrid, for the data provided.

Agrademimientos

Parece que fue hace pocos meses cuando empecé mi colaboración con el grupo MATRAS de la Universidad de Jaén. Desde entonces, siempre me ha parecido la etapa más profesional de mi vida y ahora siento que me ha aportado mi especialidad. Además de haber sido una etapa fructífera ha sido muy divertida. No pensaba que podía fascinarme tanto un trabajo.

David, como ya se ha comentado en más de una ocasión, nos tratas con una cercanía, preocupación, cuidado e interés que va más allá de lo laboral. Joaquín, tú, desde la trastienda, haces que todo siga funcionando y ayudas a crear este espacio de trabajo tan agradable.

Tengo la impresión de haber vivido la transición entre dos versiones del grupo MATRAS. No puedo olvidar a los que ya marcharon a otros proyectos y a los que sigo viendo con cariño y admiración: Samuel, José Antonio y Vicente. A los integrantes de la versión más moderna del grupo los considero, además, mis amigos: Miguel Ángel, Antonio, Miguel y Paco. Con ellos se está tan a gusto que las horas de trabajo no suponen un esfuerzo. Y las que no han sido de trabajo se me han quedado cortas. Me lo paso tan bien con ellos y me hacen sentir tan integrado y valorado que creo que va a ser difícil igualar esta sensación en otros trabajos. La idea de compañerismo y amistad quedan cortas para expresar mi relación con Clara. Además de sus aptitudes profesionales, ella hace que uno se sienta mejor, más

centrado, tranquilo y feliz. Todos estos años a su lado han sido maravillosos.

No puedo olvidar a mi otros amigos. Me siento tan bien con ellos que me cuesta adaptarme a que vivan lejos. Ojalá pronto pueda disfrutar de ellos, si no con la frecuencia que quiero, al menos más a menudo. A ellos les debo momentos inolvidables que todavía dibujan una sonrisa cuando los recuerdo.

Finalmente, estoy muy orgulloso de mi familia. Les debo mucho, y es que a veces soy un poco difícil de llevar. Además de mis padres y mis hermanos, de los que tanto he aprendido y más quiero aprender, están las nuevas incorporaciones, mis preciosos sobrinos, que cada día me fascinan más.

A todos ellos, gracias.

Abstract

Nowadays, the traditional power system is moving toward a newone with higher renewable energy penetration. Nevertheless, such power system presents serious challenges when becoming a reality. The problem arises because renewable energy generation, particularly wind and solar based resources, is weather dependent, and this dependence is translated into the power supply. Independently of the alternative proposed with the aim of paliating this effect, the forecasting of resources is showed to be a promising tool in order to design the managing strategy. This study aims at promoting, particularly, the solar energy penetration into the electricity grid through: 1) ease the transmission system operator tasks, 2) enhance the participation of solar energy into the electricity market, and 3) ease the solar plant managing system operator tasks. All of these purposes cover different prediction horizons but mainly those of nowcasting (0 to 3 h) and short-term forecasting (0 to 6 h). These forecasting horizons are increasingly becoming important both for solar plants management and also for the grid operation purposes. To this end, two different solar resource prediction technologies are mainly assessed in this thesis: sky-camera and satellite imagery. Due to the fact that there are different solar-based generation technologies in the current electricity market, the evaluation is done in light of both global horizontal irradiance and direct normal irradiance solar components.

On the one hand, the development of solar nowcasting models

based on sky-cameras has increased during the last years. As a consequence, several approaches have emerged, concerning from the own capture devices (total sky imagers, low-cost surveillance cameras) until particular algorithms (e.g. cloud detection methods) and purposes (solar resource nowcasting, cloud type classification). The thesis includes the detailed description and assessment of state-of-the-art sky-camera-related methods in order to perform all of these approaches.

On the other hand, satellite imagery is showed to be a suitable tool for short-term and also for nowcasting purposes. The particular solar resource forecasting satellite-based procedure here developed is inspired in Heliosat-2 method, which uses Meteosat Second Generation (MSG) imagery. Morever, a new algorithm is proposed in order to enhance the accuracy of the solar forecasts. This alghorithm tries to reproduce the observed streamlines, i.e, the particles flow within the satellite image. In addition to the standard MSG visible channels 1 and 2, the HRV channel is also used to compute high spatial resolution forecasts.

The forecasting models performance usually depends on features such as latitude, geography or topography. Moreover, in the studies included in this thesis, the results are assessed in light of cloudiness, particularly, the cloud type regime. The relationship between the solar resource variability and different weather patterns (being the latter related with the cloudiness) is thorough analyzed.

Finally, once a detailed assessment of several short-term solar prediction models is conducted, it is explored the performance of blended models, computed using machine learning techniques, that use the former as predictors. The reached performance of blended approaches definitely outperforms that of the original models.

Resumen

Actualmente, el sistema enegético tradicional está migrando hacia uno nuevo con una penetración más alta de renovables. Sin embargo, este sistema energético presenta importantes retos a la hora de convertirse en una realidad. El problema aparece porque la generación de energía renovable, en especial la basada en recursos solar y eólico, es dependiente del tiempo, y esta dependencia es transferida al suministro energético. Con independencia de la alternativa que se proponga para paliar este efecto, la predicción de los recursos demuestra ser una herramienta prometedora para diseñar estrategias de contingencia. Este estudio quiere promover, en particular, la penetración de la energía solar en la red eléctrica a través de: 1) facilitar las tareas de los operadores del sistema de transmisión, 2) aumentar la participación de la energía solar en el mercado eléctrico, y 3) facilitar las tareas de los operadores del sistema de gestión de plantas solares. Todos estos propósitos repercuten en diferentes horizontes del pronóstico pero principalmente aquellos de predicción a muy corto plazo (0 a 3 horas) y predicción a corto plazo (0 a 6 horas). Estos horizontes de predicción se están convirtiendo en cada vez más importantes para los propósitos de gestión de plantas solares y operación de la red. Con este propósito, se evalúan principalmente en esta tesis dos tecnologías diferentes de predicción solar: cámara de cielo e imágenes de satélite. Debido al hecho de que hay diferentes tecnologías de generación basadas en energía solar en el actual mercado eléctrico, la evaluación es realizada en virtud de ambas componentes solares de

irradiancia global horizontal e irradiancia directa normal.

Por un lado, el desarrollo de modelos de predicción solar a muy corto plazo basados en cámaras de cielo se ha incrementado durante los últimos años. Como consecuencia, han aparecido varios enfoques, que conciernen tanto a los propios dispositivos de captura (cámaras de cielo, cámaras de seguridad de bajo coste) como a algoritmos especiales (e.g. métodos de detección de nube) y propósitos (predicción a muy corto plazo del recurso solar, clasificación del tipo de nube). La tesis incluye la descripción detallada y evaluación de métodos en el estado del arte relacionados con cámara de nube con el fin de abordar todos estos enfoques.

Por otro lado, las imágenes de satélite han mostrado ser una herramienta adecuada para propósitos de predicción a corto plazo y también a muy corto plazo. El procedimiento particular de predicción del recurso solar basado en satélite aquí desarrollado está inspirado en el método Heliosat-2, que utiliza imágenes de Meteosat Second Generation (MSG). Además, se propone un nuevo algoritmo para aumentar la precisión de las predicciones solares. Este método intenta reproducir las líneas de corriente observadas, i.e., el flujo de partículas en una imagen de satélite. Además de los típicos canales del visible 1 y 2 del MSG, el canal HRV es también utilizado para calcular predicciones de alta resolución espacial.

El desempeño de los modelos de predicción dependen usualmente de características tales como la latitud, la geografía o la topografía. Además, en los estudios incluidos en esta tesis, los resultados son evaluados en virtud de la nubosidad, en particular, del régimen de tipo de nube. La relación entre variabilidad del recurso solar y diferentes patrones de tiempo (estando estos últimos relacionados con la nubosidad) es analizada en profundidad.

Finalmente, una vez se realiza que una valoración detallada de diferentes modelos de predicción a corto plazo, se explora el rendimiento de modelos combinados, calculados usando técnicas de *machine learning*, que emplean a los primeros como indicadores. El rendimiento alcanzado por los enfoques de combinación mejora definitivamente al de los modelos originales.

Contents

List of Figures

List of Tables

Índice

CHAPTER 1

Introduction

1.1 Motivations

At least until 2018, the fossil fuels still represented the majority of energy consumption resources in the world (BP, 2019). However, in addition to their harmful impacts on Climate Change and on human health, their availability is limited. In addition, the production of important materials also depends on some fossil fuels. Due to the expected growing of the population and economy, by 2040 the energy demand is expected to rises more than 25% (OECD/IEA, 2018). So that, the energy supply may be a future challenge. To this end, the development and deployment of feasible renewable energy alternatives mae be a key to replace traditional fossil fuel energy resources. Taking into account, not only the current energy needs but also the future growing in the energy demand, the renewable energies are shown to be able to provide the requested power (Jacobson and Delucchi, 2011). Withing the set of renewable energies, the most important in terms of electricity generation share are wind and solar energy, being the latter the focus of this study.

There are two main technologies in the field of solar energy elec-

tricity generation: photovoltaic (PV) and concentrating solar power (CSP) plants. The former directly produces electricity by means of photoelectric cells. Therefore, only the global horizontal irradiance (GHI) solar component is required. Thus, this technology still works under cloudy skies. On the other hand, CSP technology makes use of the direct normal irradiance (DNI) solar component. In this case, a set of mirrors concentrates the solar beam into a reduced area, where a heat transfer fluid is flowing. The electricity production is based on a thermal cycle thereafter. Finally, an additional solar technology may be cited: the concentrated photovoltaic (CPV), which concentrate the solar beam into a photovoltaic cell.

Unlike fossil-fuels-based energy resources, the availability of wind and solar energy is intermittent since they are dependent of the weather. This fact represents and important drawback in the their penetration into the electricity grid because the balance between generation and consumption is requested all the time. Traditionally, some reserves are needed with the the aim of handling generation mismatches (Yang et al., 2018). However, the uncertainty provided by weather-dependent renewables forces to enhace the reserves, thus increasing the costs and hampering the transmission system operators (TSOs) tasks.

The solar energy fluctuations in the top of the atmosphere are accurately estimated because they are produced by deterministic causes: the combination of the annual and diurnal cycle, and the latitude. On the other hand, fluctuations at the ground surface are slighly conditioned by the aerosol load and mainly affected by clouds (Matuszko, 2012). The particular features of the latter (e.g. shape, optical properties, extension) determine the attenuation of the solar radiation, being its DNI component more sensitive than the GHI. Due to the critical impact of clouds in the radiation at the ground surface, their accurate representation has become a crucial task. Traditionally, some cloud properties from human ground observations have been collected, but this method is expensive and the accuracy is poor. Nowadays some cloud properties could be retrieved from satellite products from EUMETSAT (Kniffka et al., 2014) or the APOLLO project (Kriebel et al., 2003; Wey and Schroedter-Homscheidt, 2014). However, these resources might be limited. Alternatively, since recent years, captures from ground-based sky cameras have been used to this end. They may be a feasible data source to infer some information of clouds, like cloud type and cloud coverage, in a limited

area.

Cloudiness produces solar radiation variability and, depending on cloud properties, the variability changes in space and time. Several studies addressed the temporal and spatial variability and its relation (Gueymard and Wilcox, 2011; Lave et al., 2017; Barnett et al., 1998; Perez et al., 2012). As a conclusion, correlation increases whith temporal aggregation and when the distance is reduced (Hoff and Perez, 2012; Mills and Wiser, 2010; Perez et al., 2016). The time and space scales determine the area of application of this knowledgement. A limited area and time ranges of seconds and minutes are related with plant management (Lave et al., 2012). An area of hundreds of kilometers and times about tens of minutes or hours are related with grid integration purposes (Perez et al., 2016). Finally, plant project development requires time ranges of seasons and years (Lohmann et al., 2006).

The difficulty that variability of solar resource poses for its grid integration will increase in the future as the participation in power mix of this energy also increases. In the Iberian Peninsula, despite that its potential for solar energy (Santos-Alamillos et al., 2017a), the electricity demand covered by solar installed capacity is about 4.8% (www.ree.es, 2018). However, in the next years it is expected an important deployment of renewables around the world (International Energy Agency, 2018).

In order to paliate the output solar energy variability, some storage solutions can be used. However, their costs are still elevated. Alternatively, spatial aggregation could also smooth this variability. In this aspect, the case of the Iberian Peninsula is particularly challenging due to the fact that it is highly isolated from Europe in terms of electricity interconnections (General Court, 2015). In any case, the anticipation is a key question. To this end, accurate solar forecast is a tool that eases the management tasks of solar energy. Depending on the lead times, the forecast is called in a different way and its potential users also differ. The nowcasting covers up to 1 hour lead time (Chow et al., 2011) and it is useful for power plant management issues. If the purpose is to participate in the electricity market or the integration into the grid, the suitablest predictions may be the short-term forecasts, that cover up to 6 hours lead time (Schroedter-Homscheidt et al., 2009). Lastly, the solar predictions called forecasting ranges from 6 hours to 48 hours (Lorenz et al., 2009) and also

may cover the same purposes than the short-term forecasting.

The predicted value from forecasts may not only be the solar radiation but also its variability. The existence of clouds and even some of their features could be related to the synoptic weather pattern that rules an area. Thus, some atmospheric variables might be used as descriptors/predictors of weather pattern, and therefore, of the associated variability. In the Iberian Peninsula, the atmospheric circulation, that generates the weather patterns, was largely addressed in several studies (Castro-Díez et al., 2002; Trigo et al., 2002a).

In the context of solar forecasting, there are four main types of models.

- Image-based models. They have two main steps: (1) the pixel selection at certain future time lapse through cloud motion vector (CMV) field computation and, (2) the feature-to-radiation step, where some key feature is extracted from the image in order to infer the ground surface radiation. Both satellite-based and sky-camera-based models fit into this type. While the pixel selection task could be extrapolated between sky-camera and satellite, the needed feature to infer the radiation is different. In case of sky-camera models the prediction is computed by the combination of the red, green, and blue channels in a regular image. The predictions obtained by sky-cameras are showed to be suitable for nowcasting purposes. On the other hand, satellite models have probed to be suitable for short-term horizons. In this case, the most popular key feature is the so-called clearness index (K_c) which is computed by the Heliosat method (Cano et al., 1986).

- Numerical weather prediction (NWP) models. These models try to reproduce the intrincate physical interactions between several atmospheric variables (boundary conditions) along time. They deal with the continuity equation and the ideal gas law and try to integrate these differential equation in both space and time. NWP models can simulate the generation, dissipation and advection of clouds, optionally with some particular parameterization, and the interaction with the solar rays is computed. However, they still lack the ability to accurately compute the cloud amount and location (Deng et al., 2014; Arbizu-Barrena et al., 2015). The forecast based on NWP mod-

els are traditionally suitablest from short-term scale onward (Mathiesen and Kleissl, 2011).

- Hybrid models. They combine the information from satellite and NWP models. For instance, the cloud fields of satellite are advected using the winds computed by a NWP in the CIRACast forecasting algorithm (Miller et al., 2018). Another hybrid models are the so-called MADCast (Descombes et al., 2014) model or the CIADCast algorithm (Arbizu-Barrena et al., 2017). In the latter, contrarily to the CIRACast approach, the cloud index variable computed by Heliosat method is advected and diffuses within the WRF model (Skamarock and Coauthors, 2008).

- Blended models. In this approach, some statistical techniques are used in order to optimally combine several forecasting models (Lorenz et al., 2012a; Wolff et al., 2016). Based on the premise that not a single model better performs at all conditions (i.e. at all locations, over all terrains, at all forecasting horizons or under all sky conditions), it is suggested that all of them can complement each others. So that, an optimally combination of several models can outperforms the individual ones.

1.2 State of the art in image-based solar radiation forecasting

Image-based solar radiation forecasting methods are commonly used in nowcasting (i.e. from minutes to tens of minutes) and short-time forecasting (i.e. up to 3-6 hour lead time). Two main data sources are here evaluated: sky-cameras and satellite, traditionally being more suitable for nowcasting and short-term purposes, respectively. Despite the different data source, the forecasting procedure is similar and it is composed by two main steps: the CMV computation and the image-to-ratiation calculation. From the first step, some pixels in the image are selected as current and future predictors of solar radiation. At the second step, the image is processed in order to compute the feature that is then translated into radiation. An additional component is also used in both prediction algorithms: the

clear sky-model. The state of the art and main parts of models is here presented:

1.2.1 CMV

The objective of this algorithm is to compute a vector field that represents the displacement between some consecutive images. The state of the art covers two main techniques: PIV (Mori and Chang, 2003) and Optical Flow (Horn and Schunck, 1981) algorithms. Despite of their different approaches: they share some similarities: the resolution of the resulting vector field depends of the window search configuration, there are a lot of parameters, they are based on single-channel consecutive images, any parameterization could not obtain a "valid" result and the reached quality also depends on the own images.

The PIV algorithm compute the displacement of a sub-image in the previous image by the better correlation between itself and several neighbouring sub-images in the current image. A free implementation of this algorithm is used in this work, called OpenPIV.

There have been several improvements since the first optical flow algorithm. A popular version of this algorithm was proposed by Lucas et al. (1981). This algorithm is based on the premise that the intensity of same pixels in consecutive images is almost constant. In recent years, others technologies have been applied over this research field and, as a result, the DeepFlow (Weinzaepfel et al., 2013) algorithm has been released. The most important advantages of this algorithm are: (1) with no particular parameterization, it has demonstrated to be robust when applied over very different images, and (2) the resulting resolution is the maximum available.

1.2.2 Clear-sky models

Along the last decades, several clear-sky models have been developed with the aim of accurately estimate the solar radiadiation reaching the ground under clear sky conditions. Their performance differences raises from the capability of manage some atmospheric features. For instance, the European Solar Radiation Atlas (ESRA) model (Rigollier et al., 2000) is used in conjunction with the monthly Linke tur-

bidity (Remund et al., 2003) parameter. This clear sky model is used as a part of Heliosat-2 (Rigollier et al., 2004) solar radiation forecasting model. Other atmospheric parameters such as water vapour or aerosol optical depth are used by more elaborated models (Gueymard, 2004; Ineichen, 2008). The recent McClear (Lefèvre et al., 2013) clear-sky model, which uses a direct radiative transfer model to infer the solar radiation, has been included as a part of the newly released Heliosat-4 (Qu et al., 2017) solar radiation forecasting model.

1.2.3 Sky-camera

A sky imager is a camera that produces images of the sky. With the aim of capture the whole celestial dome within a sigle image, these cameras includes a fish-eye len, which introduces an expected distortion in the resulting images. Due to certains areas are needed to be represented accurately in the images, the inherent distortion must be taken into account. To this end, two main approaches must be highlighted, being both based on the relation between known positions in the real world and their representation on the images. Marquez and Coimbra (2013) suggested to infer the distortion through the theoretical sun position and its representation in the captures. As a result, a transformation could be applied into the raw images in order to project them. Alternatively, the intrinsic parameters of camera could be inferred through control images (Gauchet et al., 2012) where it is captured a known object with special features. A popular resource is a chessboard, where the distances of the inner corners are known. Within the inferred intrinsic parameters, a transformation of raw images could be applied and similar projection than in case of that reached by Marquez and Coimbra (2013) is obtained. Despite the fact that the intrinsic parameters could be helpful to another tasks, in the works here presented, it is used the approach proposed by Marquez and Coimbra (2013) due to the georeferentiation issue, i.e. the geo-location of each pixel within the image, is done in a intuitive way. To this end, it is requested to know the cloud height by menas, for instance, of a ceilometer.

One of the most important and challenging task in a sky-camera forecasting algorithm is the segmentation of the image to discern into cloudy and cloud-free pixels. This method is not dependent of the distortion correction step and its handicap is the managing of a com-

plex casuistry, for instance, overcast but bluish skies, burned pixels, droplets, flares, reddish clouds of sunrise, etc. The source of data is limited, being the simpliest the red green and blue channels of a common image. The state of the art includes several approaches, since simple thresholding technique on single-channel image, an elaborated set of thresholds (Kazantzidis et al., 2012) applied over the raw image, cloudy sky libraries (Chow et al., 2011), etc. However, neither of those can be exported to a different sky camera without a change on its parameterization. Due to its extrapolation approach, the method proposed by Li et al. (2011) is here used as a reference. Once configured, its performance has been successfully assessed. However, some inner algorithm could not be intuitive enought. So that, this issue is here replaced by a simple clustering algorithm but keeping the overall approach.

Previous pixel-to-irradiation step approaches was fully based on the own image. However, in the lastest version, this step is different for DNI and GHI forecasting, but in both cases, the cloud fraction (CF) computation is required. In order to compute the DNI forecasting, the premise that the DNI solar component is produced just by those solar rays comming directly from the sun. Thus, the DNI prediction is computed by product of the current/future clear-sky DNI estimation and the complementary of the current/future cloud fraction. The latter is assessed over the location of interest within the image and with the same shape of the sun. On the other hand, GHI prediction at a certain lead time, is also based on 30-min backward 2-groups K_c distribution (Urquhart et al., 2013). The K_c is the ratio between a measured GHI over the clear-sky estimation. Once the distributions are computed, the previously computed current/future CF is used to select the suitablest group. Then, the K_c-centroid of the selected group is used to infer the predicted GHI in the same way as in case of DNI forecast.

1.2.4 Satellite

Solar radiation forecasting based on satellite imagery is commonly computed by the Heliosat method. Since the first version (Cano et al., 1986), several improvements have been included. The most recent release, the Heliosat-4 method (Qu et al., 2017), includes the McClear clear-sky model, which uses a radiative transfer algorithm (Lefèvre

et al., 2013). In addition to Meteosat Second Generation satellite (MSG) (Schmetz et al., 2002), retrievals from project AVHRR Processing scheme Over cLouds, Land and Ocean (APOLLO) are collected with the aim of taking into account some cloud properties such as cloud optical depth, cloud coverage, cloud type, etc. Nevertheless, due to the sophistication of the Heliosat-4 method and the difficulty to manage all of the involved data sources, a previous release of Heliosat method is here used, Heliosat-2 method (Rigollier et al., 2004).

Based on the MSG visible channels, the premise of the Heliosat-2 method is that if the apparent albedo is higher than the ground albedo, it is as a consequence of clouds, with some exception such as snow-covered areas, for instance. Thus, the ground albedo is firstly computed. As a particular implementation, and following Kleissl (2013), the diurnal variability is taken into account in this work. Then, the ground albedo can be discounted from the apparent albedo and the result is normalized using the albedo produced by most reflective cloud. The reached result is the so-called cloud index (n) and its nearly-complementary is the so-called clearness index (K_c). Thus, the GHI is computed as the product of K_c by the GHI clear-sky estimation coming from ESRA model.

Unlike the sky-camera case forecasting model, where only an overall vector is used in the prediction procedure, the complete CMV field is here requested, but it is not directly used. Alternatively, inspired by Nonnenmacher and Coimbra (2014), we have developed an algorithm with the aim of improving the accuracy of the satellite-based solar forecasts. To this end, first, the streamlines of the atmospheric movement are derived from the CMVs. Second, the streamline overpassing the station of interest is used for estimating the future cloud index at the required horizon. Finally, the solar forecast is computed from this future cloud index.

To derive DNI forecasts from the satellite images, we have used the DirIndex method (Perez et al., 2002). Note that, independently of the temporal resolution parameterized for the forecasting procedure, the GHI inputs has a time step of 15 min.

1.3 Objectives

The overall objective of the thesis is the improvement of short-term solar radiation forecasting methods, i.e. from 0 to 6 hours, based on sky-cameras and satellite. The thesis may be split in three different parts that do not represent the chronologycal progress. In the first part, the state-of-the-art technologies are provided. In case of sky-camera, several methods compose the forecasting process and all of them have been updated during years, being the latest available presented in this first part. In addition, a particular use of sky-camera with the purpose of providing an automatic cloud type recognition is presented and assessed. On the othe hand, the Heliosat-2 method is implemented and the up-to-date technology in terms of CMV is provided. The latter is the results of previous studies about comparison and calibration of other popular technologies. After the computation of CMV, an additional process is added to the forecasting procedure: the streamlines calculation. This process aims at providing more reliable, intuitive and traceable solar radiation predictions.

The second part is the study of the particular modes of weather patterns that determine the solar radiation variability in the Iberian Peninsula. The objective of this study is to anticipate the mode of variability of solar radiation as an alternative to the own solar radiation forecast. The proposed method encloses an unsupervised classification techique, non-parametric statitical analyses and a particular cluster number decision making.

The last part accounts for a comprehensive assessment of four different state-of-the-art short-term solar radiation forecasting models with independent approaches. After that, a comparison of several blending approaches based on machine learning techniques is presented, in order to choose the suitablest blended solar radiation forecasting model.

Therefore, the particular objectives of this study are:

- The development of sky-camera images distortion correction techniques.
- The study and assessement of image segmentation techniques focused on cloudy pixel recognition.
- The develpment, calibration and assessment of CMV approaches.

- The development and evaluation of several image-to-irradiation procedures for both GHI and DNI.
- The implementation and evaluation of the of Heliosat method, in particular, the Heliosat-2 method taking into account the diurnal variability. In addition, the implementation and evaluation of the HRV variant.
- The development of a streamline method based on CMV fields aimed at improving solar forecasting accuracy based on satellite images.
- The study, development and assessment of a method to discern the main modes of solar resource varibility.
- The exhaustive evaluation of several state-of-the-art solar radiation forecasting models.

1.4 Thesis summary

The results reached by the completion of this thesis may improve the penetration of solar energy into the electricty share. An improved knowledgement of solar resource uncertainty may provide tools to: (1) ease the TSOs decision making, (2) enhance its potential in the electricity market, and (3) ease the power plant management.

Chapters 3 - 7 of this thesis are a set papers that are published, submited or in preparation. The abstracts of the chapters are:

Chapter 3. The comparison of solar radiation high-temporal resolution nowcasts based on: 1) a set of three sky-cameras simultaneously, 2) Meteosat Second Generation (MSG) satellite, and 3) smart persistence approach, is presented in this paper. Two different satellite imagery datasets are used, the standard MSG channels 1 and 2, and also the HRV channel. All of the methods involved in the forecasting procedures are detailed. The reached results are assessed in light of different cloud types (altocumulus, cirrocumulus, cumulus, stratocumulus, multi-cloud, and the aggregation of all of the samples) and solar radiation component (GHI and DNI). Overall, the Smart Persistence approach is the best performing model until early horizonts but

then it is outperformed by: first sky-cameras and second satellite up to 60 min lead-time. The Satellite HRV model is not more suitable than standard Satellite model.

Chapter 4. A methodology, aimed to be fully operational, for automatic cloud classification based on the synergetic use of a sky camera and a ceilometer is presented. The random forest machine learning algorithm was used to train the classifier with 19 input features: 12 extracted from the sky camera images and 7 from the ceilometer. The method was developed and tested based on a set of 717 images collected at the radiometric stations of the Univ. of Jaén (Spain). Up to nine different types of clouds (plus clear sky) were considered (clear sky, cumulus, stratocumulus, nimbostratus, altocumulus, altostratus, stratus, cirrocumulus, cirrostratus, and cirrus) plus an additional category multicloud, aiming to account for the frequent cases in which the sky is covered by several cloud types. A total of eight experiments were conducted by: (1) excluding/including the ceilometer information; (2) including/excluding the multicloud category and (3) using six or nine different cloud types, aside from the clear-sky and multicloud category. The method provided accuracies ranging from 45% to 78%, being highly dependent on the use of the ceilometer information. This information showed to be particularly relevant for accurately classifying "cumuliform" clouds and to account for the multicloud category. At this regard, the camera information alone was found to be not suitable to deal with this category. Finally, while the use of the ceilometer provided an overall superior performance, some limitations were found, mainly related to the classification of clouds with similar cloud base height and geometric thickness.

Chapter 5. The intra-day modes of variability of the solar resources in the Iberian Peninsula, their associated weather patterns and their impact on the solar power output are assessed in this work. The analysis is performed for yearly and seasonal variability. Firstly, the modes of variability are identified by means of hierarchical cluster analysis. It is computed with two years of measured global horizontal irradiance (GHI) and direct normal irradiance (DNI) data gathered at four stations. Notably, three-hour statistics describing mean and variability of solar radiation are used as input to the cluster analysis. Secondly, synoptic weather patterns associated with each group

resulting from the cluster analysis are assessed using sea level pressure and cloudiness data. Finally, the solar PV power yield associated with each mode is evaluated. The yearly analysis reveals the existence of four modes of variability of the solar resource in the study area. The four modes are shown to have a distinctive weather pattern and also specific impacts on solar power generation in the study area. Seasonal analyses show results similar to the annual analysis, but with marked seasonal differences.

Chapter 6. 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 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). 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 conditions 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, attributed 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, complementary. The advantages attained by this complementarity are further explored in a companion paper (Part II).

Chapter 7. In this article we explore the blending of the four models (Satellite, WRF-Solar, Smart Persistence and CIAD-Cast) studied in Part 1 by means of Support Vector Machines with the aim of improving GHI and DNI forecasts. Two blending approaches that use the four models as predictors have been studied: the horizon approach constructs a different blending model for each forecast horizon, while the general approach trains a single model valid for all horizons. The influence on the blending models of adding information about weather types is also studied. The approaches have been evaluated in the same four Iberian Peninsula stations of Part 1. Blending approaches have been extended to a regional context with the goal of obtaining improved regional forecasts. In general, results show that blending greatly outperforms the individual predictors, with no large differences between the blending approaches themselves. Horizon approaches were more suitable to minimize rRMSE and general approaches work better for rMAE. The relative improvement in rRMSE obtained by model blending was up to 17% for GHI (16% for DNI), and up to 15% for rMAE. Similar improvements were observed for the regional forecast. An analysis of performance depending on the horizon shows that while the advantage of blending for GHI remains more or less constant along horizons, it tends to increase with horizon for DNI, with the largest improvements occurring at 6 hours. The knowledge of weather conditions helped to slightly improve further the forecasts (up to 3%), but only at some locations and for rRMSE.

CHAPTER 2

Introducción

2.1 Motivaciones

Al menos hasta el 2018, los combustibles fósiles todavía representaban la mayoría de los recursos de consumo energético en el mundo (BP, 2019). Sin embargo, además de sus impactos perjudiciales en el Cambio Climático y en la salud humana, su disponibilidad es limitada. Además, la producción de importantes materiales también depende de algunos combustibles fósiles. Debido al crecimiento esperado de la población y la economía, se espera que la demanda energética crezca más de un 25% en 2040 (OECD/IEA, 2018). Así que, el suministro energético puede ser un problema complicado. Con esta finalidad, el desarrollo y despliegue de factibles alternativas de energía renovable puede ser una clave para reemplazar a los tradicionales recursos de combustibles fósiles. Teniendo en cuenta, no solo las actuales necesidades energéticas sino también la futura creciente demanda de energía, las energías renovables muestran ser capaces de proveer la energía requerida (Jacobson and Delucchi, 2011). Dentro del conjunto de energías renovables, las más importantes en el conjunto de generación eléctrica son la eólica y solar, siendo esta última el foco de este estudio.

Existen dos tecnologías principales en el campo de la generación eléctrica solar: las plantas fotovoltaicas (PV) y las de concentración (CSP). Las primeras producen energía directamente mediante células fotovoltaicas. Por lo tanto, solo se requiere la componente solar de irradiancia global horizontal (GHI). Así, esta tecnología continúa funcionando bajo cielos nubosos. Por otro lado, la tecnología CSP hace uso de la componente solar de irradiancia directa normal (DNI). En este caso, un conjunto de espejos concentra el haz solar en un área reducida, por donde un viaja un fluido caloportador. Desde este punto en adelante, la producción eléctrica está basada en un ciclo térmico. Finalmente, una tecnología solar adicional cabe ser mencionada: la fotovoltaica concentrada (CPV), que condensa el haz solar en una célula fotovoltaica.

Al contrario que los recursos energéticos basados en combustibles fósiles, la disponibilidad del viento y del sol es intermitente dado que dependen del tiempo. Este hecho representa un obstáculo importante para su penetración en la red eléctrica porque se requiere del balance entre generación y consumo en todo momento. Tradicionalmente, se necesitan algunas reservas para manejar desajustes de generación (Yang et al., 2018). Sin embargo, la incertidumbre arrojada por las renovables dependientes del tiempo fuerza a aumentar las reservas, incrementándose así los costes y dificultando las tareas de los operadores del sistema de transmisión (TSOs).

Las fluctuaciones de la energía solar en el tope de la atmósfera son estimadas con precisión porque están producidas por causas determinísticas: la combinación del ciclo anual y diurno, y la latitud. Por otro lado, las fluctuaciones en la superficie terrestre están algo condicionadas por la carga de aerosoles y principalmente influida por las nubes (Matuszko, 2012). Las características particulares de estas últimas (e.g. forma, propiedades ópticas, extensión) determinan la atenuación de la radiación solar, siendo su componente DNI más sensible que la GHI. Debido al impacto crítico de las nubes en la radiación en la superficie terrestre, sus representación precisa se ha convertido en una tarea crucial. Tradicionalmente, algunas propiedades de las nubles han sido obtenidas de observaciones humanas desde tierra, pero este método es costoso y la precisión es pobre. Hoy en día algunas propiedades nubosas podrían ser adquiridas de productos de satélite desde EUMETSAT (Kniffka et al., 2014) o el proyecto APOLLO (Kriebel et al., 2003; Wey and Schroedter-Homscheidt, 2014). Sin embargo, estos recursos pueden ser limitados. Alternati-

vamente, en los últimos años, han sido usadas para este fin capturas de tierra de cámaras de cielo. Estas pueden ser una fuente de datos factible para inferir alguna información de las nubes, como el tipo de nube y la cobertura nubosa, en un área reducida.

La nubosidad produce variabilidad de la radiación solar y, dependiendo de las propiedades de las nubes, esta cambia en espacio y tiempo. Algunos estudios han abordado la variabilidad temporal y espacial y su relación (Gueymard and Wilcox, 2011; Lave et al., 2017; Barnett et al., 1998; Perez et al., 2012). Como conclusión, la correlación aumenta con la agregación temporal y cuando la distancia se reduce (Hoff and Perez, 2012; Mills and Wiser, 2010; Perez et al., 2016). Las escalas temporales y espaciales determinan el área de aplicación de este conocimiento. Un área reducida y rangos temporales de segundos y minutos están relacionados con gestión de plantas (Lave et al., 2012). Un área de cientos de kilómetros y tiempos alrededor de decenas de minutos u horas están relacionados con propósitos de intergación en la red (Perez et al., 2016). Finalmente, el desarrollo de proyectos de plantas requiere rangos temporales de estaciones y años (Lohmann et al., 2006).

La dificultad que la variabilidad del recurso solar supone para su integración en la red se incrementará en el futuro como también crezca la participación de esta energía en el conjunto energético. En la Península Ibérica, a pesar de su potencial para la energía solar (Santos-Alamillos et al., 2017a), la demanda eléctrica cubierta por la capacidad solar instalada es de alrededor del 4.8% (www.ree.es, 2018). Sin embargo, en los próximos años se espera un importante depsliegue de renovables por todo el mundo (International Energy Agency, 2018).

Para paliar la variabilidad de la energía solar producida, se pueden usar algunas soluciones de almacenamiento. Sin embargo, sus costes son todavía elevados. Alternativamente, la agregación espacial también podría suavizar esta variabilidad. En este aspecto, el caso de la Península Ibérica es particularmente complejo debido al hecho de que está altamente aislada de Europea en términos de interconexiones eléctricas (General Court, 2015). En cualquier caso, la anticipación es una cuestión clave. Con este fin, la predicción solar precisa es una herramienta que facilita las tareas de gestión de la energía solar. Dependiendo de los horizontes temporales, la predicción es llamada de una forma diferente y sus potenciales usuarios también

cambian. La predicción a muy corto plazo cubre hasta 1 hora de horizonte (Chow et al., 2011) y es útil para cuestiones de gestión de plantas energéticas. Si el propósito es participar en el mercado eléctrico o la integración en la red, los pronósticos más adecuados pueden ser las predicciones a corto plazo, que cubren hasta 6 horas de horizonte (Schroedter-Homscheidt et al., 2009). Finalmente, los pronósticos solares llamados predicciones abarcan desde 6 hasta 48 horas (Lorenz et al., 2009) y también cubren los mismos propósitos que las predicciones a corto plazo.

El valor a predecir de los pronósticos puede no ser solo la radiación solar sino también su variabilidad. La existencia de nubes e incluso alguna de sus características pueden ser relacionadas con el patrón de tiempo sinóptico que gobierne un área. Así, algunas variables meteorológicas podrían ser usadas como descriptores/indicadores del patrón de tiempo, y por lo tanto, de la variabilidad asociada. En la Península Ibérica, la circulación atmosférica, que genera los patrones de tiempo, fue ampliamente abordada en varios estudios (Castro-Díez et al., 2002; Trigo et al., 2002a).

En el contexto de la predicción solar, hay cuatro tipos principales de modelos.

- Modelos basados en imágenes. Tienen dos grandes partes: (1) la selección de un pixel a un cierto lapso temporal futuro a través del cálculo de los campos de vectores de desplazamiento de las nubes (CMV) y, (2) el paso de característico a radiación, donde algún característico clave es extraído de la imagen para inferir la radiación en la superficie terrestre. Ambos modelos basados en satélite y cámara de cielo encajan en este tipo. Mientras que la tarea de selección del pixel puede ser extrapolable entre cámara de cielo y satélite, el característico necesario para inferir la radiación es diferente. En el caso de los modelos de cámara de cielo la predicción es calculada por la combinación de los canales rojo, verde y azul en una imagen común. Las predicciones obtenidas por cámaras de cielo muestran ser adecuadas para propósitos de predicción a muy corto plazo. Por otro lado, los modelos de satélite han probado ser apropiados para alcances a corto plazo. En este caso, el característico clave más popular es el denominado índice de claridad (K_c), que se calcula por el método Heliosat (Cano et al., 1986).

- Modelos numéricos de predicción meteorológica (NWP). Estos modelos tratan de reproducir las intrincadas interacciones físicas entre varias variables atmosféricas (condiciones de contorno) a lo largo del tiempo. Tratan con las ecuaciones de continuidad y la ley de los gases ideales e intentan integrar estas ecuaciones diferenciales en espacio y tiempo. Los modelos NWP pueden simular la generación, disipación y advección de nubes, opcionalmente con alguna parametrzación particular, y se calcula la interacción con los rayos solares. Sin embargo, les continúa faltando habilidad para calcular con precisión la cantidad y localización de la nube (Deng et al., 2014; Arbizu-Barrena et al., 2015). Las predicciones basadas en modelos NWP son tradicinalmente las más adecuadas desde escalas a corto plazo en adelante (Mathiesen and Kleissl, 2011).

- Modelos híbridos. Combinan la información de satélites y modelos NWP. Por ejemplo, los campos de nubes del satélite son advectados usando los vientos obtenidos por un modelo NWP en el algoritmo de predicción CIRACast (Miller et al., 2018). Otros modelos híbridos son los denominados modelo MADCast (Descombes et al., 2014) o algoritmo CIADCast (Arbizu-Barrena et al., 2017). En este último, contrariamente al enfoque del CIRACast, la variable del índice nuboso calculada por el método Heliosat es advectada y difundida dentro del modelo WRF (Skamarock and Coauthors, 2008).

- Modelos combinados. En este enfoque, son usadas algunas técnicas estadísticas para combinar óptimamente varios modelos de predicción (Lorenz et al., 2012a; Wolff et al., 2016). Basado en la premisa de que un único modelo no es el mejor en todas las condiciones (i.e. en todas las localizaciones, sobre todos los terrenos, en todos los horizontes de predicción o bajo todas las condiciones de cielo), se sugiere que puedan complementarse unos a otros. Así que, una combinación óptima de varios modelos puede superar a los individuales.

2.2 Estado del arte en predicción de la radiación solar basada en imagen

Los métodos de predicción de la radiación solar basados en imágenes son usados comúnmente en predicción a muy corto plazo (i.e. desde minutos a decenas de minutos) y corto plazo (e.e. hasta horizontes de 3-6 horas). Dos principales fuentes de datos son aquí evaluadas: la cámaras de cielo y el satélite, tradicionalmente siendo más adecuados para propósitos de predicción a muy corto plazo y corto plazo, respectivamente. A pesar de las diferentes fuentes de datos, el procedimiento de la predicción es similar y está compuesta de dos pasos principales: el cálculo de CMV y el cálculo de imagen a radiación. Del primer paso, algunos píxeles en la imagen son seleccionados como indicadores actual y futuros de la radiación solar. En un segundo paso, la imagen es pocesada para calcular el característico que es luego traducido a radiación. Un componente adicional se usa en ambos algoritmos de predicción: un modelo de cielo despejado. El estado del arte y principales partes de los modelos es presentada aquí:

2.2.1 CMV

El objetivo de este algoritmo es calcular el campo de vectores que representa el desplazamiento en algunas imágenes consecutivas. Es estado del arte abarca principalmente dos técnicas: los algoritmos PIV (Mori and Chang, 2003) y Optical Flow (Horn and Schunck, 1981). A pesar de sus diferentes enfoques, comparten algunas similitudes: la resolución del campo de vectores resultante depende de la configuración de la ventana de búsqueda, existen muchos parámetros, están basados en imágenes consecutivas monocanal, cualquier parametrización puede no obtener un resultado "válido" y la calidad alcanzada también depende de las propias imágenes.

El algoritmo PIV calcula el desplazamiento de una sub-imagen en la imagen anterior por la mejor correlación entre ella misma y varias sub-imágenes vecinas en la imagen actual. Una implementación libre de este algoritmo se usa en este trabajo, llamada OpenPIV.

Existen diferentes mejoras desde el primer algoritmo *optical flow*. Una popular versión de este algoritmo fue propuesta por Lucas et al. (1981). Este algoritmo está basado en la premisa de que la intensidad

de los mismos píxeles en imágenes consecutivas es casi constante. En años recientes, otras tecnologías han sido aplicadas sobre este campo de investigación y, como resultado, se ha lanzado el algoritmo DeepFlow (Weinzaepfel et al., 2013). Las ventajas más importantes de este algoritmo son: (1) sin una parametrización particular, ha demostrado ser robusto cuando se aplica sobre imágenes muy diferentes, y (2) la resolución resultante es la máxima disponible.

2.2.2 Modelos de cielo despejado

A lo largo de las últimas décadas, se han desarrollado varios modelos de cielo depejado con el propósito de estimar con precisión la radiación solar que alcanza la superficie terrestre bajo condiciones de cielo depejado. Las diferencias de rendimientos aparecen por la capacidad de manejar algunas características atmosféricas. Por ejemplo, el modelo Atlas Europeo de Radiación Solar (ESRA) (Rigollier et al., 2000) es usado junto con el parámetro mensual de la turbiedad de Linke (Remund et al., 2003). Este modelo de cielo depejado se usa como parte del método Heliosat-2 (Rigollier et al., 2004). Otros parámetros atmosféricos como el vapor de agua o la profuncidad óptica del aerosol se usan en modelos más elaborados (Gueymard, 2004; Ineichen, 2008). El reciente modelo de cielo despejado McClear (Lefèvre et al., 2013), que utiliza un modelo de transferencia radiativa para inferir la radiación solar, ha sido incluido como parte del recién lanzado método Heliosat-4 (Qu et al., 2017).

2.2.3 Cámara de cielo

Una cámara de cielo es produce imágenes del cielo. Con el propósito de capturar la bóveda celeste completa en una sola imagen, estas cámaras incluyen una lente de ojo de pez, que introduce una esperada distorsión en las imágenes producidas. Dado que se necesita que ciertas áreas estén representadas con precisión en la imagen, la distorsión inherente debe ser tenida en cuenta. Para ello, dos enfoques principales deben ser señalados, estando ambos basados en la relación entre posiciones conocidas en el mundo real y sus respresentaciones en la imágenes. Marquez and Coimbra (2013) sugirió inferir esta distorsión a través de la posición teórica del sol y su representación en las capturas. Como resultado, una transformación puede ser aplicada

a las imágenes brutas con el objetivo de proyectarlas. Alternativamente, los parámetros intrínsecos de la cámara podrían ser inferidos a través de imágenes de control (Gauchet et al., 2012) donde se captura un objeto conocido con características especiales. Un conocido recurso es un tablero de ajedrez, donde se conocen las distancias de las esquinas interiores. Con los parámetros intrínsecos inferidos, se puede aplicar una transformación de las imágenes brutas con una proyección similar a la alcanzada por Marquez and Coimbra (2013). A pesar del hecho de que los parámetros instrínsecos podrían ser útiles para otross propósitos, en los trabajos aquí presentados, se emplea el enfoque propuesto por Marquez and Coimbra (2013) debido a que el problema de la georreferenciación, i.e. la geo-localización de cada pixel en la imagen, es realizada de forma intuitiva. Para este fin, se requiere conocer la altura de la nube mediante, por ejemplo, un ceilómetro.

Una de las tareas más importantes y complejas en el algoritmo de predicción con cámara de nubes es la segmentación de la imagen para discernir entre píxeles nubosos y libres de nube. Este método no es dependiente de la etapa de corrección de la distorsión y su dificultad es manejar una casuística compleja, por ejemplo, cielos cubiertos pero azulados, píxeles quemados, gotas, destellos, nubes rojizas del amanecer/atardecer, etc. La fuente de datos es limitada, siendo la más sencilla los canales rojo, verde y azul de una imagen común. El estado del arte incluye varios enfoques, desde simples técnicas de umbrales en imágenes monocanal, un conjunto elaborado de umbrales (Kazantzidis et al., 2012) aplicados sobre la imagen bruta, librerías de cielo despejado (Chow et al., 2011), etc. Sin embargo, ninguno de ellos puede ser exportado a una cámara de cielo diferente sin cambiar su parametrización. Debido a su enfoque de extrapolación, el método propuesto por Li et al. (2011) es aquí usado como referencia. Una vez configurado, su rendimiento ha sido valorado satisfactoriamente. Sin embargo, algunos algoritmos internos podrían no ser demasiado intuitivos. Así que, este obstáculo es aquí sustituido por un simple algoritmo de clasificación aunque manteniendo el enfoque general.

Enfoque previos de paso de pixel a irradiancia estaban completamente basados en la propia imagen. Sin embargo, en la versión más posterior, este paso es diferente para la predicción de la DNI y GHI, pero en ambos casos, el cálculo de la fracción nubosa (CF) es requerido. Para calcular la predicción de DNI, la premisa es que la componente solar DNI es producida justo por aquellos rayos solares

que provienen directamente del sol. Así, la predicción de DNI se calcula mediante el producto de la actual/futura estimación de DNI de cielo despejado y el complementario de la actual/futura fracción nubosa. Esta última se evalúa sobre la localización de interés en la imagen y con la misma forma del sol. Por otro lado, la predicción de GHI a un cierto horizonte temporal, se basa también en la distribución en dos grupos del K_c de los últimos 30 minutos (Urquhart et al., 2013). El K_c es el cociente entre la GHI medida sobre la estimación de cielo despejado. Una vez que las distribuciones son calculadas, el actual/futuro CF previamente calculado es usado para seleccionar el grupo más adecuado. Luego, el centroide del grupo de K_c seleccionado se usa para inferir la GHI predicha del mismo modo que en el caso de las predicciones de DNI.

2.2.4 Satélite

La predicción de la radiación solar basada en imágenes de satélite se calcula tradicionalmente por el método Heliosat. Desde las primeras versiones (Cano et al., 1986), se han incluido varias mejoras. El lanzamiento más reciente, el método Heliosat-4 (Qu et al., 2017), incluye el modelo de cielo despejado McClear, que utiliza un algoritmo de transferencia radiativa (Lefèvre et al., 2013). Además del satélite Meteosat Second Generation (MSG) (Schmetz et al., 2002), las se emplean las extracciones del proyecto AVHRR Processing scheme Over cLouds, Land and Ocean (APOLLO) con el propósito de tener en cuenta algunas propiedades de nube como la profundidad óptica de la nube, la cobertura nubosa, el tipo de nube, etc. Sin embargo, debido a la sofisticación del método Heliosat-4 y a la dificultad de manejar todas las fuentes de datos involucradas, se usa aquí un lanzamiento anterior del método Heliosat, el método Heliosat-2 (Rigollier et al., 2004).

Basado en los canales visibles del MSG, la premisa del método Heliosat-2 es que si el albedo aparente es mayor que el albedo terrestre, esto es como consecuencia de las nubes, con alguna excepción tal como las áreas cubiertas de nieve, por ejemplo. Así, se calcula primero el albedo terrestre. Como implementación particular, y siguiendo a Kleissl (2013), la variabilidad diurna es tenida en cuenta en este trabajo. Luego, el albedo terrestre puede ser descontado del albedo aparente, y el resultado es normalizado utilizando el albedo producido por la nube más reflectiva. El resultado alcanzado es el

denominado índice nuboso (n) y su casi complementario es el denominado índice de claridad (K_c). Así, la GHI se calcula como el producto del K_c por la estimación de GHI de cielo despejado proveniente del modelo ESRA.

Al contrario que en el caso del modelo de predicción de cámara de cielo, donde solamente un vector general es usado en el procedimiento de predicción, el campo completo de CMV es aquí requerido, pero no es directamente usado. Alternativamente, inspirado por Nonnenmacher and Coimbra (2014), hemos desarrollado un algoritmo con el propósito de mejorar la precisión de las predicciones solares basadas en satélite. Con este fin, primero, las líneas de corriente del movimiento atmosférico son derivadas de los CMVs. Segundo, la línea de corriente que cruza la estación de interés se usa para estimar el futuro índice nuboso en el horizonte requerido. Finalmente, la predicción solar es calculada a partir de este futuro índice nuboso.

Para derivar las predicciones de DNI de las imágenes de satélite, hemos usado el método DirIndex (Perez et al., 2002). Nótese que, independientemente de la resolución temporal parametrizada para el procedimiento de predicción, las entradas de GHI tienen un paso temporal de 15 minutos.

2.3 Objetivos

El objetivo general de la tesis es la mejora de los métodos de predicción de la radiación solar a corto plazo, i.e. desde 0 a 6 horas, basados en cámaras de cielo y satélite. La tesis puede ser dividida en tres partes diferentes que no representan el progreso cronológico. En la primera parte, se propone el estado del arte de las tecnologías. En el caso de la cámara de cielo, varios métodos componen el proceso de predicción y todos ellos han sido actualizados durante años, siendo los últimos disponibles presentados en esta primera parte. Además, un uso particular de la cámara de cielo para elaborar un reconocimiento automático del tipo de nube es presentado y evaluado. Por otro lado, se implementa el método Heliosat-2 y se propone la tecnología más actualizada en términos de CMV. Esta última es el resultado de estudios previos sobre comparación y calibración de otras tecnologías populares. Tras el cálculo del CMV, se añade un proceso adicional al procedimiento de predicción: el cálculo de las líneas de corriente.

Este proceso pretende proveer de predicciones de radiación solar más fiables, intuitivas, y rastreables.

La segunda parte es el estudio de los modos particulares de patrones de tiempo que determinan la variabilidad de la radiación solar en la Península Ibérica. El objetivo de este estudio es anticipar el modo de variabilidad de la radiación solar como alternativa a la propia predicción de radiación solar. El método propuesto encierra una técnica de clasificación no supervisada, análisis estadísticos no paramétricos y una particular toma de decisiones sobre el número de grupos.

La última parte abarca la valoración exhaustiva de cuatro modelos de predicción de la radiación solar en el estado del arte con diferentes enfoques. Tras esto, se presenta una comparación de varios enfoques de combinación basados en técnicas de *machine learning*, para escoger el modelo combinado de predicción de la radiación solar más adecuado.

Por lo tanto, los objetivos particulares de este estudio son:

- El desarrollo de técnicas de corrección de la distorsión de imágenes de cámara de cielo.
- El estudio y evaluación de técnicas de segmentación de imagen centradas en reconocimiento de pixel nuboso.
- El desarrollo, calibración y valoración de enfoques de CMV.
- El desarrollo y evaluación de varios procedimientos imagen a radiación para ambas GHI y DNI.
- La implementación y evaluación del método Heliosat, en particular, el método Heliosat-2 teniendo en cuenta la variabilidad diurna. Además, la implementación y evaluación de la variante HRV.
- El desarrollo de un método de líneas de corriente basado en campos CMV con el objetivo de mejorar la precisión de las predicciones basadas en imágenes de satélite.
- El estudio, desarrollo y evaluación de un método para discernir los modos principales de la variabilidad del recurso solar.

- La evaluación exhaustiva de varios modelos de predicción de la radiación solar del estado del arte.

2.4 Resumen de la tesis

Los resultados alcanzados por la finalización de esta tesis pueden mejorar la penetración de la energía solar en la participación eléctrica. Un conocimiento mejorado de la incertidumbre del recurso solar puede proveer herramientas para: (1) facilitar la toma de decisiones de los TSOs, (2) aumentar su potencial en el mercado eléctrico, y (3) facilitar la gestión de plantas energéticas.

Los capítulos 3 - 7 de esta tesis son un conjunto de artículos que están publicados, enviados o en preparación. El resumen de los capítulos son:

Capítulo 3. La comparación de predicciones a muy corto plazo de alta resolución temporal basadas en: 1) un conjunto de tres cámaras de cielo simultáneas, 2) satélite Meteosat Second Generation (MSG), y 3) enfoque de persistencia inteligente, se presenta en este artículo. Se emplean dos conjuntos diferentes de datos de imágenes de satélite, los canales comunes 1 y 2 del MSG, y también el canal HRV. Todos los métodos involucrados en los procedimientos de predicción son detallados. Los resultados alcanzados son evaluados en virtud de diferentes tipos de nube (altocúmulos, cirrocúmulos, cúmulos, estratocúmulos, multi-nube, y la agregación de todos los casos) y componentes de radiación solar (GHI y DNI). En general, el enfoque de Persistencia Inteligente es el modelo que mejor se comporta hasta horizontes tempranos pero luego es superado por: primero cámaras de cielo y segundo satélite hasta un plazo de 60 minutos. El modelo Satélite HRV no es más adecuado que el modelo Satélite estándar.

Capítulo 4. Se presenta una metodología, que pretende ser completamente operacional, para la clasificación automática de nubes basada en el uso sinérgico de una cámara de nubes y un ceilómetro. El algoritmo *random forest* de *machine learning* fue usado para entrenar el clasificador con 19 característicos de entrada: 12 extraídos de las imágenes de cámara de cielo y 7 del ceilómetro.

El método fue desarrollado y probado basándose en un conjunto de 717 imágenes recogidas en la estación radiométrica de la Universidad de Jaén (España). Hasta nueve tipos diferentes de nubes (más cielo despejado) fueron considerados (cielo despejado, cúmulos, estratocúmulos, nimbostratos, altocúmulos, altostratos, estratos, cirrocúmulos, cirrostratos, y cirros) más una categoría adicional de multinube, para tener en cuenta los frecuentes casos en los que el cielo está cubierto por varios tipos de nube. Un total de ocho experimentos fueron llevados a cabo: (1) excluir/incluir la información del ceilómtro; (2) incluir/excluir la categoría de multinube y (3) usando seis o nueve tipos diferentes de nube, a parte de la categoría de cielo despejado y multinube. El método arroja precisiones que van desde 45% hasta 78%, siendo muy dependiente del uso de la información del ceilómetro. Esta información mostró ser particularmente importante para clasificar con precisión nubes "cumuliformes" y para tener en cuenta la categoría multinube. A este respecto, solamente la información de la cámara mostró no ser adecuada para tratar con esta categoría. Finalmente, mientras que el uso del ceilómetro produjo un rendimiento general superior, se hallaron algunas limitaciones, principalmente relacionadas con la clasificación de nubes con una altura de base de la nube y espesor geométrico similares.

Capítulo 5. Los modos intra-diarios de variabilidad del recurso solar en la Península Ibérica, sus patrones de tiempo asociados y su impacto en la energía solar producida son evaluados en este trabajo. El análisis es realizado para la variabilidad anual y estacional. Primero, los modos de variabilidad son identificados por medio de un análisis de clasificación jerárgico. Se ha calculado con dos años de datos medidos de irradiancia global horizontal (GHI) e irradiancia directa normal (DNI) recolectados en cuatro estaciones. Notablemente, se usan estadísticos de tres horas que describen la media y la variabilidad de la radiación solar como entrada del análisis de clasificación. Segundo, se valoran los patrones de tiempo sinóptico asociados con cada grupo resultante del análisis de clasificación usando datos de presión a nivel del mar y nubosidad. Finalmente, se evalúa la producción energética solar PV asociada con cada modo. El análisis anual revela la existencia de cuatro modos de variabilidad del recurso solar en el área de estudio. Los cu-

atro modos muestran tener un patrón de tiempo distintivo y también un impacto específico en la generación de energía solar en el área de estudio. Los análisis estacionales muestran resultados similares al análisis anual, pero con marcadas diferencias estacionales.

Capítulo 6. Se evalúa la habilidad de cuatro modelos para producir predicciones a corto plazo (hasta 6 horas en adelante) de GHI y DNI en la Península Ibérica basada en dos años de datos recogidos en cuatro estaciones. Los modelos siguen (mayormente) enfoques independientes. Un modelo puramente estadístico (Persistencia Inteligente), un modelo basado en CMV derivado de imágenes de satélite (Satélite), un modelo NWP (WRF-Solar) y un modelo híbrido satélite-NWP (CIADCast). En general, los resultados muestran que Persistencia Inteligente es el mejor en los primeros pasos temporales, los modelos advectivos (Satélite y CIADCast) en los intermedios y el WRF-Solar al final del periodo de predicción. El punto de sobrepaso entre los modelos advectivos y WRF-Solar varía entre 1 y 3 horas para GHI y 3 y 5 horas para DNI. Sin embargo, un análisis detallado muestra enormes diferencias entre el rendimiento de los modelos relacionado con 1) las condiciones locales geográficas y topográficas de las estaciones de evaluación; 2) la variable valorada (GHI contra DNI); y 3) el cielo y las condiciones de tiempo sinóptico sobre el área de estudio. Dependiendo de la estación y el horizonte temporal, los valores de rRMSE se mueven desde 25% hasta 70% para GHI y desde 35% a 100% para DNI. Para las mismas estaciones y horizontes temporales, los valores de rRMSE para DNI son entre un 50% y un 100% más altos que en su contraparte correspondiente de GHI. Dependiendo del patrón sinóptico, los valores de rRMSE son de alrededor del 10/20% para GHI/DNI (horizonte de 3 horas, durante condiciones de alta presión) hasta alrededor de 80/180% para GHI/DNI (durante condiciones de baja presión). Todos los modelos muestran un rendimiento pobre en una estación costera, atribuído a una falta de habilidad para predecir nubes asociadas a brisas mar-tierra. Para concluir, ningún modelo individual pueba ser el modelo que mejor se comporte y, por lo tanto, los resultados muestran que los cuatro modelos son, de algún modo, complementarios. Las ventajas alcanzadas por esta complementariedad son exploradas a fondo en un artículo

compañero (Parte 2).

Capítulo 7. En este artículo exploramos la combinación de cuatro modelos (Satélite, WRF-Solar, Persistencia Inteligente y CIADCast) estudiados en la Parte 1 mediante máquinas de vectores de soporte para mejorar las predicciones de GHI y DNI. Dos enfoques de combinación que usan los cuatro modelos como indicadores han sido estudiados: el enfoque por horizontes construye un modelo combinado diferente para cada horizonte, mientras que el enfoque general entrena un modelo único válido para todos los horizontes. La influencia sobre los modelos combinados al añadir información sobre los tipos de tiempo es también estudiada. Los enfoques han sido evaluados en las mismas cuatro estaciones de la Península Ibérica de la Parte 1. Los enfoques combinatorios han sido extendidos a un contexto regional para obtener predicciones regionales mejoradas. En general, los resultados muestran que la combinación supera con creces los indicadores individuales, sin ninguna gran diferencia entre los propios enfoques de combinación. Los enfoques por horizonte fueron más adecuados para minimizar el rRMSE y los enfoques generales funcionan mejor para rMAE. La mejora relativa en rRMSE obtenida por la combinación de modelos fue hasta un 17% para GHI (16% para DNI), y hasta un 15% para rMAE. Se observaron mejoras similares en las predicciones regionales. Un análisis de rendimiento dependiente del horizonte muestra que mientras las ventajas de los modelos combinados para GHI se mantienen más o menos constantes a lo largo de los horzontes, tiende a incrementarse con el horizonte para la DNI, con las mayores mejoras ocurriendo a las 6 horas. El conocimiento de las condiciones de tiempo ayudó a mejorar un poco más las predicciones (hasta un 3%), pero sólo en algunas localizaciones y para rRMSE.

CHAPTER 3

Comparing sky-camera versus satellite solar radiation nowcasts

Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Fernández-León, M. M., López-Cuesta M., Pamos-Ureña, M. Á., Tovar-Pescador, J., and Pozo-Vázquez, D. (2019) Comparing sky camera versus satellite solar radiation nowcasts. *In preparation*

3.1 Introduction

A considerable effort has been made along the last decades to make the solar energy a real alternative to the conventional energy generation system. There are two main technologies, Solar Thermal electricity (STE) and solar photovoltaic (PV), and many countries have already reached a notable solar share in their energy mixes. Moreover, an important growth is expected in the near future (International Energy Agency, 2018).

Contrary to conventional generation, solar electricity generation is conditioned by weather, thus being highly intermittent. Aside from the deterministic daily and yearly solar cycle, transient clouds and

aerosol intermittency lead to a considerable variability of the solar power plants yield on a wide range of temporal scales. This imposes serious issues regarding the solar power plant management and their yield integration into the electricity grid (Brouwer et al., 2014; Zhang et al., 2015a; Ela et al., 2017). 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, the use of solar radiation forecasts is the only plausible way to mitigate the intermittency. Solar forecasts are used both for the management of solar power plants and to ease the solar energy grid integration (Lorenz and Heinemann, 2012; Law et al., 2014; Zhang et al., 2015b; Antonanzas et al., 2017; Ramírez and Vindel, 2017). For these purposes, solar forecasts are needed in time horizons ranging from minutes and hours to several days ahead. The science of solar radiation forecasting has emerged as a key research topic in the solar energy community in the last years, and many different solar forecasting procedures have been proposed. Nevertheless, the current accuracy of these forecasts is still limited. 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).

Solar radiation forecasts are derived from a wide range of procedures that produce forecasts at different time and spatial scales. Solar radiation methods are often classified according to the forecasting horizon, which is related to the application requirements in which the forecast is used. The forecasting horizon determines the most appropriate methodology to be used. Although there is not a clear agreement, solar forecast methods can be classified into three forecasting horizon: nowcasting (mostly related to the first hour horizon), short-term forecasting (which accounts for forecasts up to 6 hours ahead) and forecasting (which is usually associated with day ahead forecasts based on numerical weather prediction models). The techniques and procedures used in the alternative horizons, as well as, their spatial and temporal resolutions are essentially different (see for a review (Diagne et al., 2013; Inman et al., 2013)).

Nowcasting accounts for methods aiming to provide minutes ahead forecasts, with very high spatial and temporal resolution. The reference methods to obtain these forecasts are mainly based on all-sky cameras image processing (Urquhart et al., 2013; Kazantzidis et al., 2017). Sky-camera nowcasts are obtained by consecutive images ob-

tained with a sky-camera. The comparison of these images allows to derive the cloud motion vector (CMV) of the clouds (previously recognized). Given the CMV field, the future position of the clouds can be estimated by the simple advection (frozen cloud hypothesis). From this estimation, and given a clear-sky model, GHI and DNI forecast are obtained. Typically, temporal resolution of these forecasts is about 1 minute, and spatial resolution, despite the fact that it can be arbitrary, it is usually in the order of meters (Kuhn et al., 2017, 2018). Although Sky-camera based nowcasts have received a notable attention in the last decade (Marquez and Coimbra, 2013; Quesada-Ruiz et al., 2014; West et al., 2014; Alonso-Montesinos and Batlles, 2015; Schmidt, Thomas et al., 2017; Kuhn et al., 2018), their reliabilities are relatively low (Schmidt, Thomas et al., 2017), because the process used to derive these nowcasts is prone to many uncertainties (Nouri et al., 2019). Forecast horizon of the sky-camera based methods depends on the type of clouds, velocity of displacement and the cloud-base height. Sky-camera based methods provide very high spatial (order of meters) and temporal resolutions (order of minutes). Also the latency of the forecasts is very high, in the order of 1 minute.

Short-term prediction is associated with forecasting horizons of several hours ahead (typically up to 6 hours) and reference methods are based on satellite imagery processing (Lorenz and Heinemann, 2012; Perez et al., 2013; Blanc et al., 2017; Ramírez and Vindel, 2017). In a similar procedure to that used for the sky-camera nowcasts, consecutive cloud index (CI) (Rigollier et al., 2004) images derived from geostationary satellite are used to derived the CMV. Then the underlying atmospheric flow is estimated to finally provide the cloudiness and solar radiation forecasts (Kühnert et al., 2013; Perez and Hoff, 2013a; Arbizu-Barrena et al., 2017). Theoretically, the forecast horizon can be extended beyond 6 hours, but in practice, after 4 to 6 hours the forecast skill of this method decreases bellow the skill of other approaches. Satellite short-term forecasts have typical spatial resolution of a few kilometers and a latency of 15 minutes.

Although the sky-cameras and satellite-based forecasting procedures are qualitatively similar, the characteristics of the forecasts show significant differences. These differences are related with the spatial and temporal resolution of the images and their spatial coverage. Thus, sky-cameras are able to provide solar forecast maps with high spatial (order of meters) and temporal (minute or less) resolution (Kuhn et al., 2018). Latency of the forecasts is related with the

time step used for the sky images, usually about 1 minute in case of sky-cameras. Nevertheless the temporal horizon of its forecasts is limited to less than 15 minutes in most cases. This limitation is related with the scarce field of view of the sky-cameras, the cloud base height and the cloud speed.

On the other hand, satellite-based predictions have a greater forecasting horizon. In addition, satellites can provide forecast over extensive areas, not being just limited to the location where the camera is located. Due to the fact that satellite-based predictions are based on geostationary satellite images, their nominal resolution is that of the satellite. In the case of MSG (Schmetz et al., 2002) channels 1 and 2, used in the Heliosat-2 method (Rigollier et al., 2004), the spatial resolution is of $3 \cdot 3$ km at zenith. In addition, the latency of the forecasts is 15 minutes, i.e. the time step of the MSG satellite.

To conclude, sky-camera and satellite forecasts show some advantages and drawbacks which are worth to compare. There are two main difficulties to perform such comparison.

First, there is a forecasts non-overlapping issue. As mentioned, sky-camera horizon is limited to a few minutes most of the time. Therefore, most the cases, sky-camera forecasts do not reach the first "natural" time horizon of the satellite-based forecasts (owing to its acquisition time), which usually is 15 minutes, making the comparison impossible to be performed. Some attempts have been made in order to extend the sky-camera nowcasts horizon. For instance, Peng et al. (2015) used a set of three co-located sky-cameras to obtain a merged image. This allowed to extend the forecasting horizon of GHI nowcast up to 15 minutes. Nevertheless, the relative position of the three sky-cameras and the validation stations did not allow to obtain a larger forecasting horizon.

Second, there is a lack of coordination issue. While sky-camera forecasts are issued when required, being commonly every minute, satellite forecasts are always released every 15 minutes, as maximum. There has been some attempt to provide high temporal and spatial resolution satellite derived nowcasts. For instance, Sirch et al. (2017), used images of the High Resolution Channel (HRV) of MSG satellite to derive DNI nowcasts.

The main purpose of this work is to evaluate the relative performance of sky-camera and satellite-based GHI and DNI predictions.

The assessment is conducted at a station located at southern Spain, using a set of days covering different sky conditions and reaching up-to-120 minutes ahead horizons. To this end, and as previous step, a procedure is proposed to obtain sky-camera-based forecasts. The methodology is based on the synergetic use of a set of three co-located sky-cameras. On the other hand, the MSG HRV channel images are used to obtain satellite predictions. This channel, provides images of 1 km spatial resolution at zenith which are used to provide forecasts on time horizons ranging from 1 to 120 minutes. For the sake of comparison, the standard satellite forecasts, based on channels 1 and 2 of the MSG and with a spatial resolution of 5 km before projecting tasks, were also obtained. Finally, the sky-camera and the two satellite approaches predictions are compared. A smart persistence model is also evaluated for benchmarking purposes. Results are assessed in the light of the sky conditions.

So far, and to the best of our knowledge, no direct comparison between sky-camera-based and satellite-based forecasts has been attempted. Nevertheless, in a recent work Scolari et al. (2018) proposed a method to derived GHI estimates from sky-cameras. There, the estimates were benchmarked against the satellite estimates (Heliosat-2). Results showed that the sky-camera method outperformed the satellite method, with a RMSE relative improvement between 20 and 45%. The sky-camera proved to be superior, particularly during partly cloudy conditions, when the satellite lack the spatial and temporal resolution to capture local changes in the cloudiness. The development of accurate solar nowcasts may contribute to the improvement in the efficiency and dispatchability of both STE (Alonso-Montesinos et al., 2019; Dersch et al., 2019) and PV plants (Schmidt, Thomas et al., 2017). Mainly, by anticipating the short-term fluctuations causing ramps or by optimizing the management and size of the storage systems. This ease the participation of the solar energy in the intra-day electricity markets (Zhang et al., 2015b).

This paper is organized as follows: section 3.2 describes the ground and remote sensing dataset used. Section 3.3 introduces the four forecasting models and the assessment procedure. In section 3.4 the results are presented and discussed and, finally, in section 3.5 a summary of the main results and some conclusions are provided.

3.2 Experimental setup and database

The dataset used in this study contains data collected at several time intervals covering from 6 June 2015 to 17 October 2015. A specific cloud type rules the sky during each time interval. The different cloud types available in the dataset are: altocumulus, cirrocumulus, cumulus, estratocumulus, and an extra class called multi-cloud, where different cloud types are captured together. Each observation of the dataset contains the following features: ground truth of GHI and DNI collected by irradiance sensors, 3 simultaneous total sky images from different cameras, and cloud base height (CBH) from a ceilometer. The arrangement of the ground-truth instrumentation is showed in Figure 3.1a. Additionally, 15-min time resolution images from the Meteosat Second Generation satellite are also gathered. The next subsections describe each feature of the dataset.

3.2.1 Irradiance sensors

One-minute time resolution GHI and DNI measures are collected from a Eppley Black & White, Model 8-48 pyranometer and a WMO Secondary Standard / ISO Highest Precision Eppley NIP pyrheliometer, respectively. Raw data are quality-check analyzed according to the procedures defined by Long and Dutton (2010), related to physically possible and extremely rare limits. Data labelled as any filter are deleted. In addition, raw data of a solar zenith angle $>75^{\circ}$ are also discarded.

3.2.2 All-sky imager

A set of three commercial low-cost video-surveillance cameras are used as all-sky imagers to simultaneously provide 1-minute captures of the sky. Raw output files are JPEG RGB images of 2592 pixels width and 1944 pixels height. The set of cameras are deployed throughout the outer boundary of the cylinder-parabolic CSP platforms. The distances between sky imagers are 1, 1.84, and 2.4 km. In addition, the distances from the cameras to the irradiance evaluation station are 3.4, 4.2, and 3.6 km, as showed in Figure 3.1a.

3.2.3 Ceilometer

Fifteen-second raw measures are collected from a Jenoptik CHM 15k Nimbus ceilometer configured to discern up to three different layers. In order to paliate the fluctuation of the raw CBH signal and avoid misestimations and clear sky values, it is applied the same averaging method as in Huertas-Tato et al. (2017a) with a moving average of the last ten min. As a result 1-minute estimations of CBH is obtained.

3.2.4 Satellite retrievals

The instrument SEVIRI, which is on board the Meteosat Second Generation satellite, produces raw images every 15 minutes. From these images, only the two visible channels and the high resolution visible (HRV) channel are used in this study. These channels are transformed to fit with the azimuthal equidistant projection, thus obtaining single-layer images. As a result, over latitude 0°, longitude 0°, the visible-channels images have a spatial resolution of 5 x 5 km, while the HRV-channel image has a spatial resolution of 1 x 1 km. With the aim of deriving clearness index (K_c) estimates, the Heliosat-2 method (Rigollier et al., 2004) is applied over the two former channels. In addition, a variant of the Heliosat-2 method is also applied to the latter channel, with the proper adjusts.

3.3 Methods

3.3.1 Camera methods

Camera setup and image distortion correction estimation

With the aim of allocating some areas of real world within images, the relative position of the sky imagers has to be known. Therefore, it is needed to know the coordinates and orientation of the camera regarding the cardinal points. In addition, it is required to set the sky imager facing the zenith, thus being installed completely in parallel to the ground surface.

The fish-eye len included in the all-sky imagers, needed to capture

the whole celestial dome, introduces two particular features. Firstly, the raw images contain valid information just on an area around the center of the camera-sensor (hereafter called camera center), which could differ from the center of the image. Secondly, this valid information is captured with a mostly radial barrel distortion. Thus, the pixels close to camera center represents less real area of the sky than those close to the edges, as described at the top of Figure 3.1b.

There are several approaches available in order to infer the inherent raw-image radial barrel distortion, which are based on the relation between the known position of elements in the real world and the position of these elements in the captures. In some studies, it is used, for instance, the inner captured corners of a known chessboard to stablish this relation (Gauchet et al., 2012). However, following Marquez and Coimbra (2013), the radial barrel distortion is inferred using the theoretical sun zenith angle and the distance between camera center and sun center position in the image. As a result, a scatter plot of observations is modelized using a polynomial adjust. With the aim of obtaining a more detailed model, days when the sun path in terms of zenith angle is longer are selected as the case study. Thus, this radial distortion might be corrected applying a polynomial transformation to the position of raw pixels.

As may be inferred from section 3.2.2, neither coordinates nor orientation of the three cameras are the same. Due to fact that the three all-sky imager are of the same version, the polynomial distortion model is shared between them. However, due potential differences in the mechanic ensembles, the camera center and the area of useful information in the output raw-images is different.

Image georeferenciation

The purpose of the georeferentiation method is twofold: to crop a portion of the sky representation by the distortion-correction polynomial transformation of the raw image and to attach a size estimation for each pixel (lenght/pixel ratio).

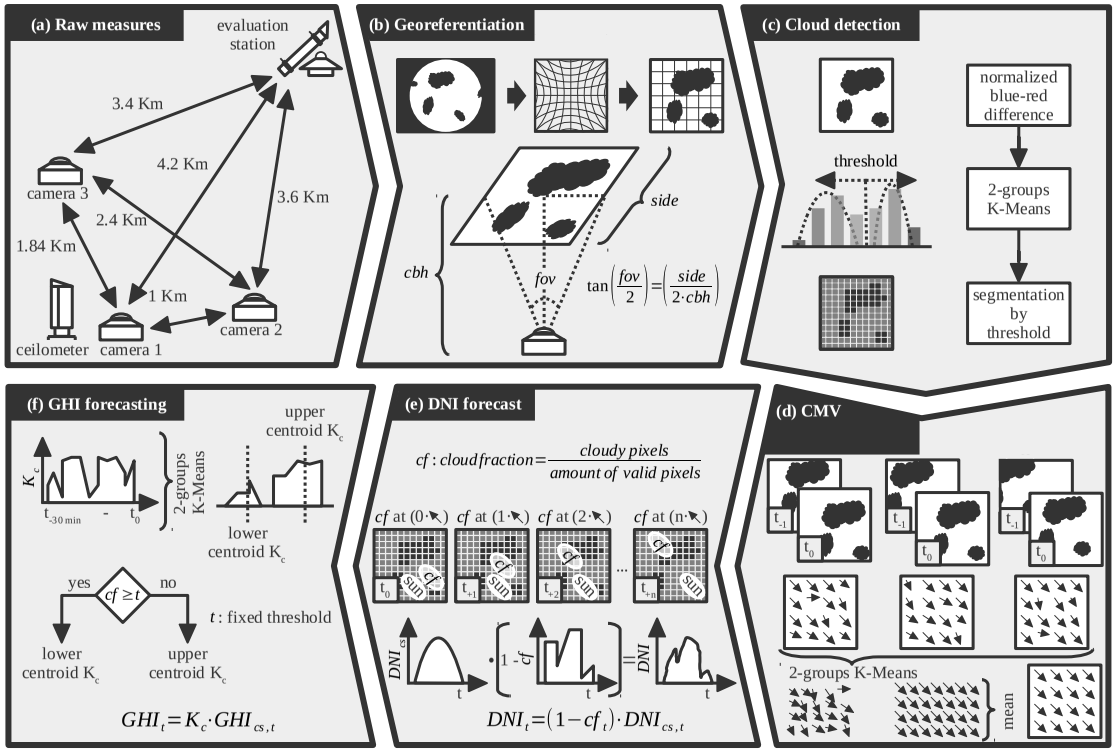

Figure 3.1: Whole camera forecasting procedure split in 6 chronologically ordered steps: (a) raw measures gathering, (b) image georeferentiations, (c) cloud detections by image segmentation, (d) CMV field computations, and finally, (e) DNI and (f) GHI forecasting.

In the raw image, each pixel is linked to a radius-zenith angle position, being the radius the distance between each pixel and the camera center. With the aim of correcting the radial distortion, this radius is transformed by applying the model resulting from the distortion correction estimation, explained in secton 3.3.1. This procedure, graphically described at the top of Figure 3.1b, has to deal with two main limitations. Firstly, due to the radial distortion, the accuracy of the distortion polynomial model decrease with the distance from the camera center. Secondly, due to the own geometry, objects in the sky, mainly clouds, might hide the information behind them. This effect increase near the horizon where, for instance, gaps between clouds might not be showed in broken sky conditions. In order to paliate these limitations, the distortion correction is only applied for a maximum field of view (FOV) angle of 130° centered at the zenith, in this case. Thus, the distortion-corrected image contains just part of the information of the raw-image, being of a square shape with an requested number of pixels.

The lenght/pixel ratio calculation does not depend on the raw-image distortion-correction step. To this end, trigonometric relations, as showed at the bottom of Figure 3.1b, apply over a right-angle triangle where two of the vertexes are: the camera and the zenith at a particular height, where the right-angle is allocated. The remain vertex (hereafter called sky boundary) is defined by the half of FOV angle from the camera. This information is transfered to the distortion-corrected image, posed at a proposed height and where the zenith is represented at the camera center. There, each pixel equally represents the double of the segment defined by the zenith and the sky boundary vertexes. Thus, the lenght/pixel ratio depends of the proposed CBH estimation.

Cloud detection

As a result of this method, a decision mask reports the type of each pixel in the input image, being one in: cloudy pixel, clear sky pixel or null value, as described at the bottom of Figure 3.1c. The latter is reserved for burned pixels, null areas, and situations where the distinction between cloudy and clear sky pixel might be computationally unavailable.

This method is based on that proposed by Li et al. (2011). In

addition to particular thresholds, two main adjustments are included: Firstly, a normalized blue-red difference score is used instead of the normalized blue/red ratio. Secondly, the bimodal branch is solved applying a 2-groups K-Means clustering algorithm instead of by the minimum cross entropy thresholding algorithm. Due to the cloud type approach of this study, all of the images are processed by the bimodal branch. The method is own-pixel dependent and applies over those pixels which are not labbelled *a priori* as null values. The overall method is graphically described at Figure 3.1c.

Cloud motion vector for all-sky camera images

Cloud motion vector (CMV) methods infer the bi-dimensional displacement through the comparison of two consecutive images (Figure 3.1d). There are several approaches available but only two of them are highlighted in this study: PIV-based and optical-flow-based algorithms. PIV methods (Mori and Chang, 2003) are based in block matching techniques between two single-channel images. In the other hand, there are several optical-flow-based approaches, being the Deep Flow (Weinzaepfel et al., 2013) an alternative development which combines the approach of large displacement optical flow (Brox and Malik, 2011) with a matching algorithm. As a result, each vector within the resulting CMV field represents the displacement of suffered on that pixel in the input images. Deep Flow approach is here used to derive the CMV through the greyscale distortion-corrected images from each all-sky camera. Thus, for each observation, three CMV fields are computed. However, due to the relatively limited area of the sky covered by a camera, only a global vector is requested. It is computed as the average vector from the most populated group resulting from a 2-groups K-Means clustering algorithm, as showed in Figure 3.1d.

GHI and DNI camera nowcasting

For each observation in the dataset, 120-min lead-time 1-min resolution predictions of GHI and DNI are computed. Both GHI and DNI nowcasting computations are determined by two shared components: clear sky estimations and the cloud fraction (CF) observed at a certain area within an distortion-corrected cloud-detection all-

sky image, where the evaluation station is located. The clear-sky estimations are calculated using the European Solar Radiation Atlas (ESRA) clear-sky irradiation model (Rigollier et al., 2000) complemented by the monthly Linke turbidity parameter (Remund et al., 2003). In the other hand, the CF at the requested area is computed as the ration between the number of cloudy pixel over the total amount of valid pixels (Figure 3.1e). This area is solar-cycle-dependent and it is defined by three features: location, size and shape.

The location is determined by three parameters, being all of them attached to areas within the image: the sun position, the relative location of the evaluation station from the camera, and the forecasting horizon. The sun position locates the camera, in other words, the observer. A pixel only affects the camera if it is somewhere in between the ray from it to the sun. This principle also applies for all of the pixels. Despite the fact that the relative position of the evaluation station from the camera is fixed, the amount of pixels of the required displacement to allocate it, changes according to the lenght/pixel ratio of the image, therefore, according to the CBH. Finally, at the same time when the image is captured, the area of interest is located just at the evaluation station. However, this location is corrected at forecasting horizons according to future sun positions and oncoming to the global vector of displacement (Figure 3.1e).

The size and shape are those of the sun representation and, therefore, only depends on the solar cycle. The size of the sun representation changes along the day in both raw and distortion-corrected images. However, the shape of the sun representation is always nearly round at the raw images but it changes at the distortion-correction ones, as a consequence of the own used algorithm to compute them.

As showed in Figure 3.1e, DNI forecasts are directly derived as the product of clear-sky DNI estimations and (1 - CFs). Nevertheless the GHI prediction at each forecasting horizon is based on the representative K_c of the latest 30 min (Urquhart et al., 2013). The overall method that produces GHI predictions is graphically described in Figure 3.1f. The K_c is computed as the ratio between a ground truth GHI measurement over the relevant clear-sky GHI estimation. Once the latest 30 min K_cs are computed for each forecasting horizon, a 2-groups K-Means clustering algorithm is applied. The center of one of the resulting groups is selected as the representative K_c. If the CF is greater or equal than a particular threshold, the center of the lower

K_c group is selected and vice-versa. A value of 0.4 is here used as CF threshold. Finally, a GHI prediction is computed as the product of the representative K_c and the relevant clear-sky GHI estimation.

The set of all-sky cameras produces individual predictions. Furthermore, an additional average GHI and DNI forecast is computed from the three individual ones. These average predictions are calculated through the available sources of information at each forecasting horizon. The predictions from the three all-sky cameras are not always available as a consequence of the size of the sky representation, which is determined by the CBH, and the displacement direction detected. Therefore, the evaluation station might not be visible by all of the all-sky cameras at the same time. This fact also could produce the forecasting horizon enhacement through the concatenation of the individual predictions at certain forecasting horizons. Thus, the average prediction could reach farthest forecasting horizons, and contain gaps or overlaps at certain lead times.

3.3.2 Satellite retrirevals methods

Cloud motion vector for satellite images

A different CMV algoritm is used for each type of satellite K_c image. However, it is applied the same principle described in Figure 3.1d. In particular, the OpenPIV algorithm (Liberzon et al., 2009) with a search window square size of 41 pixel is used to derive the CMV field from the satellite visible-based K_c images. Consequently, a vector within resulting CMV field represents the displacement of a bigger area. After the own OpenPIV algorithm, some filters are applied in order to avoid some suspicious vectors. In the other hand, the Deep Flow approach (Weinzaepfel et al., 2013) is used to derive CMV from satellite HRV-based K_c images. Similarly, as in section 3.3.1, a displacement vector is attached to every pixel within in satellite HRV K_c image. Due to the fact that the represented areas are wider than in case of 3.3.1, it is more suitable as much information as possible, instead of a global vector. Indeed, it is requested to consequently fill null vectors, if any, produced as a result of OpenPIV algorithm or by the subsequent filters.

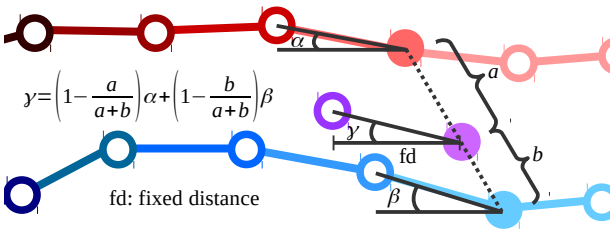

Figure 3.2: Cycle within the whole computation of the streamline of interest through two neighbour basis streamlines.

Streamline method

Inspired by Nonnenmacher and Coimbra (2014) and using as input a static CMV field at each forecasting event, the particular oncoming streamline crossing over the evaluation station, hereinafter called the streamline of interest, is computed. A streamline is here defined by consecutive pixels within an image, and attached to a velocity. In other words, a streamline contains the information about the displacement flow and the time needed to cross over each pixel. Thus, it is inferred the pixel that will cross over the evaluation station at a certain forecasting horizon. This method is here applied with the aim of obtaining K_c forecasting at 6 h lead-time with a time resolution of 1 min, according to section 3.3.1.

This method, which is graphically described in Figure 3.2, has two main steps: the basis streamlines computation and the calculation of the streamline of interest. The former step is obtained through *matplotlib.pyplot.streamplot* Python software and produces the basis for the latter. At the second step, the streamline of interest is computed as a result of a weighted average of the surrounding basis streamlines, that may be two or just one. Each element within the streamline of interest is iteratively obtained.

In order to compute the streamline of interest, the process begins at the pixel of the evaluation station and has to be computed, at least, the oncoming streamline of interest. As showed in Figure 3.2, at each cycle, two reference points are selected, being the nearest element from a neighbour basis streamline and the other streamline element in the opposite direction, if any. Two tasks have to be done: the associated velocity computation and the proposal of the next item of the streamline of interest. In case of HRV-based K_c images, the

velocity is directly computed as the module of the CMV field vector it that position. The reference points determine the location of the following item of the streamline of interest. This location is computed using a fixed distance and the weighted average of the slopes between the reference points and their next items. The distances from the reference points are used as weight, being the lower the heaviest and vice-versa. In case of satellite visible-based K_c images, a weighted average with these same weights is also applied to infer the velocity through that associated to the reference points. This cycle is repeated with the proposed item and so on, until an image boundary is reached or the proposed item crashes with another streamline, including the own streamline of interest. As a result of this method, the K_c at each forecasting horizon is obtained.

GHI and DNI satellite nowcasting

Once the K_c estimates are calculated, the own Heliosat-2 method indicates the GHI estimates computation as the product of K_c and the relevant clear-sky GHI estimates produced by ESRA clear-sky irradiation model (Rigollier et al., 2000) and tuned by the monthly turbity of Linke parameters (Remund et al., 2003). The same proposal of Heliosat-2 method is also applied with the aim of obtaining the predictions, therefore, using forecasted values of K_c and GHI clear-sky estimates. Thus, 6 h lead-time 1-min time-resolution GHI predictions are calculated.

DNI predictions are computed applying the DirIndex method (Perez et al., 2002), using the GHI predictions as input. However, each 1-min time resulution DNI forecast is obtained with its 15-min neighbour GHI forecats. Thus, for instance, 15-min DNI prediction is computed through GHI forecasts at lead-times of: 0-min, 15-min, 30-min etc. In the other hand, 16-min DNI prediction is computed trough GHI forecasts at lead-times of: 1-min, 16-min, 31-min, etc.

Despite the fact that 1-min time resolution GHI and DNI forecasts are produced, the minimum time interval between satellite prediction events is 15 min, due to the satellite retrievals updating time, as explained in section 3.2.4. In order to produce satellite predictions matching those of all-sky camera based forecasts every minute, the 14 1-min updating time predictions between two consecutive satellite-based raw forecats are filled with relevant information from the previ-

ous one, like in a operation scheme. Therefore, 1-min time resolution satellite prediction at a particular time slot covers from t_0 to t_{6h}. One min later, when no raw satellite prediction is available, a filled satellite forecast is produced. In the latter, the same forecasting horizons are covered but with an extraction of the previous prediction from t_{1min} to t_{6h} and filling the last empty position with a null value. This procedure is repeated until the 1-min gaps between two consecutive raw satellite predictions are filled.

3.3.3 Smart Persistence method

Persistence radiation forecasting models use a particular current feature in forecasting horizons. Smart Persistence model uses as a fixed feature the ratio between a ground-truth measurement over the relevant clear-sky estimate. At forecasting horizons, the predictions are computed as the product of this fixed ratio and the relevant clear-sky future estimates.

$$I(t) = \frac{I_0}{I_{clear-sky}(t_0)} \cdot I_{clear-sky}(t), \tag{3.1}$$

where I_0 is the current ground truth measurement, $I_{clear-sky}(t_0)$ represents the current relevant clear-sky estimate, and $I_{clear-sky}$ is the clear-sky radiation estimate at forecasting horizon t. I is valid to represents both GHI and DNI. This method applies for both GHI and DNI forecats. The Smart Persistence model is used as reference model to be outperformed by the other sources of prediction.

3.3.4 Evaluation procedure

For each forecasting horizon, GHI and DNI ground-truth measurements and the relevant 1-min updating-time predictions of every forecasting model are gathered only where all of the sources of information are available. Thus, all forecasting horizon-based error metrics are computed with the same number of samples. In order to describe the models performance, the root mean square error (RMSE) and the so-called forecast skill in terms of RMSE (FS_{RMSE}) (Coimbra et al., 2013) error metrics are used. The later is based on the ratio between

the RMSE of a particular forecasting model over that of Smart Persistence model.

$$\mathrm{RMSE}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right)^2} \tag{3.2}$$

$$\mathrm{FS}_{\mathrm{RMSE}}(t) = 1 - \frac{\mathrm{RMSE}_{\mathrm{forecast}}(t)}{\mathrm{RMSE}_{\mathrm{SmartPersistence}}(t)}, \tag{3.3}$$

where I stands for both GHI and DNI and N is the total amount of availables samples.

When a error metric score is assessed, the forecast from four models are compared being those based on: all-sky cameras, satellite, satellite HRV and Smart Persistence. In case of all-sky camera forecasts, just the score of the mean prediction is showed, complemented with a forecast spread descriptor. The spread is defined by the minimum and maximum prediction score produced by the Cam mean model in addition to the other three all-sky camera based individual forecasting models.

3.4 Results

This section describes the forecasting models assessment for the different cloud types analyzed. With the aim of inferring the representativeness of the results, the Table 3.1 shows the number of samples taken into account in order to compute the error metrics at some significant forecasting horizons, depending on the cloud type assessed. The right most column labelled as 'All samples', shows the aggregation of all of the samples in other cloud types.

Table 3.1: Number of samples at significant forecasting horizons depending on cloud type.

	Altocumulus	Cirrocumulus	Cumulus	Stratocumulus	Multi-cloud	All samples
0 min	590	78	264	277	1040	2249
5 min	599	78	329	317	1151	2474
10 min	583	77	401	269	1026	2356
15 min	451	55	367	226	676	1775
30 min	342	48	216	129	231	966
45 min	226	45	183	84	198	736
1 h	142	41	183	51	169	586
1 h 15 min	135	37	176	29	136	513
1 h 30 min	134	27	176	20	105	462
1 h 45 min	125	23	171	17	93	429
2 h	6	3	11	1	5	26

As showed in Table 3.1, the cirrocumulus clouds are the poorest represented cloud type, due to two main reasons: (1) this type of cloud are very uncommon in this area, and (2) it it unusual to capture the whole celestial dome with just this type of clouds. On the contrary, the multi-cloud type is much more usual than the others. Despite the fact that the study of a particular cloud type ruling the sky may allow to describe key issues of forecasting models, in an operational environment, the multi-cloud skies might be the most important issue. Under image-based prediction model approaches, the decreasing number of samples with the forecasting horizon growth is a common feature when predictions are assessed owing to two main factors: (1) the pixel of interest is out of the bounds and/or (2) the forecasting horizon belongs to night time, when the prediction assessment is not too relevant. However, Table 3.1 shows that the most represented forecasting horizon ranges from 0 to 15 min. This effect is produced by the set of cameras, where the evaluation station area is usually within the three simultaneous captures in the earliest forecasting horizons, but not neccessarily from the beginning. As expected, the number of samples dramatically falls at the lattest lead-times. Therefore, the prediction assessment is only applied up to 90 min forecasts.

Figure 3.3 shows the up-to-90-min GHI forecasting errors in terms of RMSE by the analyzed prediction models by cloud type. Thus, the Figure 3.3a, 3.3b, 3.3c, 3.3d, 3.3e, and 3.3f represents the result reached under altocumulus, cirrocumulus, cumulus, stratocumulus, and multi-cloud skies, respectively. Figure 3.3f describes the assessment for all of the cases. In each sub-figure, the outcomes reached by satellite, satellite HRV and smart persistence models are represented by the turquise, garnet and orange lines. In addition, purple color is used to describe the performance of camera model. The dark purple line is used to represent the camera mean prediction results and the light purple shadow describes its uncertainty by means of the minimum-maximum range reached by the four camera-based sources: the three single-camera ones plus the camera mean prediction model. Based on the scarce availability of samples at latter forecasting horizons showed in Table 3.1, the maximum lead time assessed is 90 min.

As expected, a common feature in all of the cloud type analyses is the better performance of smart persistence model at the first(s) forecasting horizon(s). This is a consequence of the own smart persistence proposed approach, that ensures no error at evaluation time

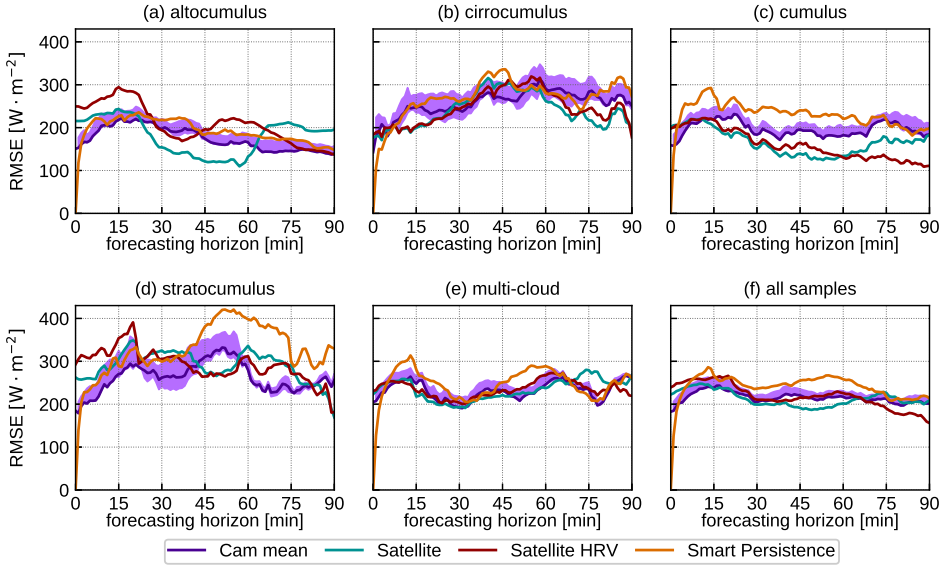

Figure 3.3: Up-to-90-min GHI RMSE forecasting errors by cloud type. Purple, turquoise, garnet and orange lines represent the forecasting errors produced by camera, satellite, satellite HRV, and smart persistence prediction models, respectively. The light purple shadow describe the uncertainty of camera-based model. Sub-figures a, b, c, d, e, and f show the assessment for altocumulus, cirrocumulus, cumulus, stratocumulus, multi-cloud, and all-samples categories, respectively.

(horizon of 0 min). In addition, depending on the sky cloud regime variability, the prediction at evaluation time could be almost the suitablest a few minutes later. However, in the majority of analyzed cloud types, the smart persistence model is outperformed at earlier lead times. In the most repeated scheme, the smart persistence is the best performing model up to a few minutes. Then, the camera model outperforms smart persistance, thus creating the first intersection. After that, a second intersection is created when the satellite model outperforms the camera and this situations is holded until an hour, approximately. From one hour on, the scheme is more variable. However, each cloud type analysis situates the intersections at different horizons. For instance, in case of altocumulus (Figure 3.3a), the first intersection is located at less than 5 min and the second one after 20 min. In case of cumulus cloud type (Figure 3.3c), the second intersection is moved since 10 min, approximately. The case of stratocumulus (Figure 3.3d) is particularly beneficial for camera model, because it is the best performing model from less than 5 min to about 40 min. After that, the satellite HRV provides the best forecasts up to less than 1 h. Despite the fact that the multi-cloud scheme (Figure 3.3e) matches the regular scheme proposed, there is almost no opportunity to camera model, and after the smart persistence, satellite model provides the better predictions until the first h, even though all of the models perform similarly. Figure 3.3e shows the all samples analysis, where the second intersection is located around 15 m. Contrarily to the proposed regular scheme, in the cirrocumulus cloud analysis (Figure 3.3b), the first intersection is at about 7 m but it is done by the satellite, instead of camera model. Then, in horizons ranging from 30 to 60 m, the camera model is a little more suitables than the other models. Finally, the camera model is showed to describe particularly hight uncertainty in case of cirrocumulus and stratocumulus cloud analyses (Figures 3.3 b and d, respectively.)

Similarly to Figure 3.3, Figure 3.4 shows the GHI FS_{RMSE} forecasting errors. Due to the fact that some features, e.g. the intersections, are the same that those showed in Figure 3.3, the scope of this figure is to describe the outperforming ratio over the smart persistence model.

In this figure, it is more clearly showed that the camera mean prediction model is, in the majority of cases, the best performing algorithm between all of analyzed camera-based models. However, its overall performing ratio does not exceed the value 0.25 in up-to-

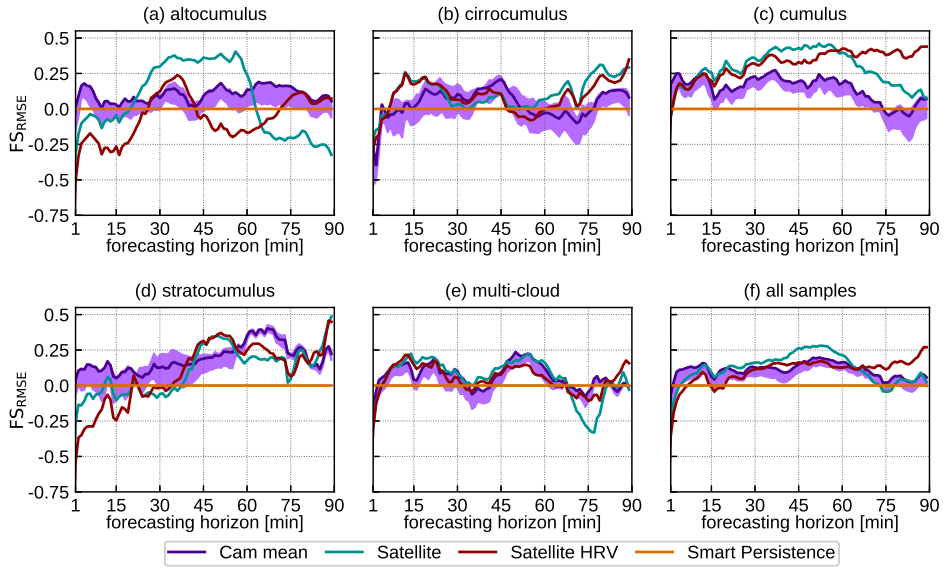

Figure 3.4: As Figure 3.3 but for FS_{RMSE} score.

60-min forecasting horizonts. Also in this lead time range, a similar maximum yield is reached by satellite model in case of cirrocumulus (Figure 3.4b) and multi-cloud (Figure 3.4e). In the rest of cloud analyses, the satellite yield is more vigorous. This feature is more evident in case of cumulus clouds (Figure 3.4c), where the outerforming ratio almost reaches a value of 0.5. In all of the cases, the satellite HRV model yields is similar or lower than that of satellite model.

Figure 3.5 is as Figure 3.3 but accounts for DNI prediction errors. As expected, the RMSE values are higher than in case of GHI analysis. In addition, the smart persistence model usually keeps as the best performing model up to slighly farest horizons than in the GHI counterparts. The regular situation described in GHI analyses only appears in case of DNI when all samples are taken into account (3.5f). There, the camera mean prediction model slightly outperforms the other models since 5 min to 13 min, approximately. After that, and up to more than 1 h lead-time, the satellite model performs better.

In case of altocumulus clouds (Figure 3.5a) the camera mean preciction model probes to be the best performing model in time horizons ranging from less than 5 min to about 25 min. After that, the satellite model performs as the best but only up to 45 min lead time. Since this point on, the more sophisticated proposed algorithms can not outperform smart persistence. Figure 3.5b (cirrocumulus) shows

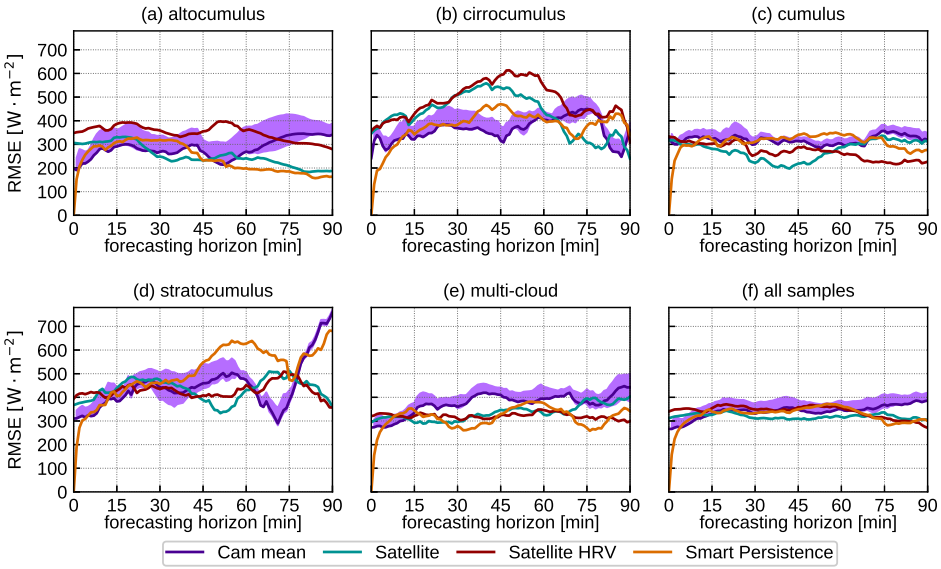

Figure 3.5: As Figure 3.3 but for DNI.

that camera men prediction model better performs since ten minutes to 1 h, approximately. Therefore, no choice for satellite or satellite HRV model in this horizon range. On the contrary, the satellite or satellite HRV model is the best performing model from 10 min on, in case of cumulus clouds (3.5c). In case of stratocumulus clouds (3.5d) the smart persistence model clearly performs better during the firsts 5 min, approximately, and even is competitive to up to about 12 min. In the other hand, in the range of 40 min to 1 h lead times, the satellite model perfoms better. In the mean times, the best performing model is changing between camera mean prediction and satellite HRV model. The multicloud analysis (3.5e) reveals that camera mean prediction model is competitive since 5 to 10 min, approximately. After that, and up to 1h lead-time, the satellite is the best performing model except from about 25 to 45 min forecasting horizons, where the smart persistence performs better. The uncertainty of camera mean prediction model is higher in cases of cirrocumulus (3.5b) and stratocumulus (3.5d) clouds, as in Figure 3.3. Nevertheless, the DNI analysis shows that the highest uncertainty of camera models is here represented by the altocumulus cloud (3.5a) analysis.

Similar evaluations may be observed in Figure 3.6. Despite the fact that overall relative yields of forecasting models are lower than in case of their GHI counterparts, some features may be highlighted.

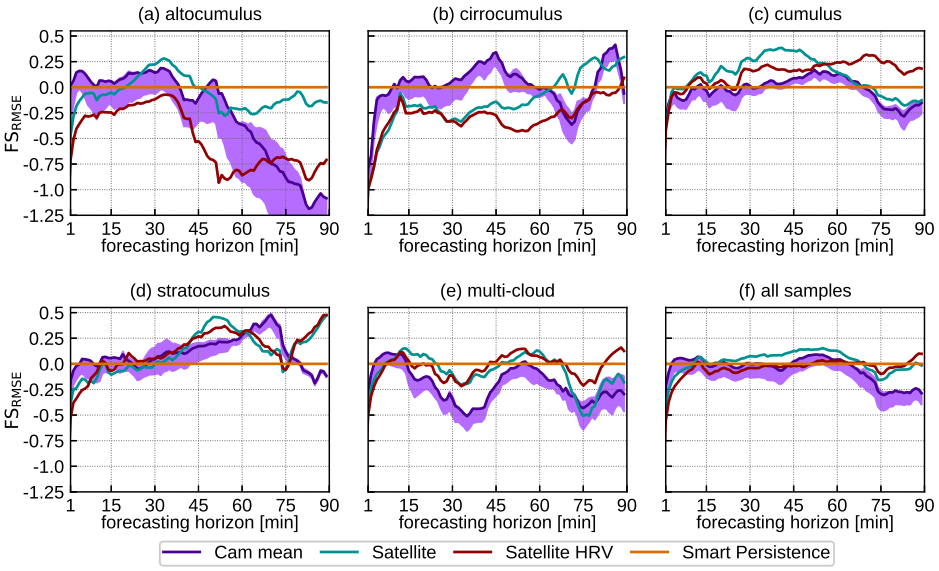

Figure 3.6: As Figure 3.4 but for DNI.

The camera mean prediction model play a key role in case of altocumulus and cirrocumulus clouds (Figure 3.6a and b, respectively). In particular, it is the only forecasting model able to outperform smart persistence in the first hour in case of the latter. The satellite model is usually shows skill along the first hour but it is exceptionally suitable in case of cumulus and stratocumulus clouds (Figure 3.6c and d). Despite the fact that in the latter the maximum ratio is reached (with a value near than 0.5), in the former, the performance is the most vigorous.

3.5 Summary and conclussions

The assessment of 4 nowscasting models is proposed in this study. Two of them are based on satellite and satellite HRV imagery. Other prediction model is based on the combination of 3 all-sky imagers and the last one is a smart persistence approach. The evaluation is conducted independently by GHI and DNI. A different assessment is made by each cloud type, discerning between altocumulus, cirrocumulus, cumulus, stratocumulus, and multi-cloud. An additional analysis which takes into account all of the available observations is included.

The performance of persistence models use to decrease when the forecasting horizon increases but it is very challenging in the early horizons. In addition, the smart persistence approach used in this study yield no error at evaluation time. Due to the fact that the nowcast maximum horizon could be alocated at 2 h, the performance of this smart persistence approach in the firsts minutes is very difficult to be overcame. Moreover, the smart persistence algorithm might be improved using a more accurate clear sky model which takes into account some "current" parameters of the atmosphere (e.g. aerosol optical depth, water vapour, etc.).

The 3 all-sky cameras produce subsequent predictions and an additional mean prediction is also computed. Alternatively to camera single-model concatenation, the mean prediction model is only computed when the three simultaneous single forecasts are available. Thus, in addition to the own solar radiation forecasting (mean prediction), the three sources let to suggest an incentainty range. As expected, in the majority of cases (i.e. GHI and DNI, forecasting horizon), the best performing camera model in terms of RMSE and, therefore, FS_{RMSE}, is that reached by the mean prediction model. The suitability of this model differs between GHI and DNI and also between each cloud type. In case of GHI prediction, the suitability of camera mean prediction model is allocated between the first minutes and a maximun lead time ranging from 7 min (multi-cloud) and 40 min (stratocumulus). Exceptionally, this model is not suitable for cirrocumulus clouds. On the contrary, the performance of this model in case of DNI is very competitive in up to 1 h horizons. Despite the fact that camera mean prediction model has no skill when cumulus clouds rules the sky, in the remain cases, this model show to be suitable for a maximum forecast horizon of 10 min, at least. There are four main step in the camera forecasting schedule: georeferentitation, cloud detection, CMV computation and GHI/DNI nowcasting. Despite the fact that improvements could be reached in each step (inaccuracies), maybe two of this steps provide the main sources of errors: the cloud detection and GHI/DNI nowcasting. The approach of the latters(s) is inspired in statistical and physical relations. An improvement in the performance of this step may mean a change of the own method approach. However, the cloud detection method might have the greatest impact on the later methods, and thus, in the final result. The method is fully dependent on the image and it is strongly affected by sun, which enhances the brightness of the

surrounding pixels, reaching a big area withing the projected image. In addition, some of the wide variaty of skies are even unmanaged, like the bluish or reddish clouds. Moreover, the k-means clustering algorithm satifies the requirement of quickly classifying an elevated numbers of elements (all of the pixels in an image), but its approach might not be the suitablest. Finally, the own binary result of valid pixels strongly determines the forecasts. Therefore, an elevated number of intermediate results might be used as a proxy of clearness index in order to tune the forecasts. This proxy of clearness index might be even modified depending on the cloud type.

In the other hand, both satellite-based models perform similarly but the satellite HRV algorithm is showed to be slighly worse, in general. In addition, its time comsumption is higher. Therefore, under the streamline approach, the satellite model might be quicker and better. Satellite model has an implicit blur if it is compared to satellite HRV algoritm and, like in other studies, this might be the cause of its better relative performance. With the aim of predicting GHI up to 1 h, the satellite model is showed to be the suitablest from early forecasting horizons approximately ranging from 7 to 22 min in case of cirrocumulus and altocumulus, respectively, in general. Exceptionally, this model starts to be useful at 40 min on in case of stratocumulus, as well as in case of DNI prediction. The satellite model is regularly showed to the suitablest for DNI up-to-1h forecasting from lead time horizons approximately ranging from 10 to 27 min in case of cumulus and altocumulus, respectively. Exceptionally, satellite model has no skill in case of cirrocumulus during the first hour. Similarly to the case of camera mean prediction model, some improvements might be reached in the main steps of the satellite-based algorithms (Heliosat-2 method, CMV computation, streamlines calculation and satellite-to-irradiation algorithm). However, some of the improvements may mean a change in the approaches. The update of the own Heliosat method to the current version (Heliosat-4) may be a feasible improvement, despite of its development sophistication. As may be expected, the so-called frozen clouds issue is a handicap. This issue is also presented in camera mean prediction model but it has a limited impatc. Mainly in case of DNI prediction, the parallax effect, which is not taken into account in this study, has an impact in the performance of satellite-based models. Moreover, the own satellite estimations (at evaluation time) have important deviations and, therefore, the forecasts are thus affected. With the aim of paliating

these inaccuracies, ground measures might be used in order to tune the predictions.

Despite the fact that prediction for a single station is not a limitation by itself, it might be required a regional forecasting (i.e. covering an area instead of a sigle pixel). This approach would limit the use of the streamline algorithm in operational environment due to the time consumption, because each streamline is computed individually. Alternatively, other approaches can be applied, like a pixel-by-pixel displacement approach. However, the sophistication of the model might be a reason to alternatively assimilate the clearness index into a numerical weather prediction model.

Acknowledgements

The authors are in debt with EUMETSAT for providing the MSG data used in this study. The Spanish Ministry of Economy and Competitiveness has supported this study by granting the projects ENE2014-56126-C2-1-R and ENE2014-56126-C2-2-R (http://prosol.uc3m.es). In addition, the team from the University of Jaen are also supported by the Junta de Andalucía (Research group TEP-220) and also by FEDER funds. The authors thanks all of the support received.

CHAPTER 4

Automatic cloud type classification based on the combined use of a sky camera and a ceilometer

Huertas-Tato, J., Rodríguez Benítez, F. J., Arbizu-Barrena, C., Aler-Mur, R., Galvan-León, I., and Pozo-Vázquez, D. (2017). Automatic cloud type classification based on the combined use of a sky camera and a ceilometer. *Journal of Geophysical Research: Atmospheres*, 122(20):11,045-11,061. 2017JD027131. DOI: `https://doi.org/10.1002/2017JD027131`

4.1 Introduction

Scientific interest in retrieving cloud information dates many decades back, and was mainly related to civil and military aviation. The attention to cloud information has grown in the framework of climate studies, since clouds play a key role in earth energy balance (Li et al., 2014; Wild et al., 2013). More recently, in the field of weather forecasting, the improvement in cloud representation has emerged as a significant research field. Mainly, because clouds are involved in multiple and strong interactions, and their misrepresentation may have large impacts and implications in the atmospheric dynamics and,

then, in the accuracy of the simulations of the numerical weather prediction models (Haiden et al., 2015; Pincus et al., 2011). Lastly, the growing penetration of the solar energy around the world has fostered a great interest in cloud information, since clouds are the main source of variability of the solar energy (Martínez-Chico et al., 2011; Mateos et al., 2014; Tzoumanikas et al., 2016). In all the previous fields of science, the establishment of proper, accurate and cheap cloud monitoring systems is crucial (World Weather Research Programme. WWRP/WGNE Joint Working Group on Forecast Verification Research (JWGFVR), 2012). Accurate and consistent cloud observations, which are globally standardized, remain an important need (World Meteorological Organization. International Cloud Atlas, 2017). Nevertheless, the type of cloud information (cloud parameters, temporal and spatial resolution, etc.) needed, greatly varies depending on the application. In some of the above mentioned applications, information about the type of cloud is crucial. Human-reported information was the first available continuous source of information on cloud type. But the high associated cost, the low accuracy and issues such as the representativeness, makes this source of information under menace in many countries. The use of satellite retrieval for cloud classification is a promising tool, because of their spatial coverage. For instance the "cloud-type" product of EUMETSAT Satellite Application Facility on Climate Monitoring or the APOLLO project (Kriebel et al., 2003; Wey and Schroedter-Homscheidt, 2014) reports operationally a coarse cloud classification. Nevertheless, the performance of these operational cloud-type monitoring systems is still limited due to limitations of the satellite platforms.

The other alternative is the use of ground-based sky cameras systems. These systems, which basically date one decade back, are now considered the reference for cloud cover estimates (Boers et al., 2010; Cazorla et al., 2008; Long et al., 2006). More recently, the automatic recognition of cloud types has emerged as possible product of these instruments.

There are two basic steps for automatic cloud classification. Firstly, the extraction from the camera images of appropriate and distinctive information of the different sky conditions and cloud types. To this end, different features can be computed on the information from camera channels. Particularly, these features account for characteristics such as cloud shape, texture, or the color of the sky/clouds. Second, once a set of distinctive features are obtained, cloud classification

relies on the use of automatic classification algorithms. Ultimately, these algorithms are trained and tested with human-supervised cloud-type databases.

The type and number of features has increased enormously in the last years, benefiting from other fields of research, such as automatic pattern recognition. For instance Calbó and Sabburg (2008) used texture properties and the Fourier Transform of the camera visible channels to classify up to 8 classes of sky conditions. The methodology achieved an accuracy of about 62%. Heinle et al. (2010) proposed the use of a combined set of textural and color features for the classification of up to 7 cloud types, with a classification success rate of about 75%. Rum (2013) proposed the use of features from the infrared channels of a camera, obtaining an accuracy of 90% in the estimation of towering cumulus and cumulonimbus cloud types. Kazantzidis et al. (2012) proposed the use of a multi-color criterion on sky images, showing an average performance of about 87% using seven cloud categories. Kliangsuwan and Heednacram (2015) used a new methodology, based on the fast Fourier transform, for feature extraction for cloud classification. The overall accuracy of this methodology was shown to be 90% for the automatic classification of 7 clouds types. Wacker et al. (2015) used, as ancillary information for cloud classification, the measured long wave radiation. They reported an improvement of up to 10%, compared to the use of just the sky camera information. The reported mean accuracy ranged from 80 to 90%. Cheng and Yu (2015) have proposed a cloud classification method based on division of the image in different blocks. In this way the authors were able to account for mixed clouds types in one image, obtaining an improved classification accuracy. Recently, Li et al. (2016) used a novel approach for cloud-type recognition, based on the analysis of image as a collection of patches, rather than a collection of pixels. The method showed an accuracy of 90% for five classes of sky conditions.

Regarding classification machine learning algorithms, the literature contains proposals ranging from artificial neural networks (Kliangsuwan and Heednacram, 2015; Lee et al., 1990; Singh and Glennen, 2005), to k-nearest neighbor (KNN) (Cheng and Yu, 2015; Heinle et al., 2010; Kazantzidis et al., 2012; Wacker et al., 2015) and support vector machines (SVM) (Schmidt et al., 2016; Taravat et al., 2015; Zhen et al., 2015). ANNs are a commonly machine learning technique used in cloud classification. It is actually a non-linear re-

gression technique that can be used for classification by setting a threshold on the output(s). Typically, standard architectures with 3 layers are used (input/hidden/output) and in a multiclass classification context, like cloud classification, there are as many output neurons as classes. KNN does not need to fit a model to the data; rather, it stores all data and classifies new instances by looking for the closest stored data instance(s). KNN does not require any adaptation for multiclass problems. The basic version of KNN may suffer more than other methods when there are many features, or some of them are irrelevant. It is also very slow for real use if the data set is large. However, there are methods for KNN that can improve both accuracy (like Weinberger and Saul (2009)) and speed (like kd-trees (Wess et al., 1994)). SVMs aim to maximize the generalization capabilities by finding separation boundaries between classes that maximize the margin. They have fewer local minima issues compared to ANNs, because SVMs solve a constrained convex optimization problem, with a single global optimum. On the other hand, the most common approach to SVM trains binary classifiers, hence requiring to train as many models as classes (one-versus-rest approach) or as many as pairs of classes (one-versus-one approach), although there are also approaches that deal with multiple classes directly (Crammer and Singer, 2001).

To sum up, performance of the different approaches and studies for automatic cloud classification varies greatly and can hardly be compared due to several reasons: different cameras, different cloud data sets, different time representativeness, different experimental setups and evaluation methods and different cloud classes.

Aside from the sky cameras, the use of ceilometers for cloud property retrieval has emerged in the last few years (Illingworth et al., 2007). Ceilometers are single-wavelength low-powered lidars (light detection and ranging) which can provide high-frequency observations of cloud profiles, including parameters such as the cloud base height (CBH), cloud top height (e.g. cloud cover). Unlike satellite imagery, which generally provides low-reliability CBH estimates at relatively low temporal resolution (very few samples per hour), ceilometers are able to provide an accurate description of the location of the cloud vertical boundaries with even several samples per minute (Arbizu-Barrena et al., 2015; Costa-Surós et al., 2014; Martucci et al., 2010; Viúdez-Mora et al., 2015).

In this work, we propose and evaluate a methodology for automatic cloud classification based on the synergetic use of the information reported by a sky camera and a ceilometer. So far, ceilometer has been not used in automatic cloud classification, so here the added value of this instrument is evaluated. Following recent bibliography, different features were derived from the sky camera images. These features, along with the information reported by the ceilometer, were used as input for a state-of-the-art machine learning classification system: random forests (RFs) Breiman (2001). This is a machine learning technique that has seldom been used for cloud classification, but which is known to be among the best performers in classification tasks, according to some empirical studies (Caruana et al., 2008; Caruana and Niculescu-Mizil, 2006). In a recent work (Cheng and Lin, 2017), RF have been used together with other algorithms (such as SVM and a Bayesian classifier) to develop a voting scheme for classifying each pixel in the image as cloud or non-cloud. This is a related, but different issue as the one addressed in the present paper, where whole images are classified, rather than individual pixels. RF belongs to the ensemble of decision trees family of algorithms. Ensemble techniques build models by training not one, but many different sub-models whose outputs are combined. In RF, randomization techniques are used to build a varied set of sub-models (decision trees). Classification is done by majority voting. RF deals with multi-class problems with no further adaptation. Also, RF training algorithms can easily take advantage of parallel computing.

The methodology is evaluated on a dataset recorded from a camera and a ceilometer located at the radiometric station of the University of Jaén (Spain) over a set of days corresponding to the period 2013-2015. The procedure here proposed aims to mimic a fully operational one. As a consequence, skies with multiple cloud types and layers at the same time are considered and accounted for. In a recent work Wacker et al. (2015) reported this kind of skies to be highly challenging in automatic cloud classification. Three analyses were conducted; the first one to evaluate the role of the ceilometer and the camera information, the second to analyze the performance of the method when skies with several clouds types are included and, finally, to evaluate the performance of the model when using an increased number of cloud types. Evaluation was conducted on the light of the different cloud characteristics and the nature of the camera and ceilometer information.

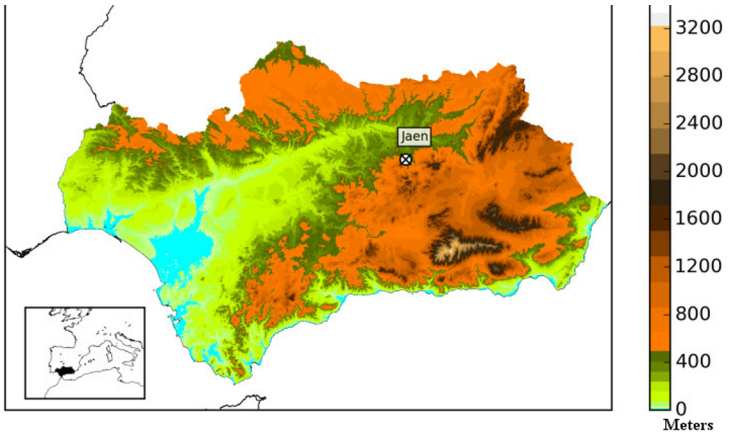

Figure 4.1: Study region and location of the meteorological station at the University of Jaén.

4.2 Data Description

In this section, issues concerning to data used in this work are explained. Particularly, the camera and ceilometer hardware characteristics, data and preprocessing procedures and the different types of clouds used in the classification are described.

4.2.1 Camera and Ceilometer Hardware Description

All the measurements used in this study were collected at the meteorological station of the University of Jaén, Andalucía (southern Spain), at coordinates 37.7877°N and 3.7782°W, and 454 m above mean sea level (Figure 4.1).

A total sky imager model Yesdas TSI-880 and a Jenoptik CHM 15k Nimbus ceilometer were installed in September 2012 (Figure 4.2). The TSI-880 is composed by solid-state CCD pointed downward at a hemispheric mirror which reflects the whole hemisphere (fish-eye vision). Reflection of the Sun is blocked by a dark strip (shadow band), thereby protecting the imager optics. The TSI provides 352 × 288 pixels images every 30 s and has been designed for climate/weather applications, showing to be robust regarding environmental conditions (Long and DeLuisi, 1998). Notably, this camera has been proven

Figure 4.2: Meteorological station of the Univ. of Jaen. Left, Ceilometer and sun tracker in the background. Right, the TSI-880 sky camera.

to be accurate for the estimation of the cloud cover (Boers et al., 2010; Kreuter et al., 2009; Long et al., 2006; Mannstein et al., 2010). In the case of high clouds, it is able to report the sky conditions over a spatial domain of about 38 km $\times$ 38 km (Mannstein et al., 2010). In the last years, this sky camera has been used as reference instrument in solar energy applications (Chow et al., 2011; Martínez-Chico et al., 2011; Quesada-Ruiz et al., 2014).

The Jenoptik CHM 15k nimbus Ceilometer uses laser pulses at wavelength of 1.064 nm, receiving the back-scattered signal over a field of view of 0.45 mrad. This instrument is able to detect up to five cloud layers simultaneously and to provide their altitude with an accuracy of $\pm$5 m, being its vertical cloud detection range from 5 m to 15 km. The sample rate is 15 s. This particular ceilometer is one of the very few ones able to detect clouds above 7.5 km and with less spurious values and better resolution in the upper cloud boundary than other similar instruments (Boers et al., 2010; Martucci et al., 2010; Wiegner et al., 2014).

4.2.2 Sky Camera Images and Ceilometer Data Preprocessing

A total of 717 TSI images, and the corresponding ceilometer estimates, were processed for this study. The images, corresponding to a total of 131 days of the years 2013 to 2015, were selected in order to have a representative sample, with different solar zenithal angles, of the 11 categories described in the following section. Every sample was meant to be representative of 5 min intervals, that is, images of each of the 11 categories were carefully selected to ensure that during the 5 previous minutes period exactly the same category was presented. First, the TSI images were masked in order to highlight the border, buildings and band in the images. Second, the images were projected following Marquez and Coimbra (2013). This procedure transforms the images from a spherical to a rectangular grid. In order to prevent horizon distortion effects, this transformation was conducted only for zenithal angles below 65°, that is, a 130° field of view of the camera. Figure 4.3 shows some examples of the TSI raw and processed images.

The ceilometer reports every 15 s cloud profiles representative of the column at the ceilometer location, namely cloud base height (CBH) and cloud penetration depth (CPD). In this work, up to three different cloud layers were considered. The CPD can be regarded as a proxy of the cloud geometrical thickness. Due to the nature of the clouds (high variability in space and time), ceilometer data should be properly processed in order to provide meaningful information linked to the TSI images, at the 5 min time interval used here. This is particularly relevant for some cloud types, such as cumulus, stratocumulus and cirrocumulus and, in general, cumuliform clouds. These clouds form patches and, therefore, the ceilometer may not report cloud information in some of the 20 samples of the 5 min evaluation period used here. In addition, the ceilometer sometimes provides spurious measures, or out of range values, which are related to the nature of the back scattering signal processed by these instruments. Given these issues, and since the methodology used in this work aims at emulating a fully operational system, ceilometer data were processed to provide meaningful cloud profile information. First, based on the 20 collected ceilometer samples, a number of candidate group of measurements are selected according to the active layers (up to three). Second, clear-sky values were removed from the 20 samples. Then,

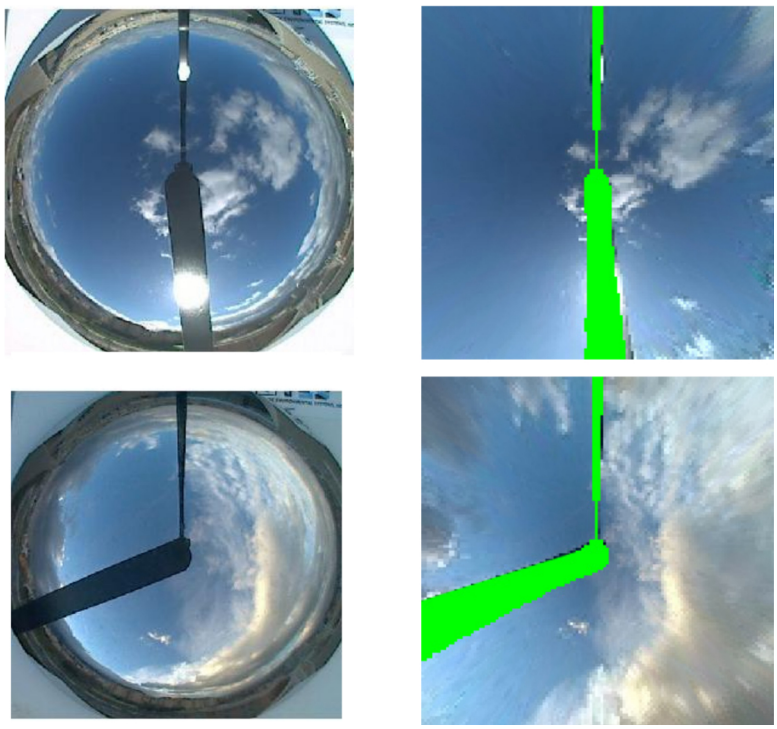

Figure 4.3: Two examples of raw and processed TSI images. Top Left/right raw/processed image corresponding to day 2015-01-22 at 12:27:33 UTC. This image was classified as cumulus cloud according to Table 4.1. Bottom left/right raw/processed image corresponding to the day 2015-02-14 at 17:22:18 UTC. This image was classified containing several cloud layers, that is, multicloud type according to Table 4.1.

based on the CBH values of the remaining samples a cluster analysis was carried out. The number of centroids in this cluster analysis provided the number of cloud layers, up to a maximum of 3. Finally, for each centroid, a mean CBH and CPD were computed, after applying a filter for outliers. If the 20 measurement are reported as clear sky (i.e. no clouds are detected), the ceilometer procedure final output is the presence of "0" cloud layers. Figure 4.4 shows an example of outputs of this processing procedure.

Since this process has been evaluated trying to mimic an operational system, some problems have been found. Particularly, in about 3% of the samples (24 images), the ceilometer reported no cloud information in cases for which the TSI-880 image was classified in some cloud category different from clear sky. A further analysis confirmed that 15 of these cases corresponded to cirrocumulus and cumulus. These clouds, in many cases, do not cover the whole sky dome and may not overpass the ceilometer column with the 5 min window here used. The other nine cases correspond to cloud types such as cirrus and nimbostratus. In these cases, the ceilometer was not able to provide the proper cloud information due to technical issues, reporting a very low detection range.

4.2.3 Sky Conditions and Cloud Classes

The sky images and the ceilometer information were used to manually classify the 717 samples according to classes displayed in Table 4.1. Particularly, two types of classification experiments were conducted. In the first ones, up to seven cloud types were used (first column in Table 4.1). These cloud categories are the most commonly used in the bibliography (Heinle et al., 2010; Kazantzidis et al., 2012), and tries to group cloud types with similar characteristics. In the second one, compound categories are decomposed into the individual cloud types, resulting in 10 cloud types (second column in Table 4.1). In both cases we have added the multicloud category, which aims to represent cases in which the sky is covered by several cloud types at the same time, including the case of several cloud layers. This category is commonly found and should be considered in fully operational systems. The multicloud category has been scarcely addressed in the literature. Wacker et al. (2015) describes the problems for automatic recognition of this category, but no attempt for classification

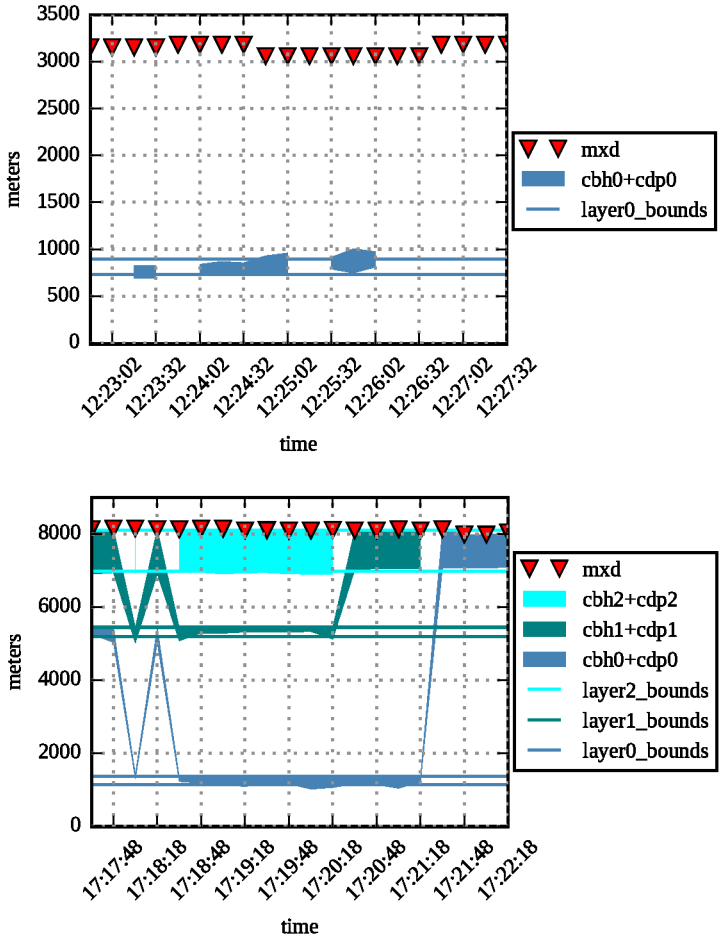

Figure 4.4: Top, Raw and processed ceilometer data corresponding to the image in Figure 4.3 top. The blue shaded area represents the range between the measured cloud base height and this values plus the cloud penetration depth. Values correspond to the 5 previous minutes at which the image was obtained. The bottom straight line shows the final estimate of the cloud base height for this sample, while the difference between the top and bottom straight lines shows the final estimate of the cloud penetration depth. Note that only some measurements were available during the 5 min interval. Triangular points at the top indicate the maximum detection range of the ceilometer for this particular measurement interval. (bottom) As in Figure 4.4, top but for the bottom image in Figure 4.3. Similarly to the previous case, the shaded areas indicate the measurements and the straight lines, the final cloud base height and cloud penetration depth estimates. Note that, in this case, three cloud layers were detected

was made. Only in Li et al. (2016) the multicloud case is considered in an automatic cloud classification procedure. This multicloud category is described as a mix of the sky conditions considered in this work, and covering more than 20% of the sky.

Table 4.1: Categories Used for the Cloud Classification and Main Characteristics Derived From the Ceilometer

Cloud types			CBH	CPD
Seven cloud types + multicloud	10 cloud types + multicloud	Number of images	Mean (SD)	Mean (SD)
Clear-Sky	Clear-Sky (CLS)	48		
Cirrus and Cirrostratus	Cirrus (ci)	131	9086 (1515)	951 (501)
	Cirrostratus (cs)	39	7684 (676)	1829 (422)
Cirrocumulus and Altocumulus	Cirrocumulus (cc)	13	6832 (2023)	469 (238)
	Altocumulus (ac)	75	4494 (2257)	726 (516)
Altostratus and Stratus	Altostratus (as)	57	6701 (1751)	1858 (607)
	Stratus (st)	53	833 (485)	295 (276)
Stratocumulus	Stratocumulus (sc)	49	1358 (372)	275 (126)
Cumulus	Cumulus (cu)	54	1121 (513)	176 (32)
Nimbostratus	Nimbostratus (ns)	42	702 (345)	448 (424)
Multicloud	Multicloud (MC)	156		

Note. In the first experiments, a total of eight classes (first column) were distinguished. In the second one, classes increased to 11 (second column). In both cases, the category multicloud is included, indicating the presence of several layers and/or different cloud types in the same image. The mean and standard deviation (in parenthesis at right), of the CBH and CPD (in meters) are displayed in the last two columns. Values correspond to the whole experimental database.

4.3 Methods and Evaluation

In this section, automatic cloud classification is addressed. First, the features extracted from the images and ceilometer information, to be used as inputs to machine learning algorithm, are described. Next, a short description of Random Forest algorithm is included. Finally, the metrics and procedure used in this work to evaluate the performance of the classifier are also presented.

4.3.1 Features for Cloud Automatic Classification

In this work, we employ a wide set of features (see Table 4.2) as inputs to the cloud classifier. They are divided into two main groups, depending on which instrument was used to compute them: image features (extracted from the ground camera, features 1 to 12) and cloud layer features (extracted from the ceilometer, features 13 to 19).

Table 4.2: Table Listing All of the Features Used as Input to the Classifier

	Feature	Type	Formula
1	μ^r; Red average	Image-spectral	$\mu^{c_1} = \frac{1}{n^2}\sum_{j=0}^{n}\sum_{i=0}^{n} M_{i,j}^{c_1}$
2	μ^b; Blue average	Image-spectral	Same as above
3	σ^b; Blue deviation	Image-spectral	$\sigma^{c_3} = \sqrt{\frac{1}{n^2}\sum_{j=0}^{n}\sum_{i=0}^{n}(M_{i,j}^{c_3} - \mu^{c_3})^2}$
4	γ^b; Blue skewness	Image-spectral	$\gamma^{c_3} = \frac{1}{n^2}\sum_{j=0}^{n}\sum_{i=0}^{n}\left(\frac{M_{i,j}^{c_3} - \mu^{c_3}}{\sigma^c}\right)^3$
5	D^{rg}; Red-green mean difference	Image-spectral	$D^{c_2} = \mu^{c_1} - \mu^{c_2}$
6	D^{rb}; Red-blue difference	Image-spectral	Same as above
7	D^{gb}; Green-blue mean difference	Image-spectral	Same as above
8	EN^b; Blue	Image-textural	$EN^{c_3} = \sum_{j=0}^{g}\sum_{i=0}^{g}\left[p_{ij}^{c_3}\right]^2$
9	ENT^b; Blue	Image-textural	$ENT^{c_3} = \sum_{j=0}^{g}\sum_{i=0}^{g} p_{ij}^{c_3}\log_2 p_{ij}^{c_3}$
10	CON^b; Blue	Image-textural	$CON^{c_3} = \sum_{j=0}^{g}\sum_{i=0}^{g}(i-j)^2 p_{ij}^{c_3}$
11	HOM^b; Blue	Image-textural	$HOM^{c_3} = \sum_{j=0}^{g}\sum_{i=0}^{g}\frac{p_{ij}^{c_3}}{1+\lvert i-j\rvert}$
12	C;% cloud coverage	Image-coverage	$Sat_{i,j} > T; T = 0.41(1)$ $C = \frac{cp}{tp}(2)$
13	h_μ^1; Mean height from layer 1	Ceilometer-height	From CBH layer 1
14	h_μ^2; Mean height from layer 2	Ceilometer-height	From CBH layer 2
15	h_μ^3; Mean height from layer 3	Ceilometer-height	From CBH layer 3
16	t_μ^1; Mean thickness from layer 1	Ceilometer-thickness	From CPD layer 1
17	t_μ^2; Mean thickness from layer 2	Ceilometer-thickness	From CPD layer 2
18	t_μ^3; Mean thickness from layer 3	Ceilometer-thickness	From CPD layer 3
19	l; Present layers	Ceilometer-layers	Number of detected layers

Note. The last column shows how the feature is obtained according to the description in section 4.3.1. c_k, indicates the color channel, using k=1, 2 and 3 fo red, green and blue, respectively; n is the size of the $n \times n$ image and g indicates the gray value of the pixel (g=256 levels).

Features From the Camera

Most of the image features used in this work are based on Heinle et al. (2010) and they have been obtained from the red, green, and blue channels of images. These channels are represented using three matrices M^r, M^g, M^b, red, green, and blue respectively. Each (i, j) location in the matrices corresponds to a pixel in the image, with integer values between *0* and *255*. There are several types of image features: spectral features, textural features, and cloud coverage.

The spectral features (rows from 1 to 7 in Table 4.2) use the color matrix M^c exclusively (where c=r, g, or b), extracting statistical measures directly from it. These are the simplest from the feature set and require very little processing.

The textural features (rows from 8 to 11 in Table 4.2) make use of a grey level co-occurrence matrix (GLCM). This is a transformation over one of the color channels. The result is a $g \times g$ matrix, g being the number of gray levels considered in the image. Thus, every element of the GLCMs in rows from 8 to 11 in Table 4.2 ($p^c_{i,j}$) represents the relative frequency of two adjacent pixel values i and j. c represents the color of the source channel. Here we use *g=256 value levels*. GLCMs represent the relative frequency of two pixel values appearing together in the image, at a given offset (in this case, $x' = x + 1, y' = y + 0$). This matrix is commonly used in image analysis for detecting textures in grey images or in a given color channel, and are supposed to give information on the spatial distribution of color, which spectral features are unable to provide. Textures are relevant in the detection of cloud types. There are several other textural features, as proposed by Haralick et al. (1973). However the four used in the article are the subset of features proposed by Heinle et al. (2010). These measure different properties of the GLCM, and are the following: energy (it measures the homogeneity of grey level differences), entropy (it measures the randomness of grey level differences), contrast (it measures local variation within the grey level matrix), and homogeneity (it measures similarity of adjacent grey levels within the matrix).

Finally, a cloud coverage statistic (row 12 in Table 4.2) is used in the procedure. To obtain this cloud coverage, firstly, the original red-green-blue image was converted to hue-saturation-value (HSV) color space following Smith (1978) and Jayadevan et al. (2015). Hue

describes the color itself, while saturation denotes the degree of difference between a color and gray and value represents the brightness. Saturation ("Sat" in Table 4.2) fits into the range [0, 1], from white, through the grays, to the most colorful hue. In this work, cloudy pixels are detected based on a threshold value *T=0.41* for the saturation value. Pixels (i, j) with a saturation greater or equal than this threshold are detected as clear sky; otherwise the pixels are classified as cloudy. The percentage of sky covered is calculated using the formula labeled as (2) in Table 4.2 (row 12), where *cp* and *tp* are the amount of cloudy pixels and the total amount of pixels, respectively.

Features From the Ceilometer

The ceilometer offers height and thickness information about the cloud type (CBH and CPD), that can help discern differences between similar-looking clouds which would be impossible to recognize otherwise.

Layers in cloud formations are numbered in order of distance from the ground. Layer 1 is the closest to the ground, then layers 2 and 3. Given this, we define six new features (CBH and CPD of each layer) plus an extra feature indicating how many actual layers (out of three) have been detected (rows from 13 to 19 in Table 4.2). We represent the information for each layer as h_μ^n or t_μ^n, to indicate the mean CBH or CPD of layer $n(n = 1, 2, 3)$, and l the number of layers detected. In sum, a total of 7 features were derived from the ceilometer to be used in the automatic classification procedure.

Machine learning algorithms require a fixed number of inputs/features. Therefore, in case the ceilometer returns just information of one or two layers, we fill the missing layers (up to 3) by replicating the information from the closest layer we have information of. In case there are no layers, values are set to an arbitrarily large number, indicating that clouds could not be detected.

4.3.2 Random Forests

For cloud classification, we use the random forest (RF) algorithm presented in Breiman (2001). RF have been reported to be one of the best algorithms for classification (Caruana and Niculescu-Mizil,

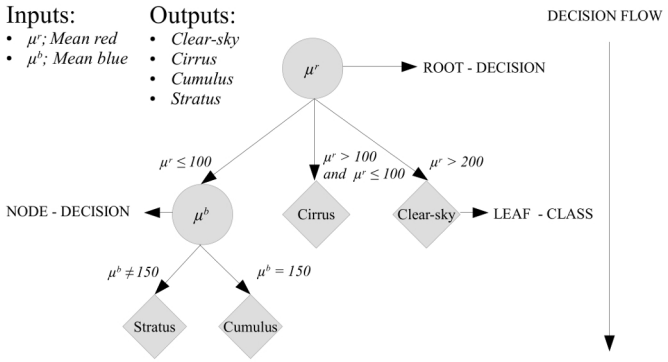

Figure 4.5: Example of decision tree with 2 input features and 4 possible classes. It has a maximum depth of 2, 2 decision nodes and 4 leaf nodes.

2006) and needs no adaptation to work in a multiclass context. This algorithm calculates N submodels (single classification trees) to form an ensemble of models that can predict the class of the given input. Every sub-model is an individual decision tree. A simple example tree is shown in Figure 4.5. To classify an instance, the tree is navigated from the root node to a leaf node. Every nonleaf node contains a decision based on an input feature which will determine the next node to be visited. The tree continues to be navigated through the nodes taking the path that decision nodes determine. Leaf nodes contain labels and, if a leaf node is reached, then the class is determined as the label of the given node.

The RF algorithm constructs multiple different trees from the same training data by means of a double randomization process. First, in order to build each tree, a new dataset with the same size as the training data is obtained by sampling with replacement. Second, instead of considering the whole set of features, each decision node of each tree uses only a random subset of them (*mtry* is the parameter name for the size of this random subset, typically much smaller than the whole set of features). The set of decision trees in the RF ensemble classify new data by majority voting. A diagram of the whole process is represented in Figure 4.6.

Before building the final model, the parameter *mtry* has to be tuned for optimal classification accuracy. This parameter must be within the range (*1, (F-1)*), where F is the total number of features.

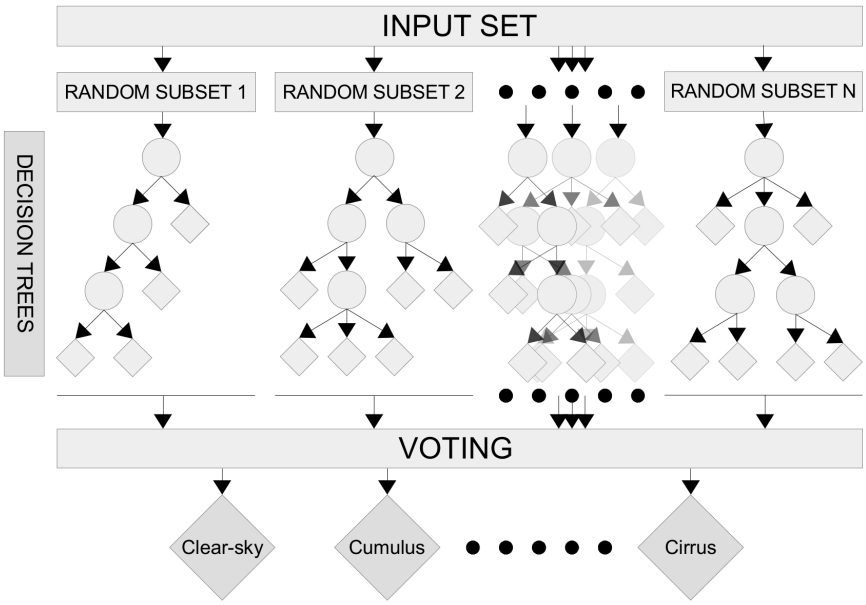

Figure 4.6: Image depicting the construction of a Random Forest ensemble by random resampling and training of several decision trees. Classification is carried out by majority voting among ensemble members.

The optimal F value is obtained by training and testing models with different values and selecting the best performing one. It is important to remark the tuning process uses the training partition only (the test partition is never used for training, parameter tuning being part of that training process). In this article, the RF implementation for R has been used (Liaw et al., 2002). RF has been used together with package *caret*, which is able to deal automatically with parameter tuning (Kuhn et al., 2008).

4.3.3 Evaluation Procedure and Metrics

In order to evaluate the performance of the RF classifier for automatic cloud classification, a cross-validation procedure is carried out. Standard cross-validation divides the available data in P equally sized folds or subsets. Then, for every fold n, a model is trained using all folds but n, and tested with fold n (i.e. a performance measure, such as accuracy, is computed for the trained model on fold n). The final cross-validation estimate is the average of the 10 accuracy values. The standard deviation can also be computed. In this work, we follow the common cross-validation practice and set $P = 10$.

However, applying standard cross-validation to cloud image datasets can be potentially problematic if the dataset contains sequences of images (cloud images, in this work) taken within short time periods, because some of the images in the sequence might be very similar. This phenomenon is called twinning and it can lead to optimistically biased cross-validation estimates if very similar images fall into both the training and test partitions. To mitigate this problem, before splitting the data into folds, cloud images are sorted chronologically. Consequently, cloud images which are close in time will most likely fall together either into the training partition or the test partition. This evaluation process avoids the optimistic bias and it will be more representative of a real situation, because it evaluates de classifier with data belonging to a time period different to that of the training data. However, this stricter validation should be expected to report worse metric values than other state-of-the-art works that use other evaluation methodologies.

The metrics used for measuring the effectiveness of the models are: accuracy and macroaverage accuracy. Accuracy is the standard classification success rate:

- Accuracy: $Acc_{Abs} = \frac{I}{S}$,

where I is the number of correctly classified instances and S is the total number of instances.

The problem with (standard) accuracy is that classes with more instances have more weight in the success rate. For instance, in an extreme case, if class A contains 95 images and class B contains just 5 images, accuracy is basically informing about class A. In order to measure the behavior of the model independently of the number of images in each class, macroaverage accuracy can be used. Macroaverage accuracy is defined as the average of the individual class accuracies.

- Macroaverage accuracy: $Acc_{Rel} = \frac{1}{t} \sum_{k=0}^{t} \frac{I_k}{S_k}$

where I_k is the number of well-classified instances for class k, S_k is the number of instances for class k and t is the total number of classes.

All experiments carried out in this work follow the same flow. First, the dataset is ordered chronologically and split into 10 different folds. Then, model evaluation is carried out with a 10-fold cross-validation. In every cross-validation iteration, the training folds are used to select the best *mtry* parameter value (see section 4.3.2) and then build the RF model with that value. Then the model is tested with the test fold. Given that RF is a stochastic algorithm, tenfold cross-validation has been repeated 10 times, each time with a different random seed (in other words, 10 tenfold cross-validations have been carried out). The results obtained are the average of these 10 different runs.

4.4 Results

In this section, results of the different experiments are presented and discussed. One of the aims of this work is to determine the relative contribution of the camera and ceilometer information for cloud classification. Therefore, baseline results were computed by training RF and testing the models using only image features from the camera (spectral, texture and coverage features). Then, RF models

Table 4.3: Overall Results for the Seven Class Experiments (Plus Multicloud)

		Seven classes	
		Featured used	
	Metric	Ca	Ca + Ce
No Multicloud	Accuracy	64.4% (0.6)	77.3% (0.6)
	Macroaverage	62.3% (0.6)	78.0% (0.6)
Multicloud	Accuracy	55.7% (0.6)	71.7% (0.6)
	Macroaverage	55.1% (0.6)	72.6% (0.6)

Note. The accuracy, macro average accuracy (in percent) and standard deviation (within brackets) are displayed separately for experiments with camera only (Ca) and with camera and ceilometer (Ca+Ce). In addition, results are presented separately for experiments excluding (7 classes) and including (8 classes) the multi-cloud category.

were trained and tested with both camera and ceilometer information. To sum up, 8 different experiments were conducted by (1) using up to 7 or up to 10 classes, as described in section 4.2.3; (2) including/excluding the multicloud category; and (3) including/excluding the ceilometer information. Results of experiments have been organized in two blocks, with 7 and 10 classes (both with and without multicloud), respectively.

4.4.1 Seven Cloud Categories Plus Multicloud

Results of the classification procedure when considering seven cloud categories (Table 4.1), with or without multicloud, are given in Figure 4.7 and Table 4.3. Results are displayed for two cases: using just the camera information (Ca), and using the camera together with the ceilometer information (Ca+Ce). For the sake of comparison, results excluding and including the multicloud category are displayed separately.

Results clearly show, first, that the use of ceilometer information (Ca+Ce) improves the performance of the classifier for both cases, without multicloud (seven classes) and with multicloud (eight classes). In the former case, the use of the ceilometer improves accuracy and macroaverage accuracy by 12.91% and 15.72%, respectively

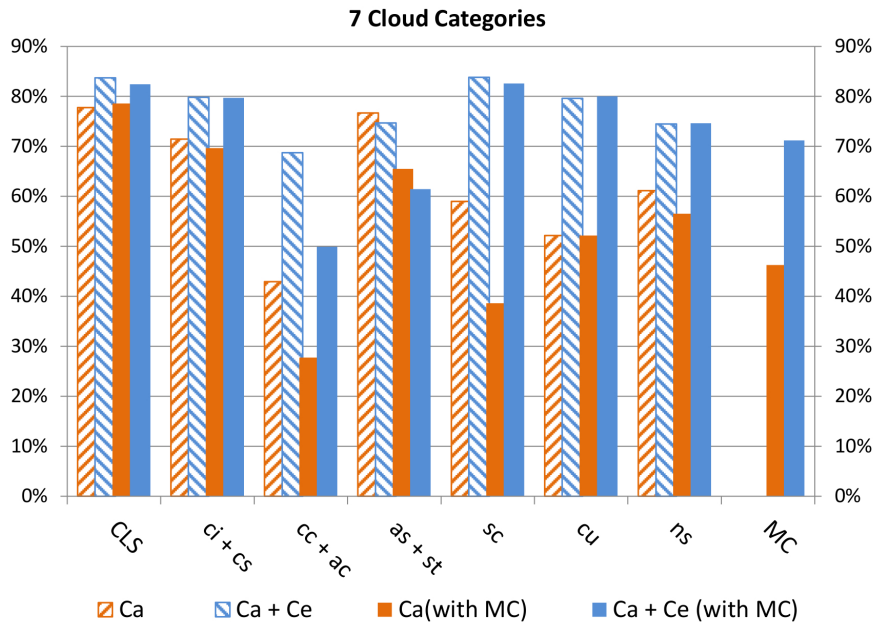

Figure 4.7: Relative frequencies (in percent) of correctly classified cloud classes for the seven cloud types (plus multicloud). Results are displayed separately for the four experiments: using just the camera information (Ca) and both the camera and the ceilometer information (Ca+Ce) but not including the multicloud class, using just the camera information and including the multicloud class (Ca with MC) and using both the camera and the ceilometer information and including the multicloud class (Ca+Ce with MC).

(Table 4.3). The improvement is even larger for the multicloud case (15.97% and 17.46%, respectively). Second, and as expected, including the multicloud class, results in a loss of approximately 5% accuracy when using all the features Ca+Ce, and about 8% when using only the camera (Ca). Interestingly, the ceilometer information allows the classifier to deal better with the extra (and noisy) multicloud class, compared to using only the camera information.

For the nonmulticloud experiments and breaking down results by cloud type, it can be observed (Figure 4.7) that the ceilometer information increases the accuracy for all cloud-types except for stratus & altostratus (in this case, it gets slightly worse by 9.1%). The best improvements are observed for cirrocumulus-altocumulus (32.7%), cumulus (30.5%) and stratocumulus (31.6%). In the rest of the classes, accuracy is also improved to a lesser degree (around 8%). When the multicloud class is included, results are similar regarding the role of the ceilometer. Particularly, the use of ceilometer helps to improve the accuracy of all classes except (again) for the stratus - altostratus class (in this case, accuracy is reduced by just 1.1%). This seems logical, as this class may contain clouds at very different altitudes. Similarly to the no multicloud case, observed improvements are: for cirrocumulus - altocumulus (21.6%), for cumulus (27.6%), for stratocumulus (44%) and for multiple cloud type (24.9%). For clear-sky, cirrus - cirrostratus, and nimbostratus the improvement is smaller (around 5% and 8%). Finally, the accuracy of the multicloud class prediction is remarkable (around 73%) when using the ceilometer; otherwise, it is just about 48%.

The comparison of the results excluding and including the multicloud class reveals some interesting features. First, when using the ceilometer information, the inclusion of multicloud reduces the accuracy of the classification of just some specific cloud types, namely: cirrocumulus-altocumulus and stratus-altostratus. For the rest of the classes, scores are similar. This result makes sense, since the multicloud category somehow includes the cirrocumulus-altocumulus and the stratus-altostratus classes, which are also composed of several cloud types and cloud layers that can be located at very different altitudes. Therefore, the multicloud type may be confused by these two cloud types. This is what is observed in Table 4.4 (the classification contingency matrix). Even though the ceilometer helps enormously in the classification, multicloud is missclassified in about 10% of the cases as cirrocumulus-altocumulus and as

stratus-altostratus. Cirrocumulus-altocumulus is classified as multicloud in 28% of the cases. Previous works have also shown the class cirrocumulus-altocumulus is the most difficult to classify correctly (Kazantzidis et al., 2012; Wacker et al., 2015). The case of the cirrus-cirrostratus class is different, given that these clouds present a quite similar morphology and, more importantly, are usually located at a very similar elevation. As a consequence, multicloud is misclassified as cirrus-cirrostratus just about 4% of the cases.

Table 4.4: Contingency Matrix Results for the Seven Classes (Plus Multicloud) Experiment That Uses the Ceilometer Information

True class	Classified as								Mean sucess rate	
	CLS	ci + cs	cc + ac	as + st	sc	cu	ns	MC	Accuracy	Macroaverage
CLS	**82.2%**	2.3%	0%	0%	0%	9.1%	0%	0%		
ci +cs	13.9%	**79.0%**	12.5%	17.6%	0%	3.1%	0%	4.2%		
cc + ac	0.8%	2.7%	**46.2%**	1.1%	4%	4%	0%	10.6%		
as + st	0%	13.5%	3.5%	**64.2%**	2.2%	0%	8.4%	10.1%		
sc	0%	0%	8.9%	0.9%	**81%**	2.2%	6.5%	1.5%		
cu	2.6%	2.0%	1.1%	0.9%	2.2%	**81.7%**	2.3%	0.1%		
ns	0%	0%	0.1%	2.9%	2.8%	0%	**72.8%**	0.2%		
MC	0.4%	0.3%	27.8%	12.4%	7.8%	0%	10%	**73.4%**	**71.7%**	**72.6%**

Note. Rows contain the true class and columns contain RF predictions. Bold entries represent the percentage of well-classified clouds for each cloud type.

Regarding the multicloud category, (Wacker et al., 2015) reported that the inclusion of this kind of cloud class may reduce the classification rate up to a 50%. Li et al. (2016) reported this sky-category to be the most difficult to classify, nevertheless obtaining an accuracy of 79.5%. This results is similar to the here presented when using the ceilometer (73%). Nevertheless, comparison is difficult given the different sky categories used in Li et al. (2016).

To sum up, the performance of the proposed procedure is highly dependent on the ceilometer information. This dependence is particularly relevant for all the "cumuliform" clouds, whose classification accuracy reduces considerably when only the camera information is used. On the other hand, the method showed to be robust against the inclusion of the multicloud class when ceilometer information is used (only the classification accuracy for the cirrocumulus-altocumulus and the stratus-altostratus is reduced).

4.4.2 Ten Cloud Categories Plus Multicloud

Table 4.5 shows the results when considering the 10 cloud types displayed in Table 4.1. First, it is observed that the accuracy scores decrease compared to the seven classes results described in section 4.4.1. This makes sense, given that the difficulty of classification problems tends to increase with the number of classes. Similarly to the seven class evaluation, a significant increment of the accuracy is obtained when using both camera and ceilometer (Ca+Ce). Overall, these increments are higher than in the seven class case (Table 4.3), indicating that ceilometer information is even more relevant when the number of classes is increased. Particularly, the accuracy and macroaverage accuracy increase by 20.5% and 18.8%, respectively, in the multicloud case.

When the multicloud class is included, accuracy and macroaverage are reduced by a 3%, approximately, if the ceilometer information is used (if only the camera is used, the reduction is larger). This result is similar to the seven class experiment. Therefore, the multicloud type does not seem to be an issue in this case. The reduction in the overall performance of the procedure seems to be related with the other categories.

Figure 4.8 and Table 4.6 break down results per class. It can be

Table 4.5: As in Table 4.3, but for the 10 Classes (Plus Multicloud) Experiments

		10 classes	
		Featured used	
	Metric	Ca	Ca + Ce
No Multicloud	Accuracy	58.8% (0.5)	74.8% (0.7)
	Macroaverage	51.4% (0.5)	66.4% (0.7)
Multicloud	Accuracy	50.6% (0.4)	71.1% (0.6)
	Macroaverage	44.8% (0.6)	63.5% (0.7)

noticed the poor scores for the cirrocumulus and cirrostratus classes, which shows accuracies of near zero and 20% respectively, regardless of the use of the ceilometer (Figure 4.8). Nevertheless, when using the ceilometer information, for some classes (clear-sky) the accuracy increases with respect to the seven class experiment or remain substantially the same (cumulus, stratocumulus, nimbostratus and multicloud). Regarding the results for the formerly combined classes (cirrocumulus-altocumulus, cirrus-cirrostratus and stratus-altostratus), now separated, some relevant outcomes were found. For instance, when using the ceilometer and including the multicloud class (Table 4.6), the stratus and altostratus classes show a high accuracy, 78.9% and 64.8%, respectively, higher than the combined stratus-altostratus class in Table 4.4 (62.2%). Note in Figure 4.8 the very relevant information provided by the ceilometer for these two classes. Both kinds of clouds present similar morphological features. The main difference is the location: while stratus are low-level clouds with the CBH below 2 km, the altostratus CBH are typically well above this elevation (Houze Jr, 2014; Kokhanovsky, 2006). In our case, Table 4.1 data confirms these values, since the mean CBH of the stratus clouds is 833 m and the corresponding value for the altostratus is 6701 m. Therefore, it seems that the combined information derived from the camera and especially the ceilometer is able to properly discriminate between these two classes of clouds, even when the multicloud class is included.

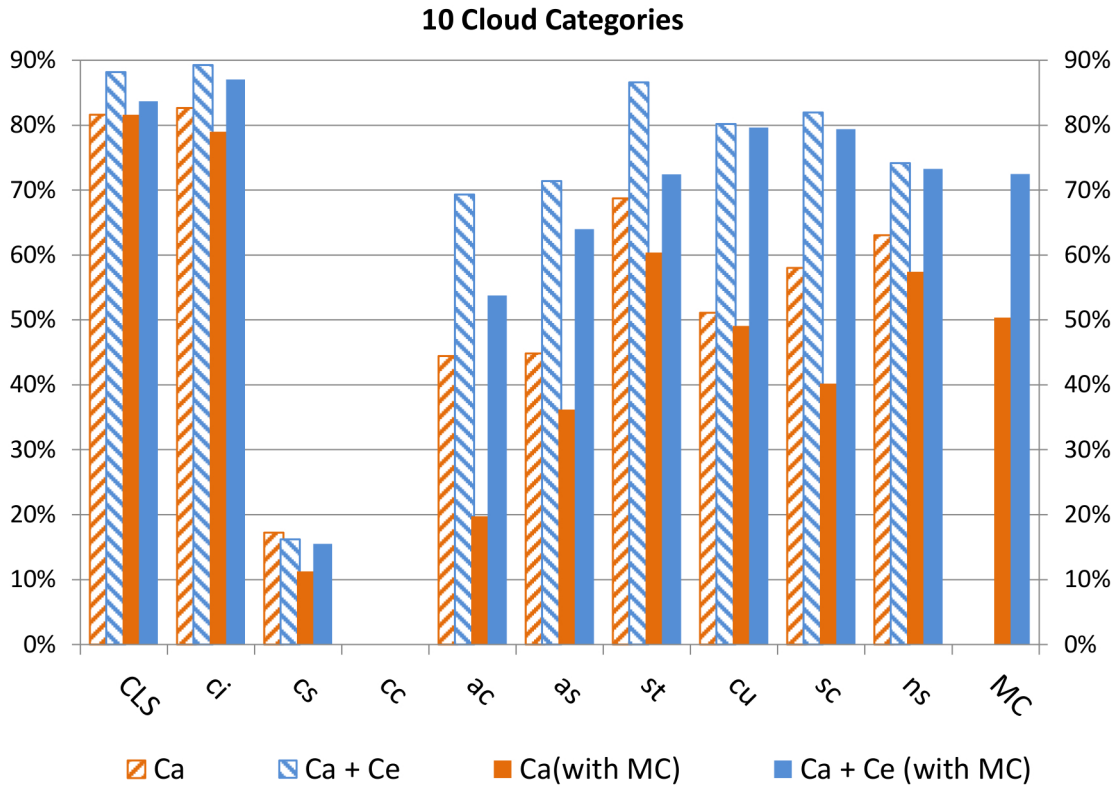

Figure 4.8: As in Figure 4.7, but for the 10 cloud categories (plus multicloud).

Table 4.6: As in Table 4.4 but for the 10 Classes (Plus Multicloud) Experiments

True class	Classified as											Mean sucess rate	
	CLS	ci	cs	cc	ac	as	st	sc	cu	ns	MC	Accuracy	Macroaverage
CLS	**84.3%**	2.9%	0%	0%	0%	0%	0%	0%	9.6%	0%	0%		
ci	11.8%	**86.4%**	22.5%	35.7%	11.4%	9.5%	0%	0%	2.7%	0%	1.9%		
cs	0%	3.1%	**21.5%**	0%	0%	14%	0%	0%	0%	0%	2.2%		
cc	0%	0.6%	0%	**0%**	0.5%	0%	0%	0%	0%	0%	0.9%		
ac	0.6%	1.2%	0%	22.1%	**54.1%**	2%	0%	4%	2.2%	0%	9.0%		
as	0%	0.8%	50%	0%	2.9%	**64.8%**	0%	0%	0%	0%	5.3%		
st	0%	0%	0%	0%	0%	0%	**78.9%**	2%	0%	9.5%	4.5%		
sc	0%	0%	0%	0%	10.4%	0%	1.7%	**80.8%**	2%	7.7%	1.7%		
cu	3.3%	2.8%	0%	7.1%	0%	0%	1.8%	2%	**81.8%**	2.3%	0.2%		
ns	0%	0%	0%	0%	0%	0%	3.9%	3.8%	0%	**72.3%**	0.2%		
MC	0%	2.1%	6%	35%	20.6%	9.8%	13.7%	7.4%	1.6%	8.1%	**74.1%**	**71.1%**	**63.5%**

Note. Bold entries represent the percentage of well-classified clouds for each cloud type.

The separation of the class cirrus-cirrostratus is not so successful. The cirrus category is reliably classified, reaching 86.4% accuracy (Table 4.6). Note that the information provided by the ceilometer is not highly relevant in this case (Figure 4.8). Nevertheless, as commented above, cirrostratus results are poor (21.5%). They are classified as altostratus in 50% of the cases and as cirrus in 22.5% of the cases (Table 4.6). These results can be explained based on of the similar characteristics of the altostratus and cirrostratus clouds. Particularly, the altostratus clouds present a mean CBH of 6701 m, with 1701 m standard deviation value (Table 4.1). The corresponding values of the cirrostratus clouds are 7684 m and 676 m. These experimental values are confirmed in the bibliography, which states that the range of elevation in middle latitudes is 2-7 km for the altostratus and 5-13 km for cirrostratus (Houze Jr, 2014; Kokhanovsky, 2006). The CPD for both types of clouds are also similar: 1858 m and 1829 m for the altostratus and cirrostratus, respectively, (Table 4.1). Therefore, these clouds cannot be discriminated just based on the CBH and the CPD. The main difference between these two kinds of clouds is the usual presence of the hallo feature in the cirrostratus clouds but not in the altostratus. This particular feature seems to be not resolved by the image characteristics used here.

Finally, the poorest results are obtained for the cirrocumulus-altocumulus discrimination. Particularly, cirrocumulus are systematically misclassified as either altocumulus, cirrus, cumulus or multicloud (Table 4.6). These poor results can be explained based on several reasons: first, because the mean and the standard deviation CBHs values (Table 4.1) of the cirrocumulus (6833 m and 2023 m), altocumulus (4494 m and 2257 m) and cirrus (9086 m and 1515 m) does not allows the use of the CBH to discriminate the cirrocumulus from the other two cloud types. Reference values of the CBHs in middle latitudes are 5-13 km for cirrocumulus, 7-10 km for cirrus 2-6 for altocumulus (Houze Jr, 2014; Kokhanovsky, 2006), therefore confirming our results. Similar inferences can be derived for the role of the CPDs, which show mean and standard deviation values (Table 4.1) that makes the CPD inadequate to discriminate between these tree cloud types. Again reference values in the bibliography confirm these findings; particularly Houze Jr (2014) and Kokhanovsky (2006) report the geometrical thickness of the cirrocumulus to be in the range 0.2-0.4 km, which overlaps the thickness of the altocumulus (0.2-0.7 km) and cirrus (0.1-3 km). Therefore, ceilometer information seems

not to be relevant to distinguish the cirrocumulus clouds from many other classes. Regarding the sky-camera information, from the morphological point of view, cirrocumulus and altocumulus are similar. In addition, cirrocumulus clouds often occur in small sheets located very high in the atmosphere (even 9 km values can be found in the experimental data set here used). As a consequence, and probably also because of the low resolution of the TSI images, the camera is not able to provide distinctive statistics values for this particular cloud class. Altocumulus results are more encouraging (accuracy 54%), although they are misclassified as multicloud in 20.6% of the cases.

4.5 Summary and Conclusions

We have presented and evaluated a methodology for automatic cloud-type classification based on the synergistic use of a sky camera and a ceilometer. The hypothesis is that, given the distinctive vertical location of the different cloud types, the use of the ceilometer may improve the classification accuracy derived just from the camera images.

The methodology here evaluated aims to be fully operational, reporting an automatic classification of the cloud/sky conditions every 5 min. Because of that, among the evaluated cloud types, we have included the multicloud category, which accounts for skies covered by several cloud types and/or cloud layers. The automatic classification was conducted by Random Forests, a state-of-the-art machine learning classification algorithm, which used as input 19 features (12 computed from the sky-camera and 7 from the ceilometer). The procedure was trained and evaluated on a set of 717 images and up to 11 different types of clouds/skies were considered. The study is performed using a 130° field of view in the sky cameras.

A total of eight experiments were conducted by: (1) excluding/including the ceilometer information in the random forest automatic classification algorithm; (2) including/excluding the multicloud category and (3) using 7 or 10 different cloud/sky types, in addition to the clear sky and the multicloud category. The comparison of results allowed to evaluate the role of the ceilometer, to analyze the effect of the multicloud type in the classification accuracy and, finally, to evaluate the performance of the model when using an increased

number of cloud/sky types (7 versus 10).

For the seven cloud/sky classes experiments plus multicloud (six cloud types + clear sky + multicloud), results showed an overall accuracy of the method of about 72% when using the ceilometer and about 55% when using just the camera information. Therefore, the ceilometer information showed to be crucial. The use of the ceilometer is particularly valuable for classifying cumuliform clouds, with an increment of the accuracy of about 30% compared to the use of only the camera information and, particularly, for the multicloud category, which is correctly estimated in about 73% (about 48% when just the camera is used). In addition, as may be expected, the inclusion of the multicloud class results in a loss of approximately 5% accuracy when using ceilometer and camera information, and about 8% when using only the camera information. The 5% reduction was accounted for just some specific cloud types (cirrocumulus-altocumulus and stratus-altostratus). For the rest of the classes, scores were similar. To sum up, the ceilometer information allowed the classifier to deal better with the extra (and noisy) multicloud class, compared to using only the camera.

Results for the augmented 10 cloud/sky classes plus multicloud experiments (nine cloud types + clear sky + multicloud) showed lower accuracy scores. This makes sense, given that the difficulty of classification problems tends to increase with the number of classes. Particularly, for the experiment using the ceilometer information and including the multicloud class, mean macroaverage reduces from about 73% (7 cloud/sky classes plus multicloud) to about 63% (10 cloud/sky classes plus multicloud). The use of the ceilometer information resulted to be even more critical for the 11 categories than for the 8. From the analysis by categories of the experiments with 10 cloud/sky classes plus multicloud, some additional conclusions were obtained. The first one is that the reduction in the accuracy was not related with inclusion of the multicloud category, but with other cloud types. Notably, the classification accuracy for cirrostratus, altocumulus, and particularly, cirrocumulus, showed to be low. Several reasons were found for this low accuracy. First, the fact that these clouds, from the morphological point of view, are similar to other cloud types: cirrostratus and cirrus, altocumulus and cirrocumulus/multicloud and cirrocumulus and cirrus. As a consequence, the camera features were not able to distinguish between these kinds of clouds. Second, many of these clouds present similar cloud base height and geometrical

thickness, making the ceilometer information not so relevant. This is the case of the cirrostratus and altostratus, and the altocumulus and cirrocumulus.

Other cloud types, as the stratus and altostratus, showed encouraging classification accuracy. Although these clouds present similar morphological characteristics, they are located at different elevation. As a consequence, ceilometer information allowed to reach better accuracy for these cloud types, even with the presence of the multicloud class.

Several applications may benefit from the here proposed automatic and operational cloud recognition system. For instance, in the field of solar energy, the here proposed method can be used to enhance the reliability of sky camera-based solar radiation estimating and forecasting procedures. Also for the characterization of the solar radiation spatial and temporal variability, that is an important issue for the solar energy grid integration. In addition, this methodology may reduce the uncertainty in the energy balance of the Earth surface, which is mainly related to the clouds. Finally, aviation weather services, that used the cloud type as a proxy of the present and for coming weather conditions, can benefit from the here proposed methodology.

Results here presented seem encouraging regarding the development of an automatic, fully operational and highly tailored cloud classification procedure. Nevertheless, some limitations were found and some challenges should be addressed. First, ceilometer information was found to be highly valuable to accurately classify certain types of cloud (especially altocumulus, cumulus, stratocumulus, stratus and altostratus) and makes the procedure robust against the inclusion of the multicloud. Camera information alone was found to be not suitable to deal with multicloud situations. Nevertheless, even the use of the ceilometer information showed some limitations. Problems are related with cloud types which present similar morphological characteristics and, at the same time, similar elevation and geometrical thickness. In these cases, the only way to increase the classification accuracy is to develop specific features, either spectral or textural ones, able to account for the differences between cloud types. An example is the "hallo" phenomena, which is present in cirrostratus but not in altostratus, and that the features used were not able to account for. Regarding this, the use of advanced sky

cameras, with enhanced resolution and/or spectral responses seems a promising tool. This will be explored in future works, as well as the role of the time window in classification accuracy.

In future works we aim to apply the here proposed methodology for different areas or patches in the image. This will allow, eventually, the classification of some of the here considered multicloud images in some specific cloud-type categories. In addition, alternative approaches to the use of the ceilometer information will be also explored. For instance, the use of stereographic methods to derived CBH (Kassianov et al., 2005; Peng et al., 2015) or the use of the cloud speed as proxy for the CBH (Peng et al., 2016; Quesada-Ruiz et al., 2014).

Acknowledgments

The authors are supported by the Spanish Ministry of Economy and Competitiveness, projects ENE2014-56126-C2-1-R and ENE2014-56126-C2-2-R and FEDER funds. The University of Jaén affiliated authors are also funded by the Junta de Andalucía (PAIDI research group TEP-220). The all-sky camera images and ceilometer data are available via COPDESS repository and at the URL: `http://matras.ujaen.es/data/Skycamera_and_ceilometer_data.nc`.

CHAPTER 5

Analysis of the intra-day solar resource variability in the Iberian Peninsula

Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Santos-Alamillos, F. J., Tovar-Pescador, J., and Pozo-Vázquez, D. (2018). Analysis of the intra-day solar resource variability in the Iberian Peninsula. *Solar Energy*, 171:374 - 387. DOI:https://doi.org/10.1016/j.solener.2018.06.060

5.1 Introduction

Great efforts have been made along the last decades to make renewable energy a plausible alternative to conventional energy generation system. In particular, the European Council has committed to achieving at least 27% renewable energy penetration in the European Union by 2030 (European Commission, Secretariat-General, 2016).

One of the most important drawbacks of integrating high levels of renewable generation in the electricity system, particularly solar energy, is the variability of the resource. Unlike conventional power generation, solar power is conditioned by weather, and is thus inter-

mittent. The power system operators try to keep the balance between generation and load in the grid. Due to the fact that the forecasts of generation and load are not completely accurate, power system operators rely on reserves to manage the anticipated and non-anticipated variability (Yang et al., 2018). The incorporation of the solar energy adds additional uncertainty to the power system from the generation side (Renné, 2014). As a consequence, an increase in the amount of balancing reserves is necessary to keep the grid stable. This increase is related to the historical solar power forecasting error distribution (Matos et al., 2017). Another consequence is the existence of specific economic costs for the integration of the solar energy. The reduction in operation reserves (on average), and consequently the integration costs, through the improvement of the solar forecast, was addressed by Zhang et al. (2015a). As solar energy penetration increases, the effect of the resource variability in power grid operation will be more challenging. The case of the Iberian Peninsula grid is particularly interesting, mainly because this grid, constituted by the Portuguese and Spanish grids, is largely isolated from the rest of Europe (General Court, 2015).

A better understanding of solar resource variability and its causes at various spatial and temporal scales would be helpful to develop efficient mitigation solutions (Coker et al., 2013; Engeland et al., 2017). Solar irradiance varies at a wide range of temporal scales (from seconds and minutes to years and decades), associated with processes that operate at spatial scales from thousands of kilometers to a few meters (solar cycle, synoptic-related cloud variability, convection and orographic clouds, aerosol load, etc.).

In general, temporal scales from seconds to minutes are related to local processes, such as convective or orographic clouds, whose area of influence is limited. These scales are relevant for optimal management of a solar power plant (or set of plants) in a limited area (Lave et al., 2012). Temporal scales from tens of minutes to hours (intra-day variability) are often associated with distinct synoptic weather patterns. As the area of influence of these patterns usually reaches hundreds of square kilometers, these time scales are relevant to grid integration objectives (Perez et al., 2016). Finally, climatic time scales (seasons to years), characteristic of the solar cycle, are more appropriate for plant project development (Lohmann et al., 2006; Pozo-Vázquez et al., 2004; Pozo-Vazquez et al., 2011; Ruiz-Arias et al., 2016; Santos-Alamillos et al., 2017b). On this matter,

Gutiérrez et al. (2017) analyzed spatial and interannual variability of the solar PV yield of the Iberian Peninsula.

The temporal variability of solar irradiance as a function of temporal resolution was extensively analyzed (Gueymard and Wilcox, 2011; Lave et al., 2017; Moreno-Tejera et al., 2016; Perez et al., 2011). Moreover, analyses of its spatial variability and the relationship between temporal and spatial scales of variability was addressed in several papers (Barnett et al., 1998; Perez et al., 2012). A common feature emerging from this work is an increment in the correlation with reducing distance and with increasing temporal aggregation (Hoff and Perez, 2012; Mills and Wiser, 2010; Perez et al., 2016). A byproduct of the spatial variability of solar resources is the smoothing effect on PV plant output caused by the aggregation of geographically disperse PV plants (Lave et al., 2012; Marcos et al., 2012; Hoff and Perez, 2012; Wiemken et al., 2001; Zagouras et al., 2014).

Although clouds are the most important factor in solar resource variability, very few papers addressed the relationship between this variability and cloud type and coverage (Kasten and Czeplak, 1980; Salgueiro et al., 2016; Tzoumanikas et al., 2016). The main issue is the lack of appropriate observational records. Instead of observation, Reno and Stein (2013) used cloud categories derived from satellite images to understand and model the spatial variability of solar radiation. Similarly, McCandless et al. (2016b) used satellite-derived cloud information to improve short-term solar radiation forecasting.

Cloud regimes are linked to specific weather patterns. Therefore, analysis of these patterns is an important step toward understanding solar resource variability. The relationship between wind power output variability and weather patterns was addressed in several papers (Brayshaw et al., 2011; Correia et al., 2017; García-Bustamante et al., 2013; Ohba et al., 2016; Steiner et al., 2017). These studies indicated that synoptic weather patterns associated with large-scale atmospheric circulations are important to understand wind power variability and can be effective predictors of wind generation variability. Unlike the case of the wind, very few papers addressed the relationship between weather patterns and solar power variability. Recently, Köhler et al. (2017) identified critical weather patterns for solar power integration in Germany. As a consequence, understanding the influence of weather patterns on solar radiation resource variability remains limited.

The present work aims at contributing to the understanding of intra-day modes of variability of the solar resource in the central and southwestern parts of the Iberian Peninsula. A better understanding of this variability and its underlying causes may provide mitigation solutions for solar energy grid integration in the study area, which is characterized by the best solar resources in Europe (Santos-Alamillos et al., 2017b). For instance, the knowledge of the atmospheric circulation patterns may help in estimating and forecasting available resources and their intermittency in the coming hours, aiding operation of the system (McCandless et al., 2015). We hypothesize the existence of specific modes of solar radiation intra-day variability in the study area associated with specific weather regimes.

Firstly, hierarchical cluster analysis is performed to obtain the modes of variability of the solar resource. This type of analysis allows the classification of a set of solar radiation values into specific regimes. The analysis is based on global horizontal irradiance (GHI) and direct normal irradiance (DNI) data collected at four stations representative of the study area over a two-year period. Given the objectives of our work, choosing the appropriate temporal scale for the cluster analysis is a critical issue. Here, following Perez and Hoff (2013b), three-hour mean and variability statistics computed with 15-minute averages based on GHI and DNI values are obtained. These statistics are used as input to the cluster analysis.

Secondly, synoptic weather patterns responsible for the different intra-day modes of solar resource variability are obtained. To this end, we perform composite analyses of pressure reduced to mean sea level (MSLP) and cloudiness data. The solar resource modes are analyzed in the light of their associated weather patterns and cloud regimes. Finally, the aggregated solar PV power yield in the study area associated with each analyzed mode is assessed. Yearly and seasonal analyses are conducted independently.

The work is organized as follows. Section 5.2 describes the study area while Section 5.3 describes the dataset and the methodology. Results are presented and discussed in Section 5.4, while in Section 5.5 some concluding remarks are proposed. The results of the seasonal analyses are detailed in 5.6.

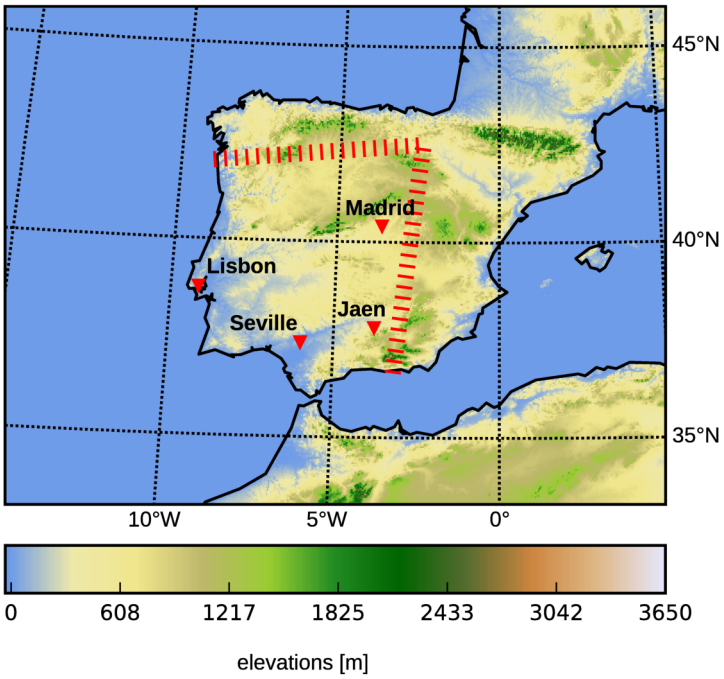

Figure 5.1: Digital elevation map showing the location of the four analyzed stations. The red dashed lines delimits the study area.

5.2 Study area

This study is based on the analysis of data collected at four radiometric stations, namely, Jaen, Lisbon, Madrid and Seville, located in the central and southwestern Iberian Peninsula (see Fig. 5.1). The study area accounts for the majority of the currently installed solar power in the Iberian Peninsula.

From a climatic point of view, the study area is located in a transition zone from temperate to subtropical climates. Several studies showed that the atmospheric circulation over the study area is ruled by the semi-permanent Azores Islands anticyclone (Castro-Díez et al., 2002; Trigo et al., 2002a, 2004). As a consequence, the study area shows the same climatic regime regarding cloudiness (Trigo et al., 2002a) and solar radiation (Pozo-Vázquez et al., 2004; Pozo-Vazquez et al., 2011; Sancho Ávila et al., 2012; Santos-Alamillos et al., 2012). The position and intensity of the Azores anticyclone changes throughout the year, giving rise to a marked seasonality in their influence area.

From a topographic point of view, the study area may be split into two different parts. The western area is a nearly flat region, with many wide valleys open to the Atlantic Ocean. The Lisbon and Seville stations are located in this part. 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). As a consequence of these different topographic characteristics, certain spatial variability in the cloudiness and solar radiation can be anticipated associated with local features (Lauret et al., 2016). In addition, in coastal areas, as for Lisbon, the availability of moisture from the ocean and the thermally driven wind may cause local cloudiness. Moreover, orographic clouds are expected in areas close to a mountain range. The importance of regional patterns in variability in contrast to local ones is here evaluated at time scales relevant for solar energy grid integration.

5.3 Data and methods

5.3.1 Data

One-minute solar irradiance data collected at four ground stations are used in this work. The three components of solar radiation, GHI, diffuse horizontal irradiance (DHI) and DNI, were gathered except in the Madrid station, where only GHI and DNI were available. The raw measurements cover the period from March 2015 to February 2017.

Table 5.1 shows the sensors of ground stations. In Jaen (University of Jaen), data are collected at a frequency of 30 seconds. The closest records to each minute time step are used as representative of the selected 1-minute resolution. To remedy a shadowing issue in Lisbon (University of Lisbon), data associated with solar zenith angle (SZA) $>75^{\circ}$ are discarded. In Madrid (Institute of Solar Energy of the Technical University of Madrid), the records for which GHI = 0 W m^{-2} are discarded. Lastly, the Seville station belongs to a private company.

The quality of the data is first verified by visual examination and then by quality tests. In particular, following Long and Dutton

Table 5.1: Sensors of ground station.

ground station	pyranometer sensor	pyrheliometer sensor
Jaen	Delta-Ohm LP PYRA 10	Delta-Ohm LP PYRHE 16
Lisbon	Secondary Standard*	First Class*
Madrid	Hukseflux SR11-05	Eppley NIP
Seville	Kipp & Zonen CMP6	Kipp & Zonen CHP1

*According to ISO 9060.

(2002), various filters are applied 1) not to allow for extremely rare values, and 2) to check for coherence of the solar irradiance components, the so-called closure-equation filter. Threshold values to the ratio of GHI over the sum of DHI and horizontal direct irradiance, i.e. $DNI \cdot cos(SZA)$, are set to 15%, instead of the 8% proposed by Long and Dutton (2002) when SZAs are lower than 75°. For the Madrid station, because of a lack of DHI data, only filters 1) and 2) are applied. After the filtering procedure, only records available simultaneously at the four stations are considered in further analyses. Thus, each station has the same number of records.

The number of stations producing high-resolution GHI and DNI measurements required for this study is scarce. Nevertheless, the use of measured data instead of satellite estimates is preferable because of the errors of those estimates, particularly in the case of DNI (Ineichen, 2014).

MSLP and cloudiness data from the North Atlantic region are used to study atmospheric circulations associated with the variability modes. The MSLP data are obtained from Global Forecasting System (GFS) analysis files with spatial resolution 0.5°. Data are available from the National Centers for Environmental Prediction at 0, 6, 12 and 18 hours Universal Time Coordinate (UTC) (National Centers for Environmental Prediction, NCEP, 2006). Cloudiness data correspond to the EUMETSAT Cloud Mask (CLM) product (Kniffka et al., 2014). This product classifies each satellite pixel as one of the following types: clear sky over water, clear sky over land, cloud, and no data (Derrien and Gléau, 2005). CLM data are available at 15-minute resolution, with a spatial resolution of 3 x 3 km at latitude 0°, longitude 0°.

The corresponding two-year data of solar PV hourly generation grouped by provinces are published by the Spanish Transmission System Operator (TSO) Red Eléctrica de España (ESIOS, 2018), to-

gether with the installed capacity. In order to obtain the solar PV hourly capacity factor (CF), the PV generation and the installed capacity of those Spanish provinces within the study area are aggregated. Then, the CF is computed by dividing the power production by the installed capacity. The solar PV hourly CF distribution associated with each mode of variability of the yearly analysis is computed and analyzed.

5.3.2 Solar irradiance data processing

The present analysis aims at highlighting the existence of common modes of intra-day variability of the solar resource across the study area. To this end, both the time interval over which the solar irradiance changes and the period over which the variability is assessed should be selected (Perez and Hoff, 2013b). Given the wide range of scales of solar irradiance temporal variability, this choice should be made in the light of the purpose of the analysis. In the present work, a time interval of 15 minutes is selected in order to average the 1-minute measured data. Based on the resulting 15-minute data series, the period considered for variability is three hours.

The time interval is selected primarily as appropriate to solar energy grid integration. The time interval defines the cloud perturbation size causing solar resource variability (Perez and Hoff, 2013b). High-frequency variability (seconds to minutes) are mainly caused by local clouds. In contrast, low-frequency perturbations are often related to large cloud structures associated with distinctive synoptic perturbations. In that case, significant correlation can be found among stations separated by hundreds of kilometers, because they have a coherent variability. To account for this variability, 15 minutes appears to be a reasonable choice, because it is assumed that this time interval could be sensitive to the solar resource variability caused by both local clouds and larger cloud structures. It should be highlighted that 15 minutes is also the temporal resolution of most satellite-derived solar irradiance databases. In areas lacking ground measurements, these satellite data can be used to apply our proposed analysis.

The three-hour period is selected based on results of previous analyses in the study area (Martínez-Chico et al., 2011; Huertas-Tato et al., 2017b) that showed that most cloudiness regimes last about

three hours. Only clear sky and certain types of clouds, mainly cirrus, generally last longer. Each cloud type has a distinctive signature (attenuation and temporal variability) of the solar irradiance (Martínez-Chico et al., 2011; Tzoumanikas et al., 2016). Therefore, by analyzing solar radiation within three-hour intervals, a more homogenous and representative variability can be expected.

Based on filtered 1-minute GHI and DNI datasets and their corresponding clear sky values, 15-minute means are computed, associated with 0, 15, 30 and 45 minutes in each hour. Each 15-minute record is obtained by averaging the 15 previous 1-minute values, including the last value that defines the time interval. Following Roesch et al. (2011), the 15-minute records are only computed when at least 20% of 1-minute values are available. The 15-minute average records are used to compute various statistics of the three-hour reference periods, i.e., 0, 3, 6... hours UTC of each day. Unlike the 15-minute average records, three-hour statistics are only computed if the 12 corresponding 15-minute records are available.

Although several statistics characterizing the solar resource are found in the literature (Gutiérrez et al., 2017; Moreno-Tejera et al., 2017), in the present work, two features are used following Lauret et al. (2016). In particular, mean clear-sky index values for both GHI ($\overline{K_c}$) and DNI ($\overline{K_{cb}}$) are computed. Additionally, in order to describe solar resource variability, the standard deviation of the increment of the clear-sky index of the GHI over the reference period (so-called nominal variability) is computed:

$$\sigma_{\Delta K_c} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Delta K_{c_i} - \overline{\Delta K_c})^2}$$

Similarly, nominal variability of the DNI ($\sigma_{\Delta K_{cb}}$) is defined with ΔK_{cb} instead of ΔK_c.

Although other measures of solar irradiance variability can be used, the nominal variability has been shown to provide more reliable measures of that variability within a reference period (Perez et al., 2016).

In summary, the database used in this study is composed of 1105 sets of values, each one representative of a period of three hours. Each set of values contains 16 attributes, i.e., four statistics ($\overline{K_c}$, $\overline{K_{cb}}$, $\sigma_{\Delta K_c}$, and $\sigma_{\Delta K_{cb}}$) computed at each of the four stations.

5.3.3 Hierarchical clustering method

In order to obtain modes of coherent variability in the dataset, a hierarchical clustering algorithm (Wilks, 2006) is applied. Using this method, the results are invariant to the number of attempts with the same number of groups. At the beginning of the procedure, each set of values constitutes a group. At each step, two groups are merged in a new group. The process continues until only one group made up of all sets of values is obtained. Ward's minimum variance method described in Wilks (2006) is used to decide which groups are merged at each step. However, not all of these aggregation stages provide intuitive clusters. Thus, to assess the suitability of the resulting groups, two subjective graphical metrics are used: silhouette coefficients (Rousseeuw, 1987) and the distance between groups at each aggregation step (Wilks, 2006).

The silhouette coefficients help to simultaneously assess the number of members belonging to each group and the suitability of each set of values to its assigned group. For each set of values i, its silhouette coefficient, *s(i)*, varies from -1 to 1. A *s(i)* value close to 1 means that the set of values fits well in the assigned group. A value of *s(i)* close to 0 means that the set of values, although it is in the best possible group, is near a neighboring group. A negative *s(i)* means that the set of values fits better in a neighboring group. The plot of silhouette coefficients is built group-by-group and the *s(i)* coefficients are sorted from largest to smallest.

The distance between groups at each aggregation step can reveal an intuitive number of groups. In the hierarchical clustering method, when two groups are merged at each stage, the distance between the two members of the newly created group is registered. The distances can be plotted as a function of the aggregation step number. Major slope changes in the curve provide evidence of the existence of a natural number of groups.

Input data to the clustering procedure are the sets of values: $\overline{K_c}$, $\overline{K_{cb}}$, $\sigma_{\Delta K_c}$, and $\sigma_{\Delta K_{cb}}$, computed for the four stations. Because these statistics vary in different ranges, this dataset is scaled (mean equal to 0 and standard deviation equal to 1).

Given the marked climate seasonality in the study area, two independent analyses are conducted: an annual analysis using the entire

dataset, and a seasonal analysis. In the latter, winter databases are composed of the months of December, January, and February, spring by March, April, and May, summer by June, July, and August, and autumn by September, October, and November.

5.3.4 Cluster number selection method

Even with use of the two aforementioned graphical tools, the selection of the number of groups is not a straightforward task and is prone to subjective choices. Instead, the final number of groups in each analysis (annual and seasonal) is selected in the light of the relevance of their associated weather patterns, from both physical and statistical points of view. Thus, the main objective in the present work, which is to identify common modes of intra-day variability of the solar resource in the study area, is reached by assessing their associated MSLP anomaly maps. A composite MSLP anomaly field of each group is determined by computing, at each grid cell, the average of differences between the corresponding observation values and the mean value over the study period (seasonal or annual).

The decision method begins by assessing MSLP anomaly fields of the last two groups of the clustering procedure. Due to the fact that the distribution of the MSLP samples is not normal, for each grid cell, the Kruskal-Wallis test (Kruskal and Wallis, 1952) and, if applicable, the pairwise-comparison Conover-Iman test (Conover and Iman, 1979) is applied, with a significance level of 0.05, in order to ensure that all groups are statistically different. In addition, the metheorological meaning of the corresponding MSLP anomaly maps are analyzed. If the patterns are found to be relevant, the next set of groups (three groups) is evaluated, and so on. The decision procedure stops when none or few of the grid cells are statistically significant. At this point, the previous set of groups is selected.

5.4 Results

Table 5.2 shows the number of groups (modes of variability) selected for the annual and seasonal cluster analyses, along with the number of sets of values assigned to each group.

Table 5.2: Absolute and relative number of sets of values per group at each time period of analysis.

Period of analysis	Group 1	Group 2	Group 3	Group 4
Annual	198 (17.92%)	256 (23.17%)	266 (24.07%)	385 (34.84%)
Winter (Dec, Jan, Feb)	25 (15.53%)	47 (29.19%)	58 (36.02%)	31 (19.25%)
Spring (Mar, Apr, May)	53 (14.4%)	85 (23.1%)	125 (33.97%)	105 (28.53%)
Summer (Jun, Jul, Aug)	58 (18.35%)	107 (33.86%)	151 (47.78%)	
Autumn (Sep, Oct, Nov)	76 (29.23%)	96 (36.92%)	88 (33.85%)	

Results of the annual, winter, spring, summer, and autumn analyses are shown in Figures 5.2, 5.5, 5.6, 5.7, and 5.8, respectively. In each figure, the first row is dedicated to the plots of the silhouette coefficient metric and the distance between groups at each aggregation step. The subsequent rows include as many columns as resulting groups. The second row contains the MSLP anomalies maps. To further assess the clustering procedure, the third row shows composite maps of CLM values. To this end, cloudiness information of each CLM file is encoded as binary values, 1 being a cloudy pixel and 0 the rest. Then, the mean of the 15-minute CLM values attached to each group is computed at each grid cell. The fourth and fifth rows display the normalized histogram of the three-hour statistics at the four analyzed stations for GHI ($\overline{K_c}$ and $\sigma_{\Delta K_c}$) and DNI ($\overline{K_{cb}}$ and $\sigma_{\Delta K_{cb}}$), respectively.

Figure 5.2 summarizes results of the annual analysis. Based on the decision method, four groups (modes) are identified. The silhouette coefficient metric (Fig. 5.2a) shows that group 1 and especially group 4 are the most consistent. Both groups exhibit positive and relatively large *s(i)* values for most of the sets of values, meaning that they fit well in their groups. On the contrary, groups 2 and 3 show less consistency, with sets of values weakly assigned to their clusters and many sets of values that would belong to other groups (i.e., sets of values with negative *s(i)* values). Figure 5.2b suggests that the number of groups may be significantly larger (at lower stages) than the number of selected groups. The distance at the selected set of groups is better shown in Figure 5.2c.

The MSLP anomaly field corresponding to the first group is displayed in Fig. 5.2d. An area of negative MSLP anomaly is observed centered over the northwestern Iberian Peninsula, indicating the presence of low-pressure centers (*low-pressure pattern* hereafter). This synoptic condition is associated with thick clouds (with large optical depth values) over the study area. Accordingly, large cloud-

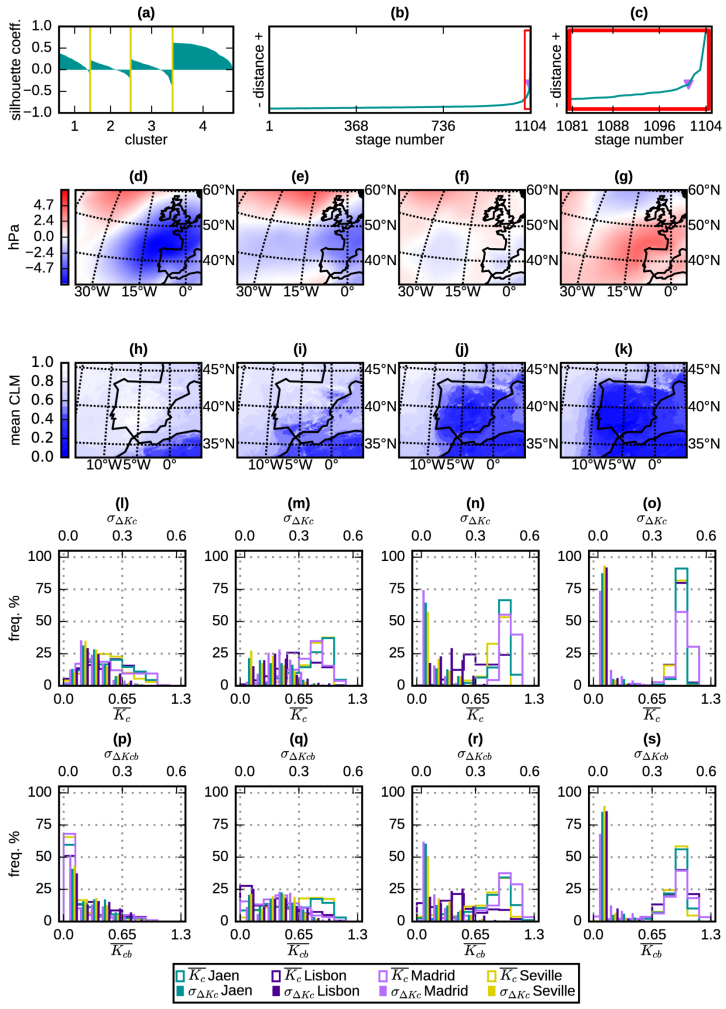

Figure 5.2: Result of annual analysis. (a) Plot of the silhouette coefficients. For each set of values (horizontal axis) its silhouette coefficient is shown in green. The vertical yellow lines separate the groups, whose coefficients are sorted from largest to smallest. (b) and (c) Figures showing the distance between groups at each aggregation step (c shows a part of the Figure b but enlarged). The purple triangle corresponds to the stage of the selected number of groups. Second row: (d), (e), (f), and (g) MSLP anomalies (in hPa) maps corresponding to the first, second, third, and fourth group, respectively. Positive areas are displayed in red while negative ones are indicated in blue. Third row: (h), (i), (j), and (k) CLM mean values maps focused on the Iberian Peninsula for the four groups. The pixels vary from blue (clear sky) to white (completely cloudy). Fourth row: (l), (m), (n), and (o) Distribution of the $\overline{K_c}$ (step-type) and $\sigma_{\Delta K_c}$ (bar-type), at the four analyzed stations, for the first, second, third, and fourth group, respectively. Variable and stations colors are indicated in the panel at the bottom. At each figure, the left vertical axis indicates the frequency distribution (in %); the bottom horizontal axis shows the $\overline{K_c}$ values while the top horizontal axis shows the $\sigma_{\Delta K_c}$ values. Fifth row: (p), (q), (r), and (s) As (l), (m), (n), and (o) but for the $\overline{K_{cb}}$ and $\sigma_{\Delta K_{cb}}$ distributions.

cover values are observed there (Fig. 5.2h). Solar radiation statistics distributions (Fig. 5.2l and p) clearly reflect this cloudiness. At the four stations analyzed, small values and reduced variability of the solar resource are evident. Nevertheless, there are notable differences between the DNI and GHI statistics distribution. In particular, the $\overline{K_{cb}}$ distribution shows a tendency toward smaller values than that of $\overline{K_c}$, whereas the $\sigma_{\Delta K_{cb}}$ distribution shows a smaller spread than that of $\sigma_{\Delta K_c}$ (Fig. 5.2p and l).

For the second group, the MSLP anomaly field (Fig. 5.2e) shows an area of weak negative MSLP anomalies covering the central Atlantic Ocean and most of the Iberian Peninsula, except the extreme southwest. This pattern appears to be the result of transient low pressure systems (*transient pattern* hereafter), tracking from the middle of the Atlantic to central Europe, but with the centers distant from the Iberian Peninsula. These synoptic conditions bring a certain amount of cloudiness over the study area. Cloud cover (Fig. 5.2i) has intermediate values over that area. Accordingly, solar radiation statistics distributions show considerable spread at all stations (Fig. 5.2m and q), but with notable differences between GHI and DNI. In particular, the $\overline{K_c}$ distribution shows larger values and the $\sigma_{\Delta K_c}$ distribution less spread than their DNI counterparts.

The MSLP anomaly map associated with the third group (Fig. 5.2f) does not show any meaningful pattern. From a meteorological perspective, it appears that the samples of this group have associated a wide range of synoptic conditions that differs from those associated with the other groups but do not show any coherent structure. That is, these samples seem to share few (if any) characteristics. On the other hand, associated cloud cover (Fig. 5.2j) has intermediate-to-small values, except near the Atlantic coast and in mountainous areas (sea breezes and orographic clouds). This indicates the dominance of local circulations associated with this group (*local conditions* hereafter). The solar radiation statistics distributions (Fig. 5.2n and r) show a relatively large spread, but with a tendency toward large $\overline{K_c}$ and $\overline{K_{cb}}$ values and small $\sigma_{\Delta K_c}$ and $\sigma_{\Delta K_{cb}}$ values, except the Lisbon station. In this case, $\overline{K_c}$ and $\overline{K_{cb}}$ values are smaller and $\sigma_{\Delta K_c}$ and $\sigma_{\Delta K_{cb}}$ values are larger than those of the rest of the stations. This highlights a non-coherent solar resource variability, according to the cloudiness and the MSLP analyses results.

The corresponding MSLP field of the fourth group (Fig. 5.2g)

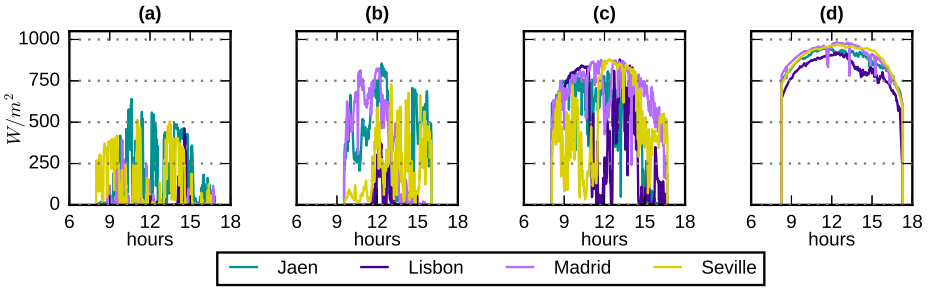

Figure 5.3: 1-minute DNI measurements at the four ground stations in a representative day of the groups 1, 2, 3, and 4 of the annual analysis shown in Figures (a), (b), (c), and (d), respectively.

shows a wide area of positive anomalies, covering the central Atlantic and western Europe, including the Iberian Peninsula. Anomalies are especially great in the northwestern Iberian Peninsula. This pattern, therefore, indicates high pressures over the entire study area (*high-pressure pattern* hereafter). The large values of the anomalies indicate that the *high-pressure pattern* allows for little local circulation and, therefore, less localized cloudiness variability. This is clearly evident in the very small cloud cover values associated with this group (Fig. 5.2k) and in the distributions of the solar irradiance statistics. In particular, the $\overline{K_c}$ and $\overline{K_{cb}}$ distributions are clearly tied to very large values, and the $\sigma_{\Delta K_c}$ and $\sigma_{\Delta K_{cb}}$ distributions to very small ones (Fig. 5.2o and s).

In order to easily understand the correspondence between the annual groups and the solar radiation variability, Figure 5.3 shows 1-minute DNI measurements at the four ground stations for days 4 October 2015 (Fig. 5.3a), 17 January 2016 (Fig. 5.3b), 9 October 2015 (Fig. 5.3c), and 10 March 2015 (Fig. 5.3d). These days are representative because all the 3-hours statistics belong to the same group. In particular, Figures 5.3a, 5.3b, 5.3c, and 5.3d are associated with the groups 1, 2, 3, and 4 of the annual analysis, respectively.

Figure 5.4 shows the solar PV power CF distribution (in %) corresponding to the Spanish part of the study area for the annual analysis. Figure 5.4a shows the CF distribution related to the *low-pressure pattern*. Relatively low values and a low spread are observed, similarly to the $\overline{K_c}$ distributions (Fig. 5.2l). The mean CF for this group is around 0.3. Larger spread is shown in the CF distribution associated with the *transient pattern* (Fig. 5.4b) in accordance with

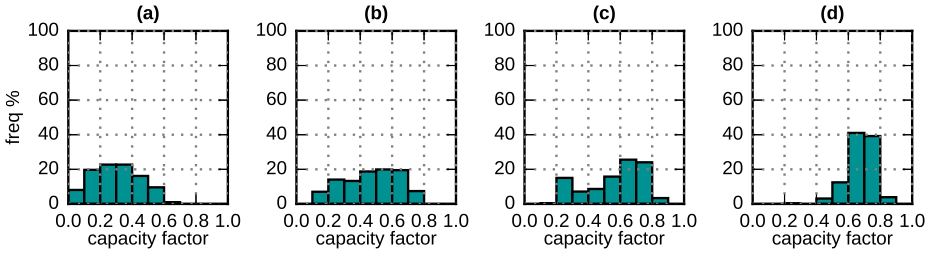

Figure 5.4: Normalized distributions of the aggregated solar PV power CFs in central and southwestern Spain. Figures (a), (b), (c), and (d) show the distributions associated with groups 1, 2, 3, and 4 (respectively) of the annual analysis.

Figure 5.2m. Notably, mean CF value is about 0.47, but values range from 0.1 to 0.8, highlighting the marked intra-day variability. Figure 5.4c shows the distribution of the CF registered under *local conditions*. The clearly non-normal distribution reflects the wide range of synoptic conditions and solar resource (Fig. 5.2n). Finally, the CF distribution related to the *high-pressure pattern* (Fig. 5.4d) is tied to very large values (mean value about 0.68), with reduced variability; as may be expected from the $\overline{K_c}$ distributions (Fig. 5.2o).

In addition to the annual analysis, seasonal analyses are performed and their results are shown in 5.6. In general, the seasonal analyses show modes of variability resembling those of the annual analysis. Nevertheless, there are notable seasonal differences, as expected by the prominent weather seasonality of the study area. Particularly, four groups are selected in the winter analysis, and the results are similar to the annual case but with larger cloudiness and lower values of the solar resource, especially for the case of the *low-pressure pattern*. Similarly, for the spring analysis, four groups are identified. Their associated weather patterns are similar to the corresponding annual patterns. However, mean CLM values are higher in the groups of *transient pattern* and *local conditions*, with a larger spread for *local conditions*. On the other hand, only three modes of variability are found for summer. These results can be explained based on the slight variability of synoptic conditions in the study area during this season. In summer, the most important mode is the *high-pressure pattern*. The *low-pressure pattern* exibits lower levels of mean CLM values and higher spread of the solar resource than in the annual case. Despite the fact that the *transient pattern* was

identified in the MSLP anomaly field of the remaining group, cloudiness and solar resource resemble those of *local conditions*. Lastly, the autumn analysis also reveals only three modes of variability. The *low-pressure pattern* and *high-pressure pattern* are identified. The remainder is similar to the *local conditions* group of the annual analysis but showing larger cloud cover and solar resource spread, thus making this group closer to the winter counterpart.

5.5 Conclusions and future work

In the present work, we analyze intra-day modes of variability of the solar resource in the central and southwestern Iberian Peninsula. To this end, a hierarchical cluster analysis, applied over radiation statistics from four stations within the study area, is used. Independent yearly and seasonal analyses are conducted. The number of groups is determined using the mean MSLP anomaly field of each group, ensuring that all the associated synoptic weather patterns are significantly different and meaningful in each cluster analysis. In addition, the modes of variability are analyzed in view of their associated cloudiness regimes, which are derived from satellite CLM estimates. Finally, solar PV yield production associated with each of the analyzed annual modes is assessed.

The yearly analysis shows the existence of four main modes of intra-day variability of the solar resource in the study area associated to the following synoptic weather patterns: *low-pressure pattern* (small values of the solar resource and very weak intra-day variability), *transient pattern* (very strong intra-day variability of the solar resource), *local conditions* (substantial spatial and temporal variability), *high-pressure pattern* (very large values of the solar resource with very weak intra-day variability). These findings are supported by the agreement between the results of the different and independent databases analyzed, such as cloudiness and solar PV power yield annual distributions.

In spite of the important meteorological seasonality of the study area, the modes found in the seasonal analyses are similar to those of the annual analysis. Furthermore, all modes prove to be spatially coherent over the entire study area, except one. The latter accounts for weak pressure anomalies, allowing local weather conditions to

arise. The impact of this local dominance depends on the period of analysis.

In order to apply the presented method in other areas, some issues should be taken into account. First, the set of stations should be representative of the main factors driving the variability of the solar radiation in the area (synoptic conditions, topographic features). Moreover, the availability of 1-minute GHI and DNI measurements could be scarce, especially in case of DNI. Furthermore, due to the fact that a particular level of significance is required in the included statistical methods, the length of the database must be large enough.

The silhouette coeficient metric and, especially, the distance between clusters at each aggregation step, suggest a higher number of modes than the finally selected modes based on the significance of the weather patterns. Nevertheless, these patterns, obtained at spatiotemporal scales relevant to grid integration, are easily identifiable by the power system operators. Thus, the results of the present analyses can contribute to better integration of solar energy in the study area. In particular, by nowcasting the dominant weather pattern, the reliability of the estimates of the solar power and its spatiotemporal variability in the coming hours can be improved. Therefore, the reserves requirement of the electric system and the associated costs would be reduced. Future work will examine temporal persistence of the different weather regimes identified in this study and the transition probability between them. This may allow an increase, beyond 3 hours, of the forecasting horizon with reliable predictions of the solar power variability. Also, we will analyze the performance of various solar radiation forecast models for the different modes of variability, allowing the development of optimal ensemble models.

Acknowledgements

The authors are supported by the Spanish Ministry of Economy and Competitiveness, project ENE2014-56126-C2-1-R (http://prosol.uc3m.es), FEDER funds and also by the Junta de Andalucía (research group TEP-220). The authors are indebted to the National Centers for Environmental Prediction (NCEP), EUMETSAT, Red Eléctrica de España, Faculdade de Ciências 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. The authors thank the support of the MATRAS group members.

5.6 Seasonal analyses (appendix)

5.6.1 Winter

Similar to the annual case, Figure 5.5 summarizes results of the winter analysis, which accounts for 161 samples (Table 5.2). Four groups are selected. The silhouette coefficient metric (Fig. 5.5a) clearly shows that groups 1 and 4 are the most consistent. In contrast, groups 2 and 3 show many sets of values weakly associated with their groups. Based on Figures 5.5b and c, a plausible number of groups might be higher than the four groups ultimately selected.

The associated MSLP pattern of the first group (Fig. 5.5d) closely resembles the *low-pressure pattern* of the annual analysis (Fig. 5.2d), but with a northward shift. This pattern is associated with low-pressure centers located over Central Europe, usually related to thick clouds over the study area. This is clearly observed in Figure 5.5h, which shows large CLM values over the Iberian Peninsula. Small values of solar resource and low variability are found at the four stations (Fig. 5.5l and p), even more enhanced than those in the annual analysis (Fig. 5.2l and p).

For the second group, the associated MSLP field (Fig. 5.5e) shows an extensive area of relatively weak negative anomalies centered over France, which includes the northern half of the Iberian Peninsula. The MSLP anomaly pattern over the Iberian Peninsula seems similar to the *transient pattern* of the annual analysis (Fig. 5.2e). Relatively large cloud cover values are associated with this group (Fig. 5.5i), and the solar radiation statistics distributions (Fig 5.5m and q) show a large spread across the four stations.

The corresponding MSLP map field of the third group (Fig. 5.5f) shows an area of weak positive anomalies covering the entire Iberian Peninsula. The associated mean CLM map (Fig. 5.5j) displays small values over the study area except for its west, including the Lisbon area, with intermediate values. Solar radiation statistics distributions (Fig. 5.5n and r) are clearly tied to large $\overline{K_c}$ and $\overline{K_{cb}}$ and small

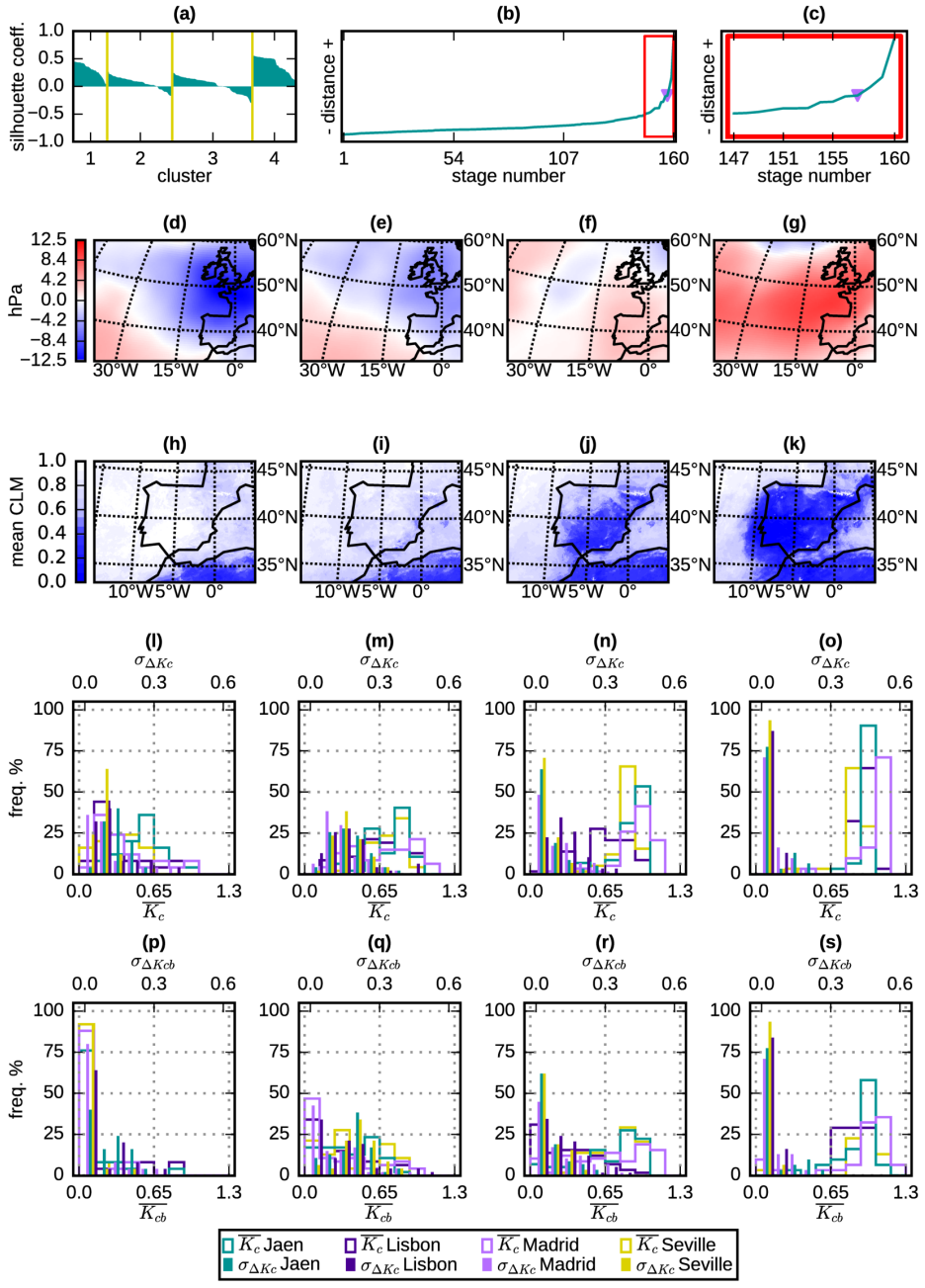

Figure 5.5: As in Figure 5.2 but for the winter seasonal analysis.

$\sigma_{\Delta K_c}$ and $\sigma_{\Delta K_{cb}}$ values for all stations except the Lisbon one. At that station, there is greater spread, in agreement with the CLM map. It appears that this kind of weak synoptic situation allows the manifestation of local climate phenomena (*local conditions*), e.g., the development of convective circulations near the coast. This produces greater variability of solar radiation that inland where no source of humidity is available. Note that this is the most common pattern during winter.

For the fourth group, the corresponding MSLP field (Fig. 5.5g) closely resembles the *high-pressure pattern* of the annual analysis (Fig. 5.2g), but anomalies are greater. Consequently, the mean CLM map (Fig. 5.5k) shows very low cloud cover over the study area, and the radiation statistics distributions (Fig. 5.5o and s) exhibit quite large $\overline{K_c}$ and $\overline{K_{cb}}$ and considerably small $\sigma_{\Delta K_c}$ and $\sigma_{\Delta K_{cb}}$ values, with little spread.

5.6.2 Spring

Figure 5.6 shows the results of the spring analysis. According to Table 5.2, it includes 368 samples, finally divided into four groups. Regarding the silhouette coefficients (Fig. 5.6a), groups 1 and 4 are clearly the most consistent, as opposed to groups 2 and 3. Based on Figures 5.6b and c the number of intuitive groups might not be far from the final selected number.

The MSLP anomaly associated with the first group (Fig 5.6d) is very similar to the *low-pressure pattern* (Fig. 5.2d). Accordingly, high mean CLM values are observed over the entire study area (Fig. 5.6h) and solar radiation statistics distributions (Figures 5.6l and p) show small mean values and low variability.

For the second group, the MSLP field (Fig 5.6e) resembles the *transient pattern* (Fig. 5.2e), but displaced southward. Consequently, the mean CLM map (Fig. 5.6i) shows intermediate-to-high values over the study area. Solar resource distribution values (Fig. 5.6m and q) are very similar to those observed in the annual case (Fig. 5.2m and q).

The associated MSLP map of the third group (Fig. 5.6f) do not show any meaningful patterns. Only in the central Atlantic, a neg-

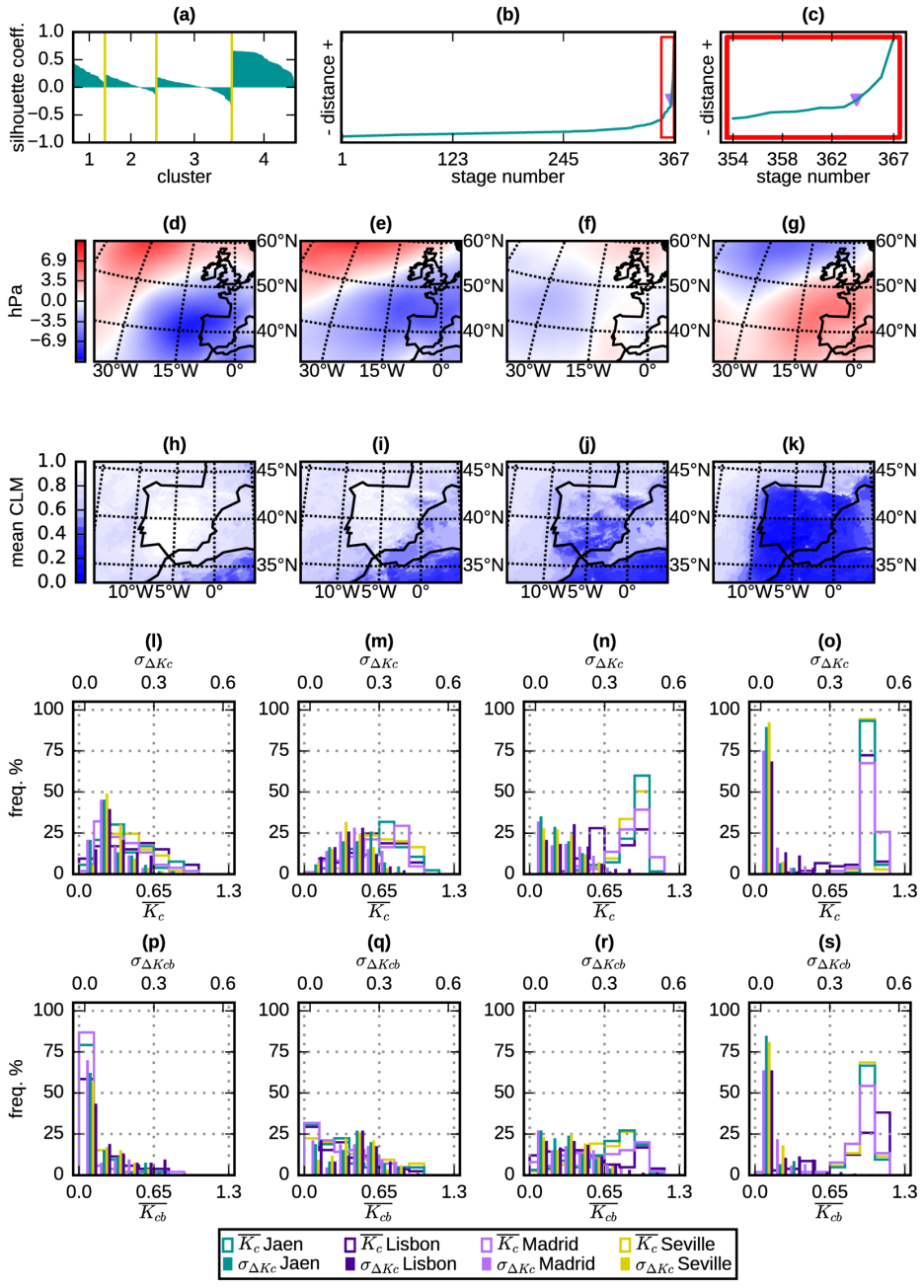

Figure 5.6: As in Figure 5.2 but for the spring seasonal analysis.

ative anomaly is observed. This group seems to account for a set of synoptic patterns observed during spring that do not show any coherence. This is a result of the highly variable synoptic conditions that characterize the study area in this season. During this part of the year, a clear weakening of the prevailing westerly winds occurs. As a consequence, a wide range of synoptic patterns, many associated with easterly circulations that rarely occur during winter and summer, are often observed. The CLM map (Fig. 5.6j) shows intermediate values over the study area, but with a high spatial variability. Solar radiation statistics display a considerable spread (the maximum observed for all the groups) at all the stations (Fig. 5.6n and r).

The MSLP field of the fourth group (Fig. 5.6g) closely resembles the *high-pressure pattern* (Fig. 5.2g). As highlighted earlier, this pattern accounts for high pressures over the entire study area. Accordingly, low cloud cover is exhibited (Fig 5.6k), and the radiation statistics distributions (Fig. 5.6o and s) indicate high values of the resource and a limited variability at the four stations.

5.6.3 Summer

A total of 316 samples are used for the summer analysis (Table 5.2) and its results are summarized in Figure 5.7. Unlike the previous analyses, 3 instead of 4 groups are selected. This fact can be explained based on the low variability of synoptic conditions over the study area during summer, caused by the presence and prominence of the Azores anticyclone during this season. The silhouette coefficients (Fig. 5.7a) visibly show the relevance of group 3, the most consistent and which accounts for most samples. As described below, this cluster corresponds to stable conditions in the study area. The other two groups show less consistency. The selected number of groups may not be in the intuitive range regarding the distance between groups (Fig. 5.7b and c).

For the first group, the MSLP pattern (Fig. 5.7d) is similar to the *low-pressure pattern* (Fig. 5.2d). The associated cloudiness displays a relatively higher spatial variability (Fig. 5.7g) than in the annual case (Fig. 5.2h). As a consequence, solar radiation statistics distributions (Fig. 5.7j and m) also show a higher spread than their annual counterpart (Fig. 5.2l and p).

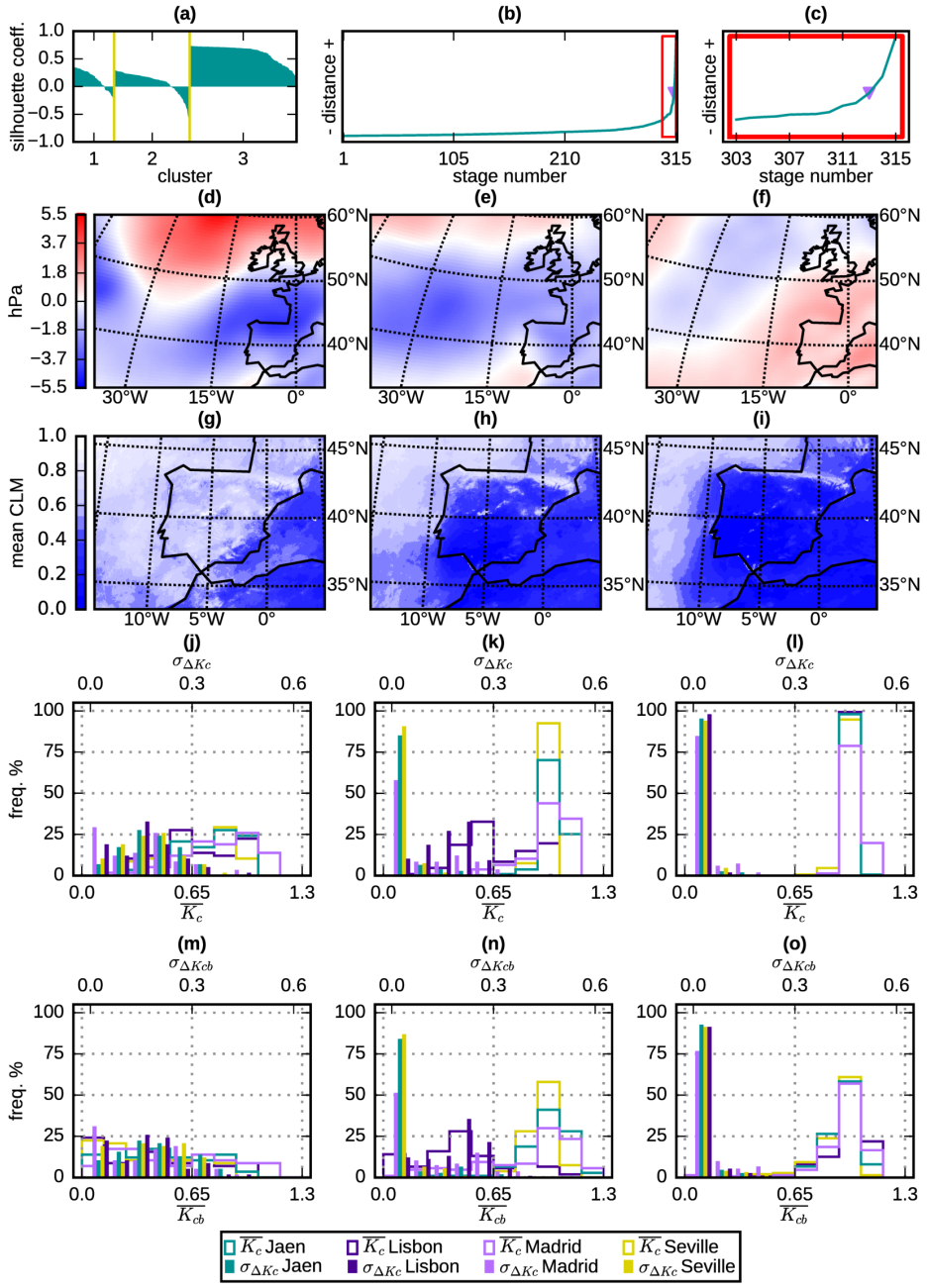

Figure 5.7: As in Figure 5.2 but for the summer seasonal analysis.

The MSLP anomaly pattern of the second group (Fig. 5.7e) closely resembles the *transient pattern* (Fig. 5.2e). However, the associated cloudiness (Fig. 5.7h) is weaker than in the annual case, except over the north-west, including the Lisbon area, similar to that of *local conditions* (Fig. 5.2j). In addition, the radiation statistics distributions reflect this spatial variability of cloudiness. Figures 5.7k and n are closer to Figures 5.2n and r, respectively, but the Jaen, Madrid, and Seville stations exhibit lower spread.

For the third group, the associated MSLP pattern (Fig. 5.7f) is like the *high-pressure pattern* (Fig. 5.2g). As expected, very low cloudiness is observed over the entire study area (Fig. 5.7i). Solar radiation statistics distribution (Figures 5.7j and o) are very similar to the corresponding annual ones (Figures 5.2o and s).

5.6.4 Autumn

Figure 5.8 represents results of the autumn analysis, which accounts for 260 samples (Table 5.2). Three groups are identified. According to the silhouette coefficient metric (Fig. 5.8a), group 3 is the most consistent. The other two groups show less consistency. Figures 5.8b and c reveal that the selected number of groups might be lower than those in the intuitive range of candidates.

The MSLP pattern of the first group (Fig 5.8d) resembles the *low-pressure pattern* (Fig. 5.2d). High CLM values are observed associated with this pattern (Fig. 5.8g). Solar radiation statistics distribution (Fig. 5.8j and m) are similar to the corresponding yearly distributions (Fig. 5.2l and p). Therefore, the same interpretation applies here.

For the second group, the MSLP pattern (Fig. 5.8e) shows a positive, but weak, center of anomalies located over the southwestern Iberian Peninsula. This pattern resembles the MSLP map associated with the *local conditions* group of the winter analysis (Fig. 5.5f). The cloud cover map (Fig. 5.8h) is also alike Figure 5.5j. This weak high pressure center produces relatively large values, and high variability, of the solar resource over the study area (Fig. 5.8k and n). As in the case of the winter analysis, this is the most common pattern during autumn.

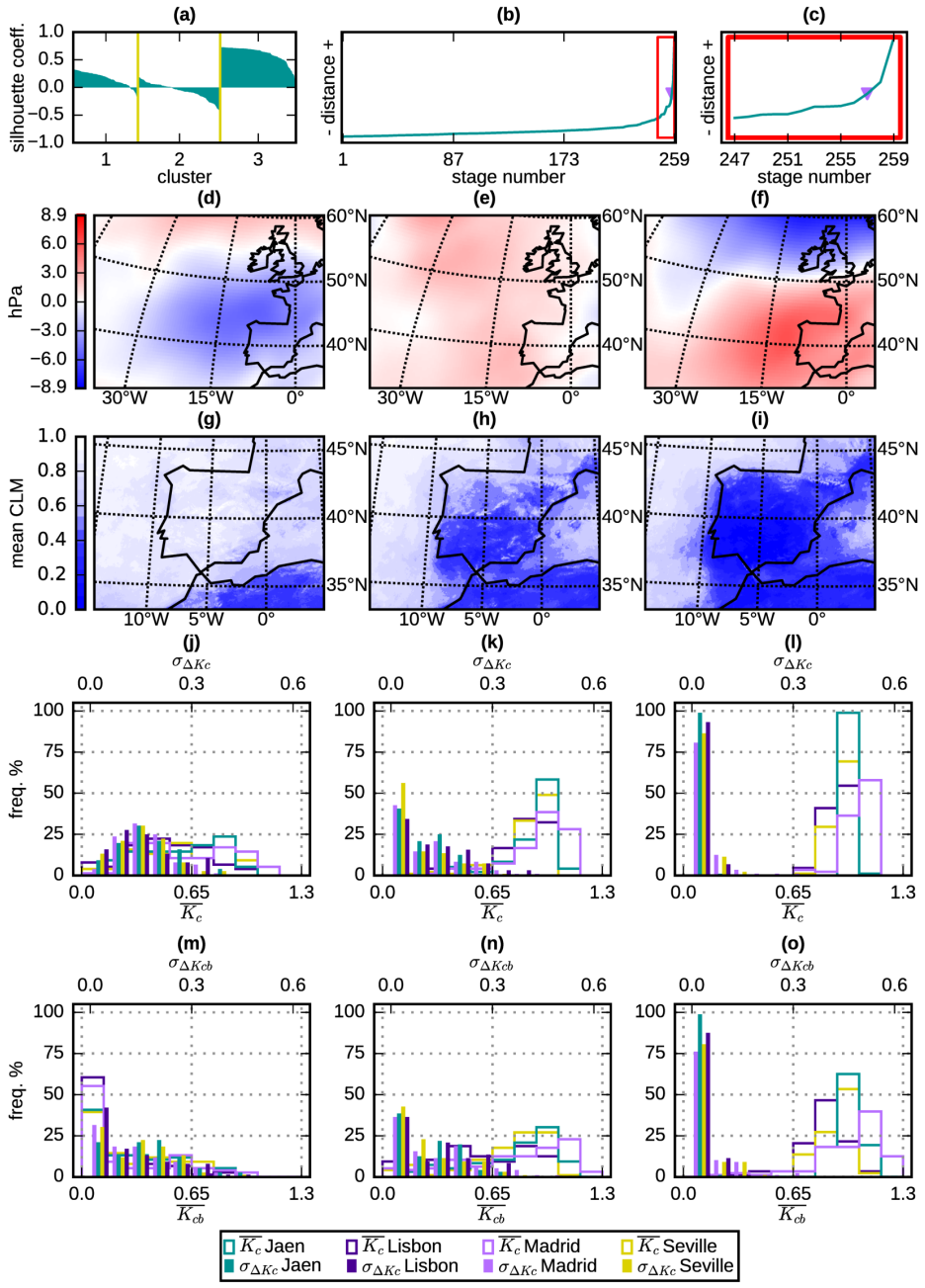

Figure 5.8: As in Figure 5.2 but for the autumn seasonal analysis.

Finally, the MSLP pattern associated with the third group (Fig. 5.8f) is very similar to the *high-pressure pattern* (Fig. 5.2g). Also the cloud cover map (Fig. 5.8i) and the solar radiation statistics distribution (Fig. 5.8l and o) resemble those of the annual analysis (Fig. 5.2k, o and s, respectively). Hence, the same features are observed here.

CHAPTER 6

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

Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Huertas-Tato, J., Aler-Mur, R., Galvan-León, I., and Pozo-Vázquez, D. (2019). A short-term solar radiation forecasting system for the Iberian Peninsula. Part 1: Models description and performance assessment. *Submitted for publication in Solar Energy, manuscript number SE-D-19-01526*

6.1 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., 2017a), 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 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 and Coauthors, 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. (2012a). 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 6.2 presents the study area and describes the ground and remote sensing dataset used. Section 6.3 introduces the four forecasting models. Section 6.4 describes the metrics and procedure involved in the models assessment. In section 6.5 the results are presented and discussed and, finally, in section 6.6 a summary of the main results and some conclusions are provided. Additional analyses are discussed in 6.7 and 6.8.

6.2 Datasets

6.2.1 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 (Figure 6.1) 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 ± 2.5% for the DNI and about ± 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., 2002b).

Figure 6.1: Shaded relief map of the study area and location of the evaluation stations (garnet dots).

Changes 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 (Figure 6.1). 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., 2002b; Rodríguez-Benítez et al., 2018).

6.2.2 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 6.3.4) 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).

6.3 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.

6.3.1 Smart Persistence

Smart Persistence model is used as the reference.

$$I(t) = \frac{I_0}{I_{clear-sky}(t_0)} \cdot I_{clear-sky}(t), \tag{6.1}$$

where I_0 is the 1-minute time resolution measured irradiance at the initial time (t_0), $I_{clear-sky}(t_0)$ is the corresponding clear sky irra-

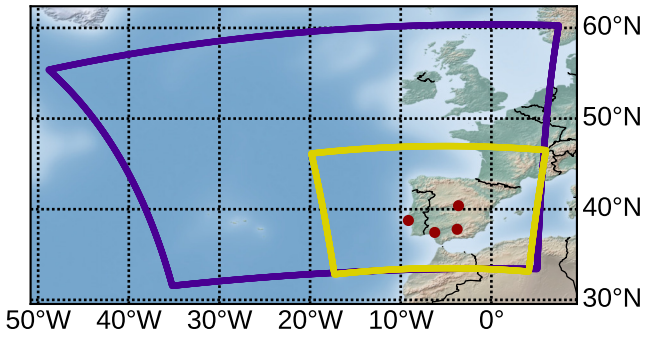

Figure 6.2: Domains of the CIADCast and WRF-Solar (yellow line), and Satellite-based model (purple line). The garnet dots indicate the location of the evaluation stations.

diance 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 $I_{clear-sky}$.

6.3.2 Satellite

The CI images are used to provide satellite-based GHI and DNI estimates, as described in section 6.2.2. 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 6.2.2. Figure 6.2 represents the area of the MSG images analyzed to derive the CMV fields.

6.3.3 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 minutes, 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 seconds of model time-step) and 50 vertical levels. The domain is drawn in Figure 6.2. Initial and boundary conditions are taken from the National Centers for Environmental Prediction (NCEP) GFS [0.5 Deg] dataset (National Centers for Environmental Prediction, 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.

6.3.4 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 6.3.3. 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 6.2.2.

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.

6.4 Evaluation procedure

Evaluation of the forecasts with ground data is carried out at 15 minutes time steps, according to the time when the MSG images were obtained for each station, and at 1-minute time resolution. Table 6.1 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.

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

Table 6.1: 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

BIAS error defined as:

$$\mathrm{RMSE}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right)^2} \tag{6.2}$$

$$\mathrm{FS}(t) = 1 - \frac{\mathrm{RMSE}_{\mathrm{forecast}(t)}}{\mathrm{RMSE}_{\mathrm{SmartPersistence}(t)}} \tag{6.3}$$

$$\mathrm{rRMSE}(t) = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right)^2}}{\frac{1}{N}\sum_{i=1}^{N} I_{\mathrm{measured}(t),i}} \cdot 100 \tag{6.4}$$

$$\mathrm{rMAE}(t) = \frac{\frac{1}{N}\sum_{i=1}^{N}\left|I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right|}{\frac{1}{N}\sum_{i=1}^{N} I_{\mathrm{measured}(t),i}} \cdot 100 \tag{6.5}$$

$$\mathrm{BIAS}(t) = \frac{1}{N}\sum_{i=1}^{N}\left(I_{\mathrm{forecast}(t),i} - I_{\mathrm{measured}(t),i}\right), \tag{6.6}$$

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 6.5.2), 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, 2013a). As opposed to deterministic forecasts (assessed in section 6.5.1), 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:

$$\sigma_{\Delta K_c} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\Delta K_{c,i} - \overline{\Delta K_c}\right)^2}, \tag{6.7}$$

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 6.5.4, 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., 2016a). 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.

6.5 Results

6.5.1 General evaluation

Figure 6.3 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 Figure 6.3) 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,

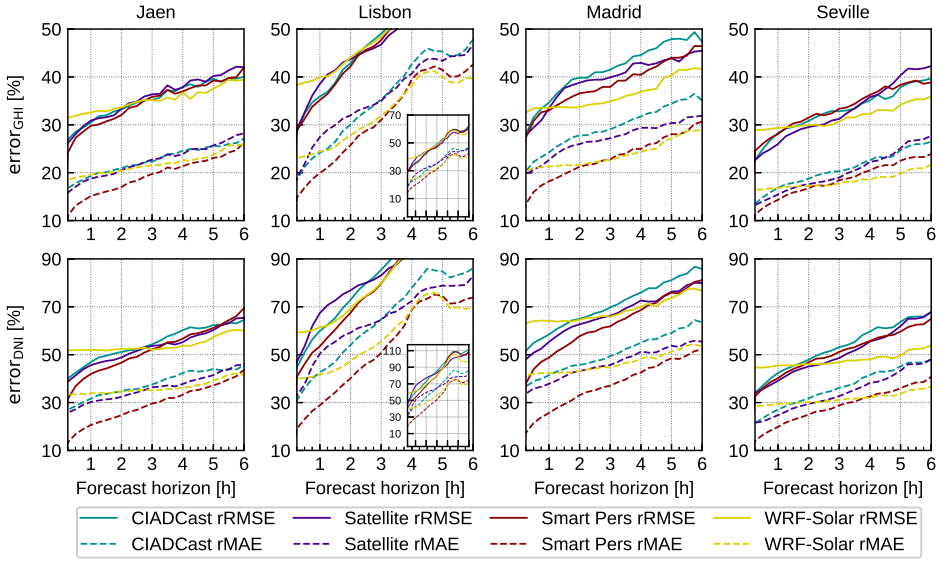

Figure 6.3: rRMSE (solid lines) and rMAE (dashed lines), depending on the forecasting horizon, for predictions of GHI (upper row) and DNI (bottom row). Values (in %) are displayed for CIADCast, Satellite, Smart Persistence, and WRF-Solar model with a turquoise, purple, garnet, and yellow line, respectively. 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.

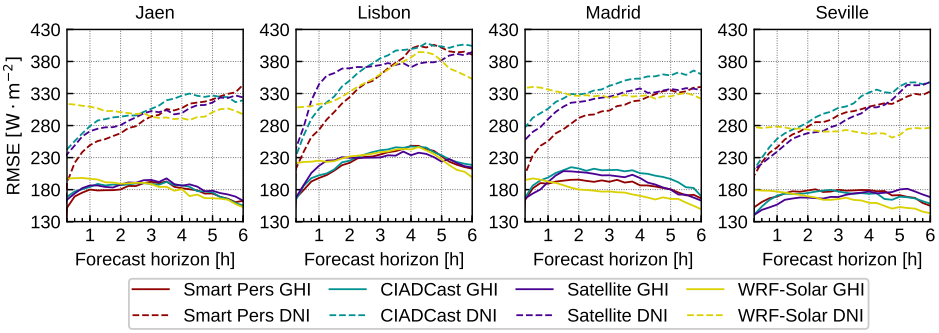

Figure 6.4: RMSE values for GHI (solid lines) and DNI (dashed lines) forecasts, depending on the forecasting horizon. The color code is the same used in Figure 6.3.

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 Figure 6.3) 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.

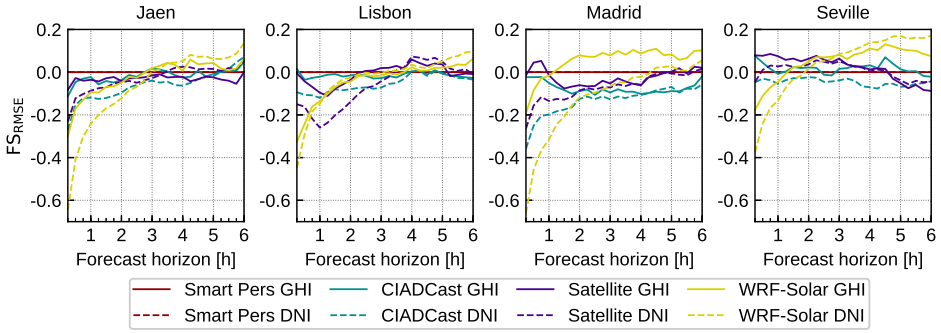

Figure 6.5: Forecast skills (FS) of the GHI (solid lines) and DNI (dashed lines), depending on the forecasting horizon. The color code is the same used in Figure 6.3.

Figure 6.4 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 (Figure 6.3). 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. Figure 6.5 shows the results for both GHI and DNI forecasts. As may be expected based on the analysis showed at Figures 6.3 and 6.4, 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. 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

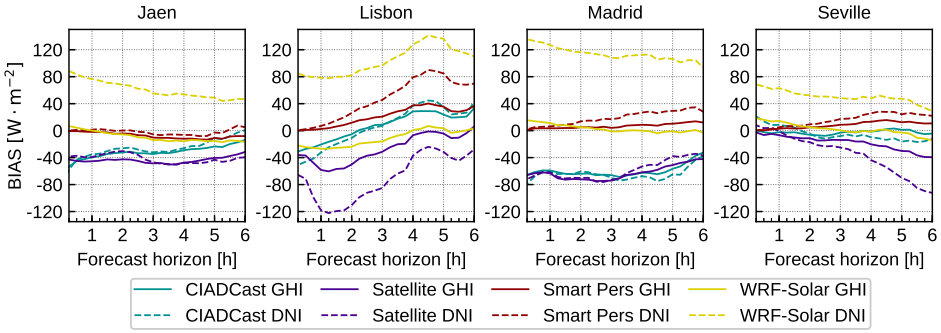

Figure 6.6: As in Figure 6.4 but for the BIAS score.

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, Figure 6.6 shows the BIAS error values for both GHI and DNI. GHI BIAS values are relatively low and negative (between 20 and -80 $W \cdot m^{-2}$) 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 6.2.1. 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 6.5.4.

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 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 (6.7). For the GHI (Figure 6.12) 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 (Figure 6.13). 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 6.5.3. 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 hours, depending of the station. These conclusions are also supported by the analysis of the Forecast Skill scores (Figure 6.14). Finally, for spring and winter, the WRF-Solar shows a superior performance, especially for the DNI (Figure 6.14).

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.

6.5.2 Variability forecasts evaluation

Figure 6.7 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.

6.5.3 Forecasts dependence on the time of day

Figure 6.8 shows the rRMSE values depending on the time of the day for two representative forecasting horizons: 1 and 4 h. Figure 6.9 shows the BIAS error for the 4 h horizon.

According to Figures 6.8 and 6.9, no significant dependence on the time of the day is observed, except at the Lisbon station (Figure 6.8). 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 max-

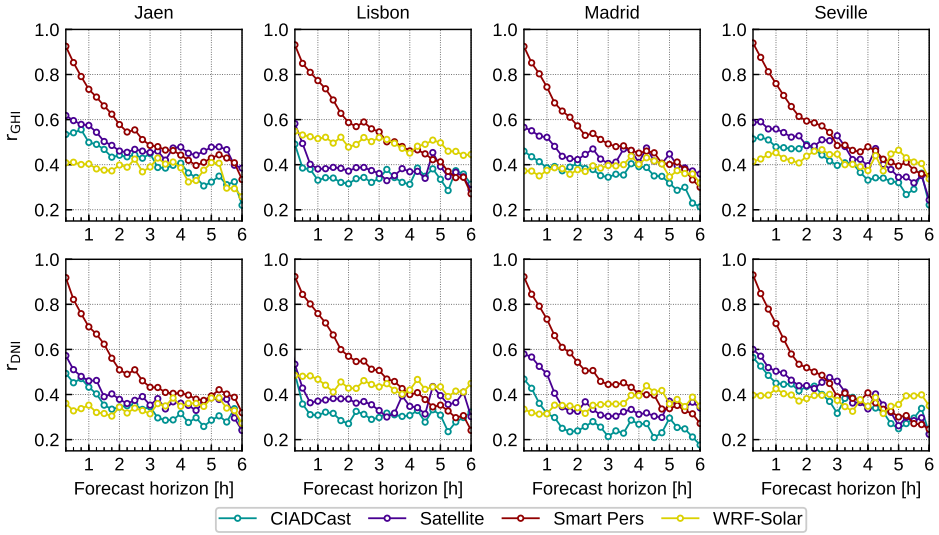

Figure 6.7: Correlation coefficient, r, between the forecasted and observed solar radiation nominal variability index, as function of the forecasting horizon. Color code, rows, and columns structure as in Figure 6.3.

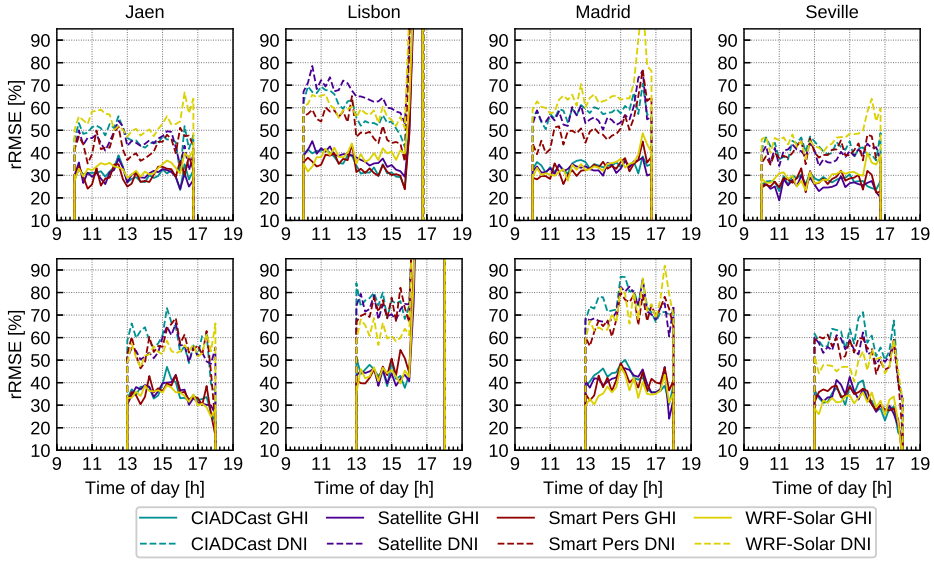

Figure 6.8: Upper row, rRMSE values (GHI/DNI solid/dashed lines) for 1 h ahead forecasts as function of the time of the day. Lower row, as upper row but for 4 h ahead forecasts. Color code as in Figure 6.3.

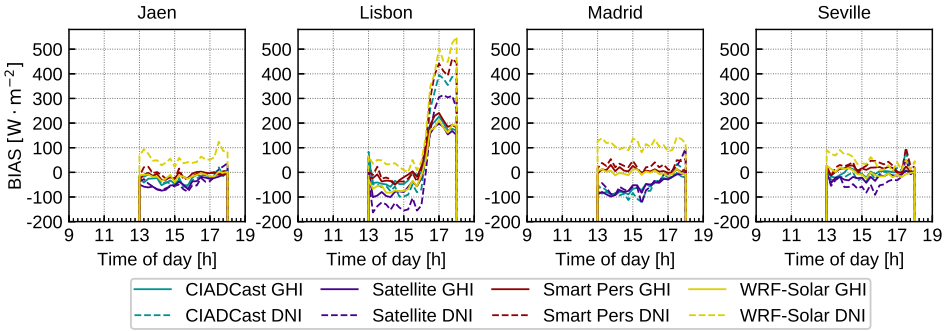

Figure 6.9: BIAS values (GHI/DNI solid/dashed lines) for 4 h ahead forecasts as a function of the time of the day. Color code as in Figure 6.3.

imum in the afternoon hours. As observed in Figure 6.9, 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 (Figure 6.8). 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.

6.5.4 Forecasts dependence on the synoptic weather conditions

Figures 6.10 and 6.11 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

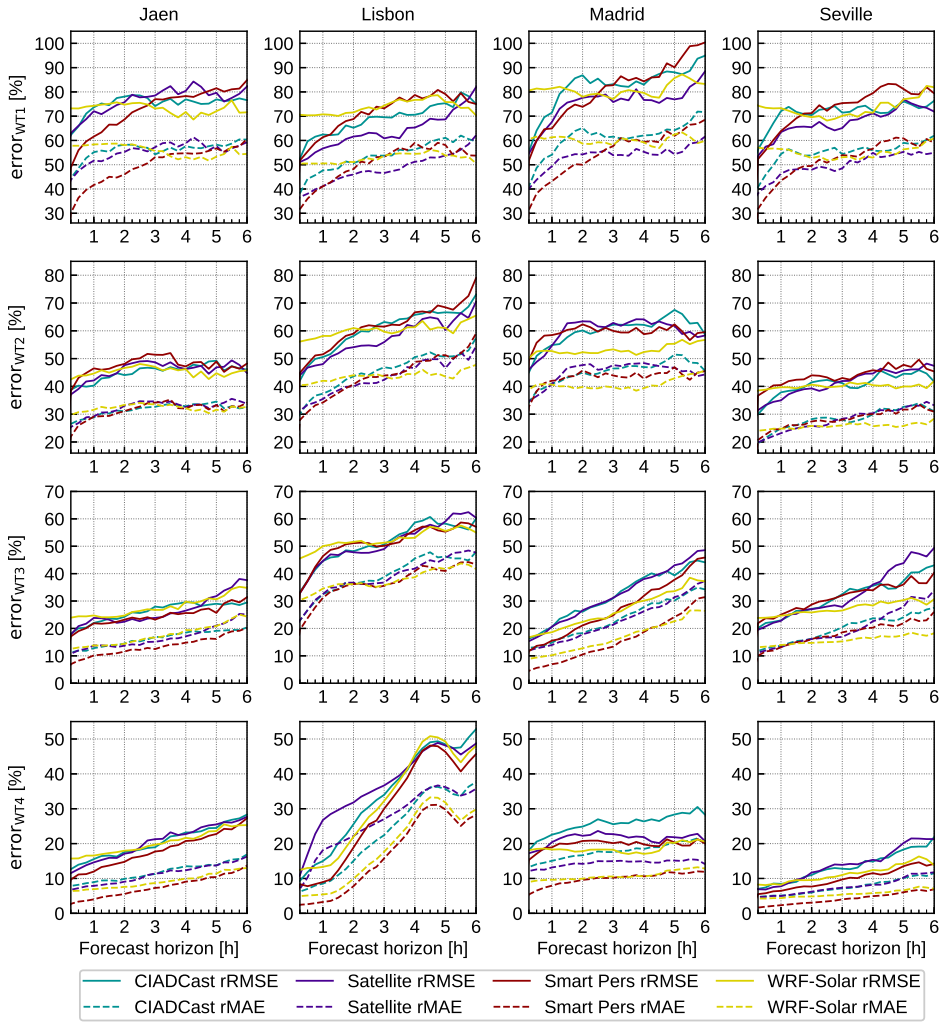

Figure 6.10: rRMSE (solid lines) and rMAE (dashed lines) depending on the forecasting horizon, for predictions of GHI classified by the weather type observed at the time when the forecasts were issued. Each row represents a weather type. Results for Jaen, Lisbon, Madrid and Seville are displayed at the first, second, third and fourth column, respectively. Colors code as in Figure 6.3. The vertical range of values (%) varies with the WTs, but is kept the same at the four stations, for the sake of comparison.

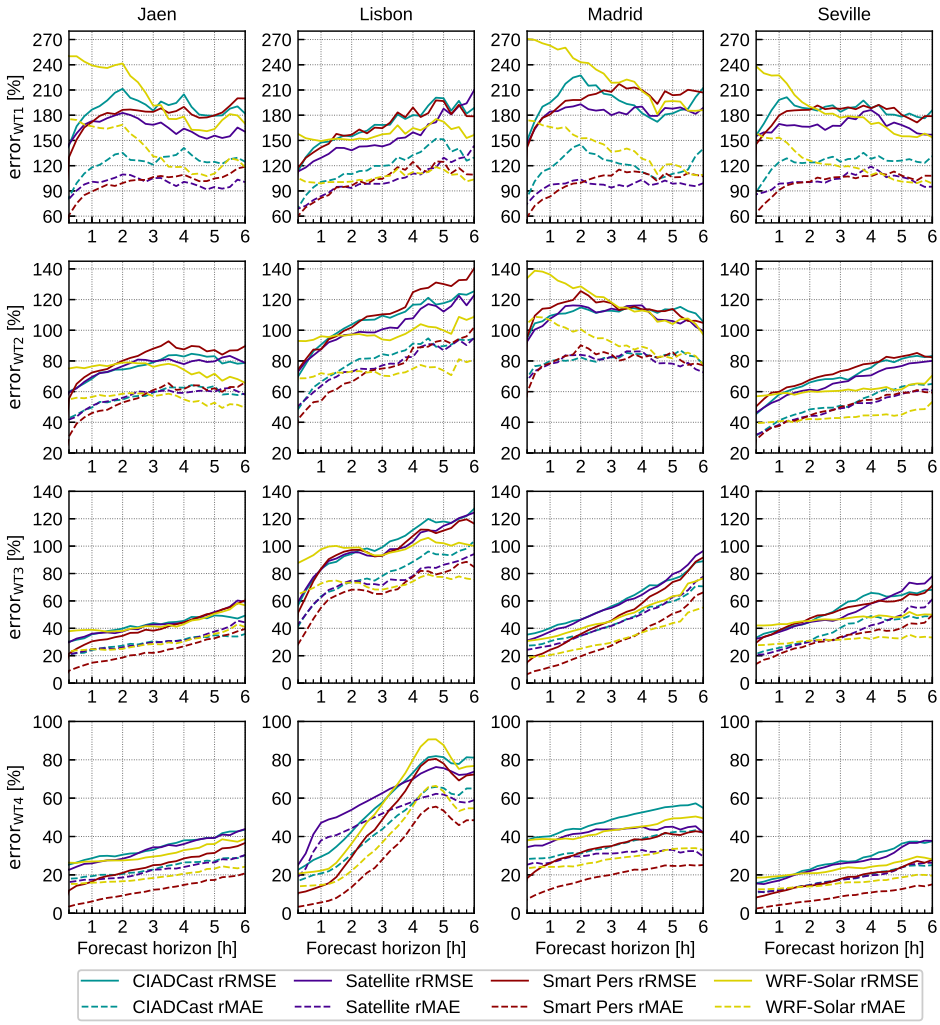

Figure 6.11: As in Figure 6.10 but for the DNI.

a 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 6.9, 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 (Figure 6.10 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 Figures 6.10 and 6.11 shows the results for GHI and DNI, respectively, for WT 3 (relative high pressures). Forecast error values 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 (Figure 6.10, 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 (Figure 6.11, 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 Figures 6.10 and 6.11 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, (Figure 6.10 last row), Smart Persistence is the best model for all stations and fore-

casting 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 (Figure 6.11, 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 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 (6.8). 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 hour. 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.

6.6 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.

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

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 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.

6.7 Seasonal analyses (appendix)

Figures 6.12 and 6.13 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.

Regarding the relative performance of the models, for the GHI (Figure 6.12) and during winter, WRF-Solar shows a notable performance. Notably, WRF-Solar presents the lowest rRMSE values at lead time higher than 1.5 hour in Lisbon and Madrid and higher than 3 hours 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 hour. 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 (Figure 6.13) 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.

Figure 6.14 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,

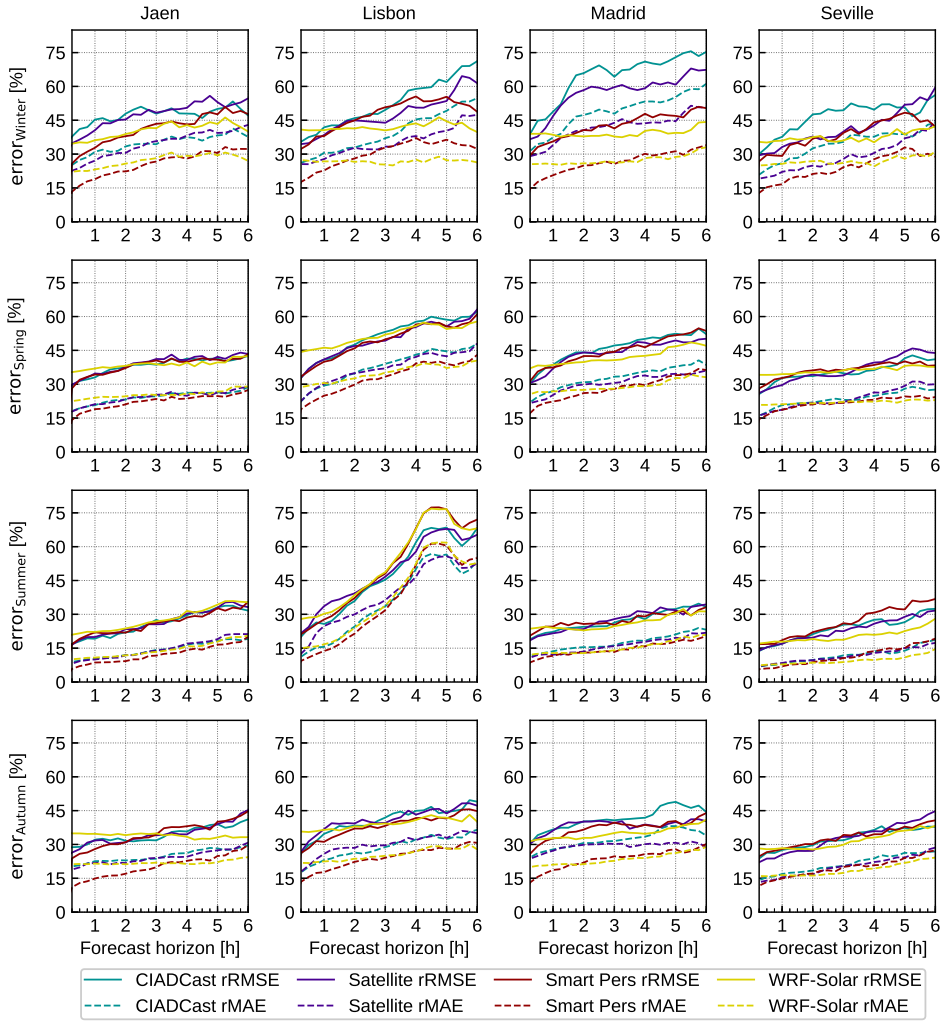

Figure 6.12: rRMSE (solid lines) and rMAE (dashed lines) depending on the forecasting horizon, for predictions of GHI classified according to the season. Results for Jaen, Lisbon, Madrid and Seville are displayed at the first, second, third, and fourth column, respectively. Colors code as in Figure 6.3.

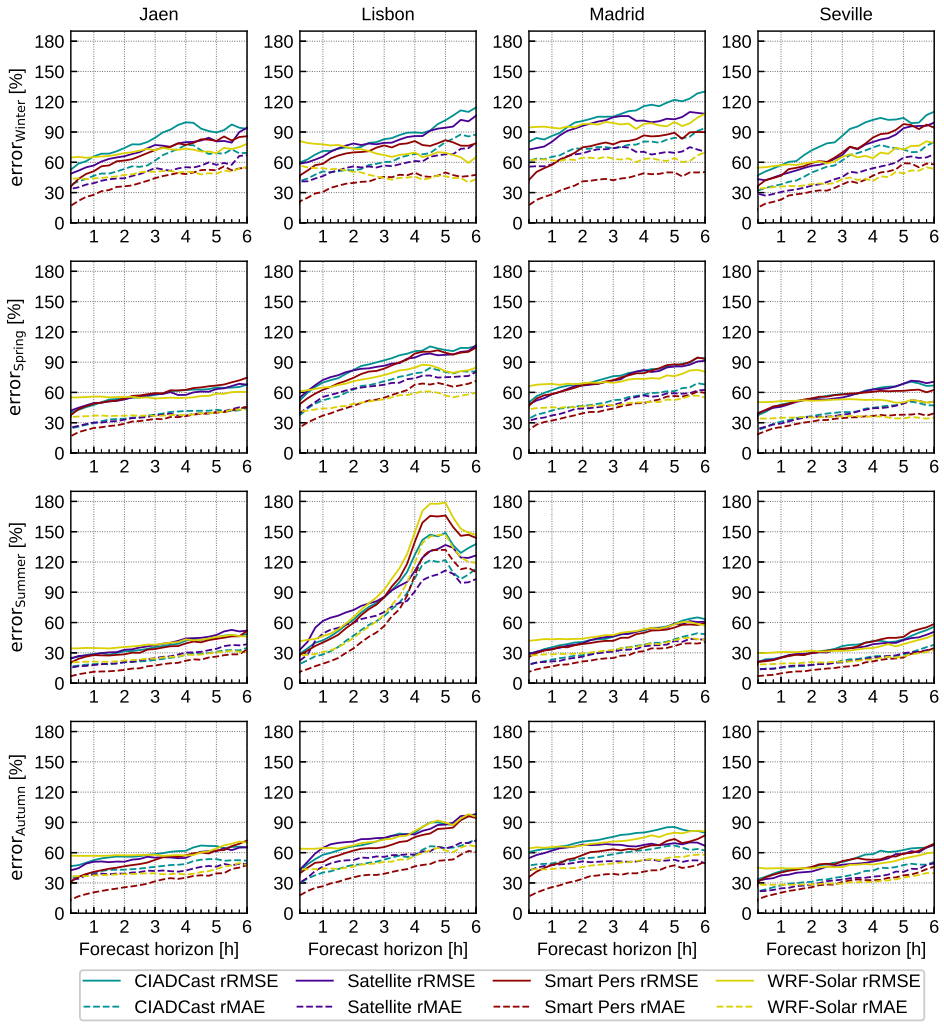

Figure 6.13: As in figure 6.12 but for DNI.

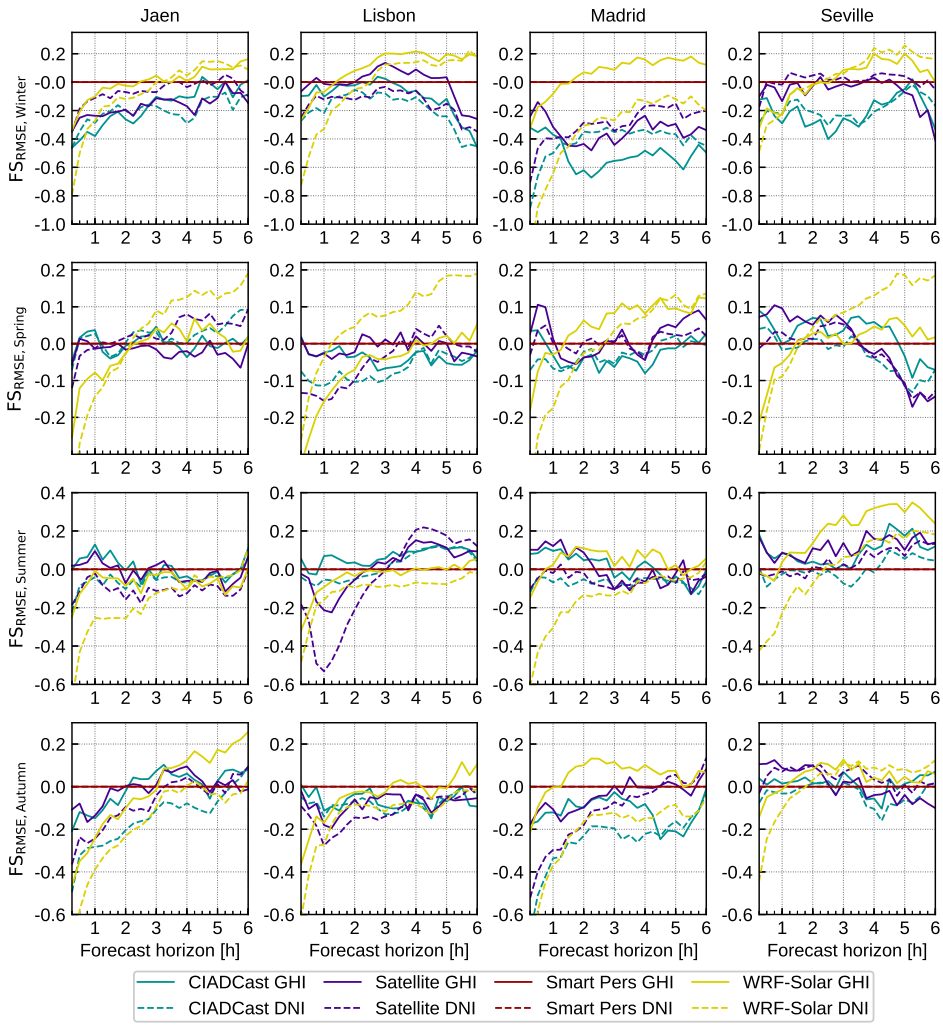

Figure 6.14: Forecast skills (FS) of the GHI (solid lines) and DNI (dashed lines), depending on the forecasting horizon, classified according to the season. The color code is the same used in Figure 6.3.

WRF-Solar shows a notable performance for all the stations and for the GHI, specially for lead times greater than 3 hours. 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.

6.8 Daily k_t distribution analysis (appendix)

Figures 6.15 and 6.16 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 bellow 0.4 as overcast conditions.

For clear-sky conditions and GHI (Figure 6.15 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 hours-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 (Figure 6.16 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 (Figures 6.15 and 6.16, 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% (Figure 6.15). For the DNI case (Figure 6.16) 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 (Figure 6.15). 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 hours, when WRF-Solar performs similarly.

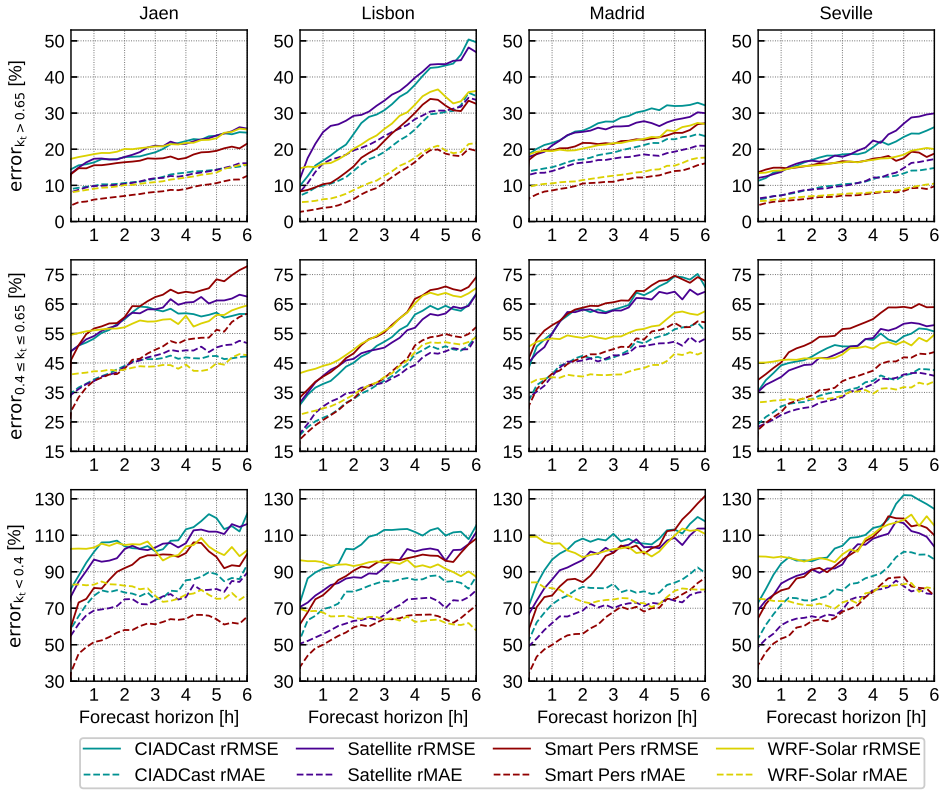

Figure 6.15: rRMSE (solid lines) and rMAE (dashed lines) depending on the forecasting horizon, for predictions of GHI classified according to the k_t distribution observed at the day when the forecasts were issued. First row represents $k_t > 0.65$ (clear sky), second row shows $0.4 \leq k_t \leq 0.65$ (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 Figure 6.3. 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.

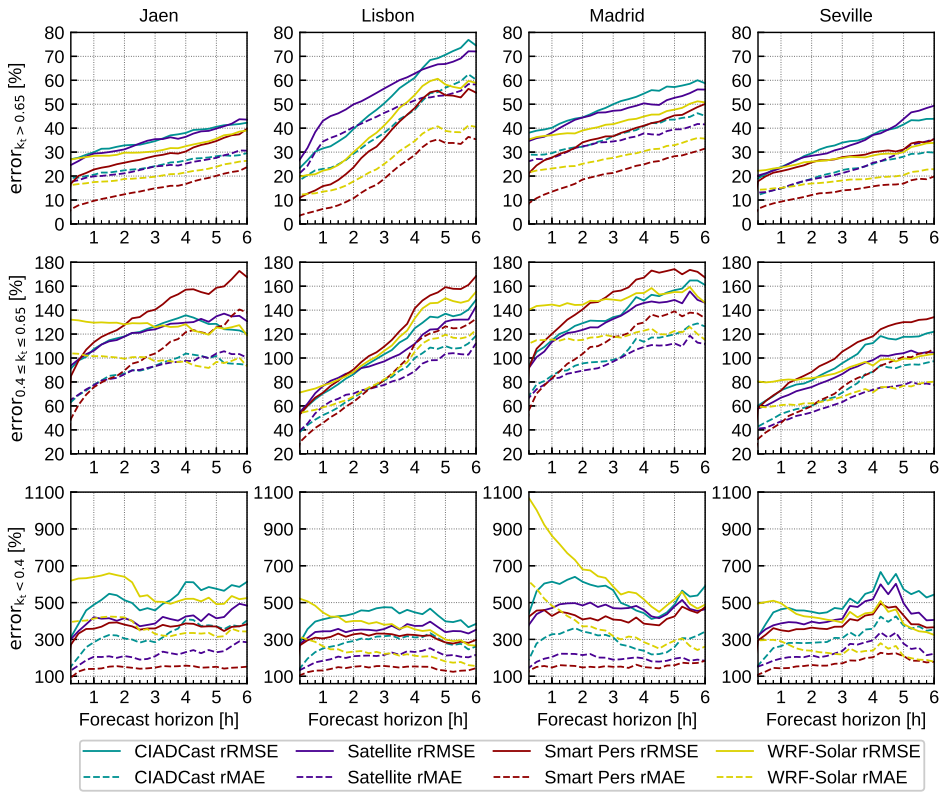

Figure 6.16: As in Figure 6.15 but for the DNI.

For overcast conditions (Figures 6.15 and 6.16 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 (Figure 6.15). For the GHI, rMAE values in the range 50% to 90% and rRMSE values between 80% and 120% are observed (Figure 6.15). For DNI (Figure 6.16), error values are considerable higher.

CHAPTER 7

A short-term solar radiation forecasting system for the Iberian Peninsula. Part 2: Model blending approaches based on machine learning

Huertas-Tato, J., Aler-Mur, R., Galvan-León, I., Rodríguez-Benítez, F. J., Arbizu-Barrena, and Pozo-Vázquez, D. (2019). A short-term solar radiation forecasting system for the Iberian Peninsula. Part 2: Model blending approaches based on machine learning. *Submitted for publication in Solar Energy, manuscript number SE-D-19-01525*

7.1 Introduction

In the companion paper (Part 1), four different short term solar (GHI and DNI) forecasting models (Satellite-based model, WRF-Solar, Smart Persistence and CIADCast) were presented and their performance was evaluated at four stations located in the Iberian Peninsula. Results showed that those models are complementary because no model outperforms the rest for all forecasting horizon, location or variable (GHI and DNI).

In this second part we aim to explore the benefits obtained by de-

veloping optimal integration (blending) models, using Machine Learning techniques. The question arises of whether an optimal blending of the four models can provide enhanced forecasts at any forecasting horizon and station compared with the most accurate forecasts derived from a single model, taking advantage of their synergies and complementarities (Vislocky and Fritsch, 1995). This question is not new in the framework of weather forecasting research. Already, the blending of different forecasting sources has been shown to provide enhanced precipitation (Xie and Arkin, 1996), temperature (Salazar et al., 2011) and wind speed (Xiao et al., 2015) forecasts.

For renewable energies in general, model blending has been identified as a great possible avenue for improving forecasting accuracy for both wind velocity (Tascikaraoglu and Uzunoglu, 2014) and solar radiation (Tuohy et al., 2015). The integration methods proposed across the literature mostly make use of statistical integration of data via linear regression (Lorenz et al., 2012b) or other estimates. A weighting of models can drastically improve prediction as was proposed in Kühnert (2016), whose results show that linear regression applied to numerical weather prediction (NWP) and cloud motion vector (CMV) models can improve the overall accuracy of forecasts on horizons from 0 to 5 hours. The system described in Haupt et al. (2018) blends several short-term forecasting models by weighting the model contributions according to their historical performance at each lead time.

Machine learning is a very popular approach to forecasting and time series in general, and it has seen some use in solar forecasting with ever growing popularity (Mellit, 2008; Voyant et al., 2017; Zamo et al., 2014a,b). Previous work shows the potential of using machine learning to combine NWP forecasts. For example, artificial neural networks (ANNs) can be applied as proposed by Voyant et al. (2012), where a NWP and meteorological current measurements are merged by means of an ANN. In another approach (Lu et al., 2015), three different NWPs are used to forecast meteorological variables, from which predictor variables are obtained and they are blended with random forest to predict day-ahead radiation. The results again show that this combination is more accurate than the individual NWPs. Wolff et al. (2016) combines irradiance measurements, satellite and NWP to produce better forecasts of PV production using support vector machines (SVM) with notable results. The work of Aguiar et al. (2016) builds an ANN where ground measurements, European

Center for Medium-Range Weather Forecasts (ECMWF) and satellite data are blended. The experimental results show a clear increase of skill when multiple information sources are used. Machine learning and integration have been also used in operational environments (Hamann, 2017) with multiple configurations of input NWP models, learning algorithms and data size. Recently, Dersch et al. (2019) have proposed and evaluated an optimal combination of forecasting models specifically designed for obtaining improved DNI forecasts.

The models proposed across the literature have shown promising results for forecasting integration. The interest of the present work is to study the integration of different forecasting models in order to obtain improved short term (i.e. six hours ahead, and time resolution of 15 minutes) GHI and DNI radiation predictions. With this aim, we propose and evaluate different blending approaches that integrate the predictions of GHI (or DNI) provided by the four models analyzed in the companion paper (Part 1). The blending of those four models is of particular interest because they follow four almost independent approaches/foundations.

Two different blending approaches are studied and evaluated, namely: horizon and general approach. The difference between them is that the horizon approach constructs one model per forecasting horizon using training data from that horizon only, while the general approach constructs a single model trained with data from all forecasting horizons. The horizon approach trains models specialized on each horizon but there is less data to train each of the models, while the general one constructs a model independent of the horizon, being trained using data from all horizons. Both approaches make use of the SVM (Cortes and Vapnik, 1995). SVM models use as input the four models presented and assessed in Part 1, which will be referred hereinafter as the four predictors. SVM has been trained using linear and non-linear (radial) kernels with the aim to study whether non-linear blending is more adequate than linear.

Blending strategies may include additional parameters, beyond the models input. For instance, Lu et al. (2015) have proposed a novel methodology for model blending that takes into account additional weather state parameters for day-ahead solar radiation forecasting. These parameters accounted for different weather categories in which the different models provided enhanced performances. Results showed improved accuracy compared to reference model blend-

ing approaches. In McCandless et al. (2016b,a) a cloud-regime dependent short-term statistical model is proposed, which also showed enhanced performance. In the companion paper (Part 1) weather conditions were found to have an important influence on the performance of the four models evaluated. Then, a specific study, that attempts to include the weather conditions in the model blending procedure, is here conducted. To this end, the weather types identified as relevant for the solar radiation in the study area (Rodríguez-Benítez et al., 2018), and already assessed in the companion paper, are the starting point.

Finally, accurate solar forecast at site (station) level are relevant for plant owners. On the other hand, regional average (or aggregated) forecasts for entire regions are important for transmission system operators (TSOs) (Pierro et al., 2017). As solar energy increases its share electric systems, an enhanced accuracy of the solar radiation forecast at grid level is needed in order to manage the electric systems (Renné, 2014). In this work a specific blended regional model is presented and evaluated.

This work is organized as follows. 7.2 presents in detail the different blending approaches (horizon, general, weather type aware and regional forecast), including at the end of this section a brief introduction to the SVM methodology. In 7.3 the data and experimental procedure are introduced, while in 7.4 results are presented and discussed. Finally, some overall conclusions are provided in 7.5.

7.2 Blending approaches

This section describes the methodology followed for obtaining the blending models. Firstly, two different blending strategies (general and by horizons) are explained. Secondly, an approach that constructs blended models conditioned to the weather types described in Rodríguez-Benítez et al. (2018) will be presented. The previously mentioned blending approaches are suitable for individual stations. Therefore, in the third subsection, regional blending approaches are described. In principle, any machine learning technique could be used for obtaining the blending models. In this article, SVMs have been used because they have shown very good performance. Therefore, finally, a short description of the SVM for regression will be provided.

7.2.1 Horizon and General Approaches

The approach to predict both types of irradiance (GHI and DNI) is to blend the four predictors described in the companion paper (Part 1) by means of machine learning models. The blending aims to find a function (f) that combines the predictors (P_i) using Eq. 7.1:

$$I = f(P_1, P_2, P_3, P_4), \tag{7.1}$$

where f is constructed using SVMs. To test the linearity of this problem, both linear and non-linear SVM kernels have been used.

The blending approach aims to calculate at a given point in time (t) an accurate prediction of irradiance (I) for different forecasting horizons (h), from 15 minutes to 6 hours, with 15 minute steps. Two different approaches have been used to optimize the combination of the four predictors. The main difference between them is that the horizon approach constructs a model for each horizon, while the general approach trains a unique model valid for all horizons. Both are described below.

Horizon approach: This approach trains a different model f_h for each horizon, hence allowing for horizon-dependent predictor blending. Therefore, there is a model for horizon 15, another for horizon 30, and so on, up to horizon 360 (thus, there are 24 models in total). Fig 7.1 displays how each model f_h is trained using data belonging to horizon h only. Data for training model f_h is made of patterns like $((P_1(t,h), P_2(t,h), P_3(t,h), P_4(t,h)),\ I(t+h)$, for all t. The $P_i(t,h)$ denote predictors issued at time t, that make forecasts for time $t+h$. Once the models have been constructed, Eq. 7.2 shows how the f_h models can be used for irradiance forecasting. For instance, if at time t we would like to know the forecast in 15 minutes time, f_{15} would be used and $f_{15}\left(P_1(t,15), P_2(t,15), P_3(t,15), P_4(t,15)\right)$ would be computed.

$$I(t+h) = \begin{cases} f_{15}\left(P_1(t,15), P_2(t,15), P_3(t,15), P_4(t,15)\right), & \text{if} h = 15 \\ f_{30}\left(P_1(t,30), P_2(t,30), P_3(t,30), P_4(t,30)\right), & \text{if} h = 30 \\ \ldots \\ f_{360}\left(P_1(t,360), P_2(t,360), P_3(t,360), P_4(t,360)\right), & \text{if} h = 360 \end{cases} \tag{7.2}$$

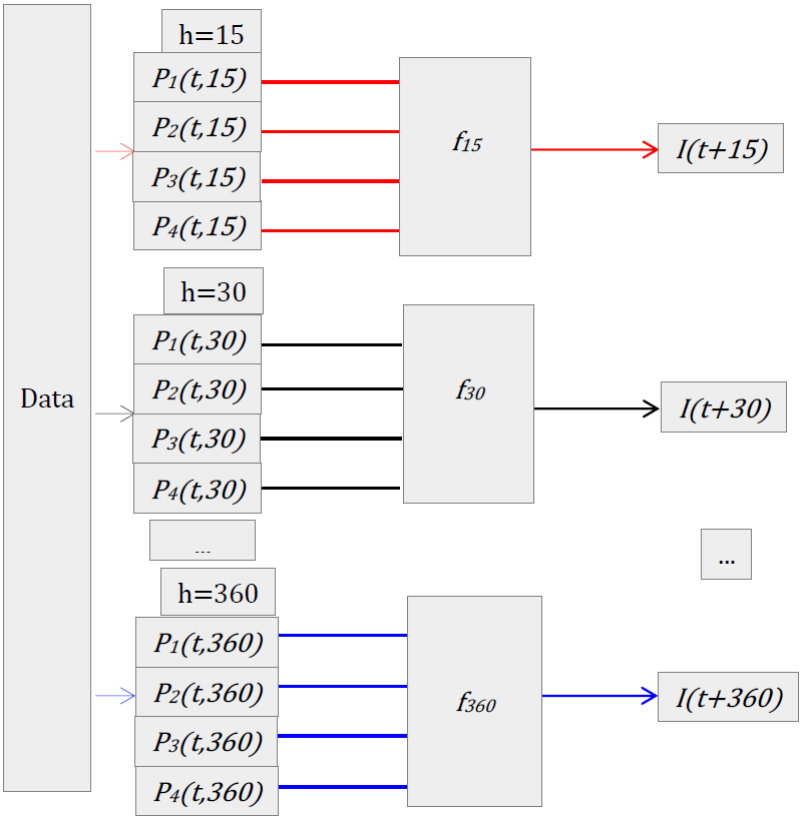

Figure 7.1: Horizon Blending Approach: there is one model f_h per forecast horizon. Each f_h is trained with data belonging to that horizon. In order to make predictions for $I(t+h)$, the appropriate model f_h and predictors $P_i(t,h)$ are used.

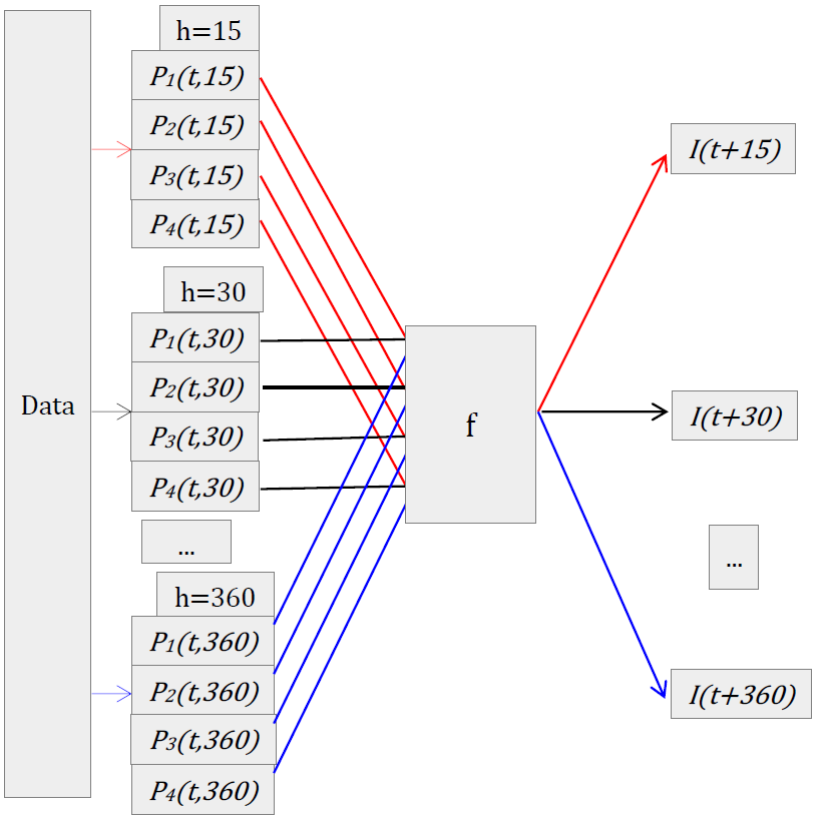

Figure 7.2: General Approach: there is a single model f valid for all horizons, which is trained with data belonging to all horizons. In order to make predictions for $I(t+h)$, the appropriate predictors $P_i(t,h)$ are used together with model f.

General approach: This approach constructs a single model f that blends the four predictors independently of the horizon. Fig. 7.2 shows how data from all horizons is combined and a single model f is trained. In this case, training data are the patterns $(P_1(t,h), P_2(t,h), P_3(t,h), P_4(t,h))$, $I(t+h)$, for all available t and h. Once the model f is trained, it is used for irradiance prediction using Eq. 7.3. It can be noticed that, although $P_i(t,h)$ is used as input for predictions at $t+h$, f itself is independent of h.

$$I(t+h) = f\left(P_1(t,h), P_2(t,h), P_3(t,h), P_4(t,h)\right) \tag{7.3}$$

7.2.2 Blending approach using weather types

In the companion paper (Part 1), a remarkable dependence of the four predictors on the weather conditions was found (section 6.5.4). Motivated by this result, a specific blending analysis depending on the weather conditions is undertaken. The same weather types (WT) described in section 6.5.4 of the companion paper have been used. There are basically four weather types: synoptic perturbation over the study area (WT1), transient weather conditions (WT2), moderate high pressure anomalies over the study area (WT3), and high pressure system over the whole study area (WT4). The aim is to construct blended models specialized in each of these weather types.

With this purpose, the general blending approach described in section 7.2.1 has been modified to consider WT's. For this analysis, four different general models have been trained grouping data by the four WT's considered. Each of those four models is constructed using the scheme shown in Fig. 7.2, that is joining data for all horizons but separating data by WT. Therefore, in this case, there will be four general blending models, one per WT. The horizon approach could also have been used, but it has not been considered because it would involve dividing the data into too many groups (one group for each WT and horizon), resulting in 4 WT's $\cdot$ 24 horizons = 96 models (versus the four models required by the general approach). This implies that there would be few data for training each of the models. That is the main reason why the general approach has been used instead.

7.2.3 Regional forecasting

Regional forecasting means that only the regional average (or aggregated) irradiance is of interest, instead of the local irradiance prediction addressed in previous sections. In principle, regional forecasting should be more accurate, as local fluctuations will be averaged. The aim is to predict the average of irradiance (GHI and DNI) over a region, which in this case, it will be the mean of the four individual stations (Jaen, Lisbon, Madrid, and Seville).

In this work, two different regional blending approaches are studied to predict irradiance at the different forecasting horizons considered. The first one performs regional forecasting by computing the

average of the four local blending models. That is, if Eq. 7.4 represents the blending model for station S (where S can be Jaen, Lisbon, Madrid, and Seville), then the regional model is given by Eq. 7.5, where r is the number of stations ($r = 4$ in this case).

$$I^S(t+h) = f^S\left(P_1^S(t,h), P_2^S(t,h), P_3^S(t,h), P_4^S(t,h)\right) \tag{7.4}$$

$$\overline{I}(t+h) = \frac{1}{r}\sum_{S=1}^{r} f^S\left(P_1^S(t,h), P_2^S(t,h), P_3^S(t,h), P_4^S(t,h)\right) \tag{7.5}$$

The second one constructs a model (Eq. 7.6) whose inputs are the four available predictors at each of the r stations. Therefore, r stations $\cdot$ 4 predictors will be used as inputs. The target output to train model f is now directly the average of irradiance (GHI or DNI) at the four locations.

$$\begin{aligned}\overline{I}(t+h) = f(&P_1^{S1}(t,h), ..., P_4^{S1}(t,h), P_1^{S2}(t,h), ..., P_4^{S2}(t,h), ...,\\ &P_1^{Sr}(t,h), ..., P_4^{Sr}(t,h))\end{aligned} \tag{7.6}$$

7.2.4 Support Vector Machines

SVM (Cortes and Vapnik, 1995) is a machine learning algorithm often used for classification and regression problems. For classification, the learning algorithm searches the optimal hyperplane dividing two different classes. This is achieved maximizing the margin between the instances of both classes. Non-linear models can be obtained by means of kernel functions (the so called kernel trick). These functions transform data into a higher dimensional space where the maximum margin hyperplane is computed.

In this article, SVM's have been used for regression. The SVM approach to regression has been described in Scholkopf and Smola (2001), and some details will be provided here. Assume a training data set $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$ where the outputs y_i are real values, and the x_i are the input variables. The equation $f(x) =$

$\langle w, x\rangle + b$ is the general form of the model, where $\langle \cdot, \cdot \rangle$ is the dot product, b a real number and w is the weights vector, to be optimized. The objective is to look for a model f as simple as possible, but achieving predictions for all instances with errors no larger than some allowed deviation ε. This can be done by minimizing the modulus of w by means of the optimization problem represented in Eq. 7.7.

$$\text{minimize} \frac{1}{2}\|w\|^2$$

$$\text{subject} \begin{cases} y_i - \langle w, x_i\rangle - b \leq \varepsilon \\ \langle w, x_i\rangle + b - y_i \leq \varepsilon \end{cases} \tag{7.7}$$

However, Eq. 7.7 assumes the existence of a valid solution that approximates all (x_i, y_i) pairs within the allowed deviation ε. This may not always be the case (for instance, if some of the instances are very noisy) and some minor errors should be allowed in order to find a feasible solution. This is achieved by introducing slack variables ξ_i and ξ_i^* (errors beyond the allowed deviation ε). The final optimization problem is presented in Eq. 7.8.

$$\text{minimize} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

$$\text{subject} \begin{cases} y_i - \langle w, x_i\rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i\rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}, \tag{7.8}$$

where $C > 0$ is a constant that determines the degree of importance of the two goals in the optimization process (the complexity of the model measured by $\|w\|^2$ and the size of the slack variables ξ_i and ξ_i^*. If C is too large, then the slack variables will be small, but this may lead to overfitting to noise. If C is too small, the optimization process will focus on minimizing the complexity of the model, and thus underfitting may be the result. C is therefore the main hyper-parameter that must be tuned in order to obtain well performing SVM's. This optimization problem can be further extended to non-linear models by means of the kernel trick, already mentioned for classification. One of the most widely used kernels is the gaussian

(or radial) one. In that case, the SVM model becomes that of Eq. 7.9.

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*)\, e^{-\frac{\|x - x_i\|^2}{2\sigma^2}} \tag{7.9}$$

Those x_i in Eq. 7.9 for which $(\alpha_i - \alpha_i^*) \neq 0$ are the support vectors, that is, the training instances selected by the optimization method for constructing model *f*. $(\alpha_i - \alpha_i^*)$ are factors computed during optimization that determine the importance of each of the support vectors, and σ is the standard deviation of the gaussian function, around each of the support vectors.

7.3 Data and experimental methodology

The four models described in the companion paper (Part 1), namely Satellite, WRF-Solar, Smart Persistence and CIADCast, are here used as inputs for the blending approaches. Also, the same forecasts dataset (Table 6.1 of Part 1) is here used for both training and evaluating the proposed blended models. Note that there is significantly less data as the horizon increases, due to the time range and the increasing time window from t to $t+h$. Forecasts are made at different time horizons, from 15 minutes to 6 hours, with 15 minute steps. Models are evaluated independently for the GHI and the DNI at the four evaluation stations: Seville, Jaen, Madrid and Lisbon.

Every approach described in 7.2 requires models to be trained and validated. In this work, cross-validation (CV) has been applied for this purpose. CV is a common practice in machine learning validation. CV creates N random partitions (or folds) of the same size. For each partition i, a model is trained with all partitions but i, and tested on i. The CV result is the average of the i testing partitions.

Standard CV partitions the dataset randomly. This is appropriate when instances are independent but in the case of time series, there is a temporal dependency between instances. However, this kind of CV risks using consecutive days for both training and testing, which may result in overly optimistic estimations of performance. In order to mitigate this problem, a variation of CV has been used here

(called grouped CV). There will be 4 different partitions (CV with $N = 4$), one for each week of every month. Therefore, partition 1 contains the first week of January, the first of February, and so on. Similarly for partition 2, with the second week of every month. This guarantees that, at least, training and testing partitions will never contain instances belonging to the same week.

Blending approaches can use both linear and non-linear models. For this purpose, experimental results have been obtained using linear and non-linear (radial) SVMs. SVMs require tuning hyper-parameter C (see Eq. 7.8). This has been done by systematically evaluating different C values ($C = \{0.25, 0.5, 1, 2, 4\}$) by means of standard CV on the training partition. It must be noted that the test partition is not used for hyper-parameter tuning.

To measure the quality of the blending models, the metrics relative root mean square error (rRMSE) and relative mean absolute error (rMAE), presented in the companion paper (Part 1) have also been used.

7.4 Results and discussion

7.4.1 Overall performance of the Horizon and General approaches

In this section, we show the empirical results of the Horizon and General blending approaches described in section 7.2.1 for the four stations: Jaen, Lisbon, Madrid, and Seville. Both approaches have been estimated using SVM with linear and radial kernels, resulting in four blending models, referred to as: SVMLinear-Horizon, SVMRadial-Horizon, SVMLinear-General, and SVMRadial-General. They are compared to the performance of the four original predictors. Both GHI and DNI are considered. The average errors (rRMSE and rMAE) for all horizons and for each station (Jaen, Lisbon, Madrid, Seville) are shown in Tables 7.1 and 7.2 for GHI and DNI, respectively.

As is observed in Table 7.1, the average error for GHI of the four predictors is consistently higher than any of the blending approaches. The improvement attained using the blending models, compared to

the best performing predictor model, is maximum at Lisbon. At this station the rRMSE of the best predictor is 49.97% and SVMLinear-Horizon model value is 41.38%, i.e., about 9% absolute improvement, or 17% or relative improvement. Jaen and Seville stations show a similar relative improvement (17% and 16%, respectively). Comparing blending techniques, it is observed that the rRMSE for different approaches is very similar, although the SVMLinear-Horizon performs best in Jaen (28.83%), Lisbon (41.38%) and Madrid (32.20%). In Seville the lowest rRMSE of 27.23% is reached using the General approach and Linear SVM, although the differences in terms of rRMSE are very small.

With respect to rMAE, blending models also outperform the four predictors. In this case all stations show around 3% of absolute improvement. The maximum relative improvement is achieved at Jaen (14%) and Seville (15%). The best blending approach is SVMRadial-General for all stations: Jaen (16.75%), Lisbon (29.02%), Madrid (20.12%) and Seville (15.92%). It is also observed that there is some variability across stations, errors in Lisbon tend to be higher than others while the error in Seville tends to be the lower than the rest.

Table 7.1: Summary for average GHI errors of the four predictors (top) and the four blending approaches (bottom). The average is computed over all forecasting horizons. Values are in %.

		rRMSE				rMAE			
GHI		Jaen	Lisbon	Madrid	Seville	Jaen	Lisbon	Madrid	Seville
Satellite		35.85	49.97	39.53	33.64	22.63	36.86	27.18	20.49
WRF-Solar		35.35	51.11	36.30	32.37	21.93	33.95	23.78	18.75
SmartPersistence		34.87	50.10	38.70	34.23	19.50	32.88	23.23	18.86
CIADCast		35.36	50.76	41.61	33.82	22.78	36.86	29.68	21.23
SVMRadial	General	29.19	41.94	32.89	27.67	**16.75**	**29.02**	**20.12**	**15.92**
SVMLinear	General	28.98	41.67	32.33	**27.23**	17.40	29.43	20.85	16.55
SVMRadial	Horizon	28.97	42.31	32.50	27.86	17.64	30.35	20.79	17.08
SVMLinear	Horizon	**28.83**	**41.38**	**32.20**	27.28	17.34	29.51	20.80	16.43

A similar performance is observed for DNI forecasting (see Table 7.2). All error metrics are lower for the blending methods than the four predictors, and comparing the machine learning methods also shows that they perform similarly. The main difference between GHI and DNI is that the errors are overall much worse than for GHI (for example, rRMSE from Lisbon goes from 41.38% to 73.45% and rMAE goes from 29.02% to 50.51%). For rRMSE, the best approach appears to be the Horizon one. For Jaen (44.24%) and Lisbon (73.45%) the Radial SVM performs best, and for Madrid (55.48%) and Seville (41.81%) the Linear SVM has the lowest error. The improvement attained using the blending models, compared to the best performing input models, is maximum at Lisbon, about 14% of absolute improvement and 16% relative improvement.

With respect to rMAE and DNI, the average results show that the SVMRadial-General approach outperforms all other models, as was also observed for GHI (see Table 7.2). The most important improvement versus the performance of predictors, is observed at the Lisbon station (about 7% in absolute value and 12% in relative value). In this case the maximum relative improvement happens at Seville station (14%).

Table 7.2: As in Table 7.1 but for the DNI.

		rRMSE			rMAE			
DNI	Jaen	Lisbon	Madrid	Seville	Jaen	Lisbon	Madrid	Seville
	Satellite	53.36	88.99	66.69	51.35	36.25	67.96	46.53	34.74
	WRF-Solar	54.36	89.21	67.86	48.72	36.64	63.33	46.18	32.48
	SmartPersistence	52.49	87.41	62.60	51.56	29.74	57.14	37.35	29.14
	CIADCast	55.21	90.95	70.73	54.46	38.24	69.17	50.99	37.37
SVMRadial	General	45.15	73.99	56.19	42.54	**27.09**	**50.51**	**35.50**	**25.14**
SVMLinear	General	46.79	78.35	57.04	42.80	28.90	55.05	36.54	27.41
SVMRadial	Horizon	**44.42**	**73.45**	55.75	42.08	27.52	51.43	36.27	25.97
SVMLinear	Horizon	44.78	74.50	**55.48**	**41.81**	28.55	54.76	36.33	27.15

Additionally, in Table 7.3 the average of RMSE-Skill for all horizons and each location is shown. The RMSE-Skill is calculated for both GHI and DNI for the four blending approaches. It is observed that the blending approaches achieve good values of RMSE-Skill for GHI and DNI, being better for GHI. As it has been observed previously concerning RMSE, there are no important differences regarding the performance of the four different blending models, although Horizon models tend to perform better in general.

Table 7.3: Average RMSE-Skill for GHI and DNI of the four blending approaches. The average is computed over all forecasting horizons. Values are in %.

		GHI				DNI			
		Jaen	Lisbon	Madrid	Seville	Jaen	Lisbon	Madrid	Seville
SVMRadial	General	16.19%	16.21%	15.04%	19.14%	13.33%	14.46%	10.19%	17.09%
SVMLinear	General	16.89%	16.72%	16.51%	**20.48%**	10.76%	10.39%	8.87%	16.79%
SVMRadial	Horizon	16.86%	15.79%	16.09%	18.52%	**14.93%**	**15.26%**	**10.89%**	18.21%
SVMLinear	Horizon	**17.17%**	**17.18%**	**16.74%**	20.33%	13.81%	13.49%	10.88%	**18.38%**

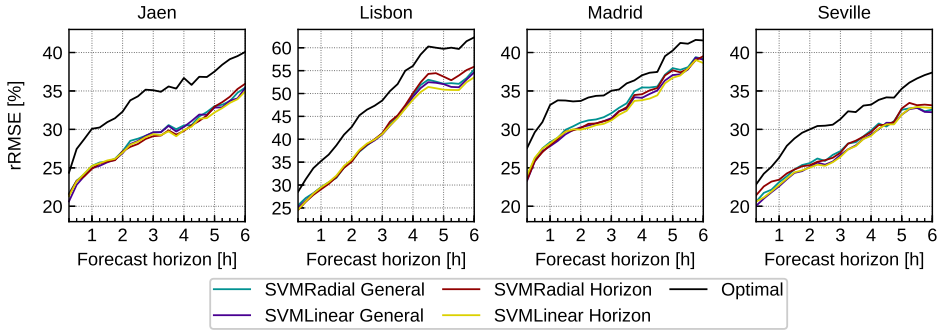

Figure 7.3: GHI rRMSE of blending models (General and Horizon) by horizon. "Optimal" displays the best performance out of the four predictors (the optimal line). Note the difference range of values for the Lisbon station.

7.4.2 Horizon and General approaches performance depending on the forecasting horizon

Here, the performance of the blending approaches and the four predictors is broken down by forecasting horizon for GHI and DNI and for the four stations. Given that machine learning methods outperform the four predictors, in the following figures the line named "optimal" represents the best performance for each horizon among the four models. Note that this reference is highly stringent, since the best performing model at each forecasting horizon is unknown beforehand.

Fig. 7.3 displays the rRMSE for GHI along the horizons for the four stations. It is observed that the blending approaches outperform the optimal line for all forecasting horizons. Blending machine learning models reach relative improvements with respect to the optimal line of 6% in the worst case and 18% in the best one. On the other hand, the errors of the blending machine learning models overlap with each other, it being difficult to decide the best one approach for every horizon. However, SVMLinear-Horizon approach seems to perform better than the other approaches at least in some horizons. This is true for all stations, except for Seville, where this approach performs similar to SVMGeneral-Linear model. In any case, the difference in terms of rRMSE is very small. The rRMSE in every station increases with the horizon, without large differences between blending models.

In Fig. 7.4 the evolution of the rMAE for GHI is shown. In all

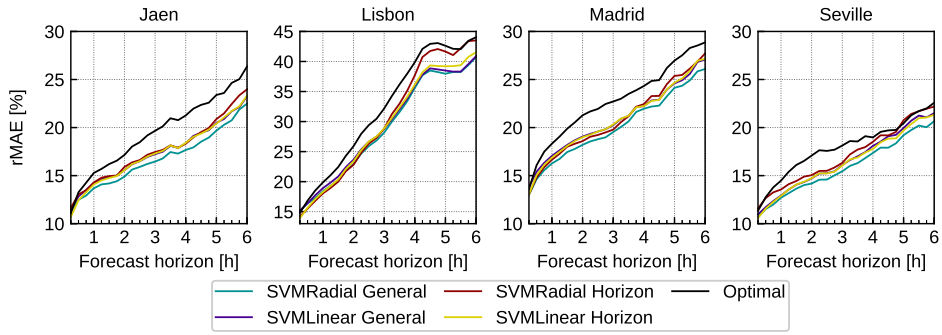

Figure 7.4: As in Fig. 7.3 but for the GHI rMAE forecasting errors.

stations, the optimal line is also worse than any other model, except at some horizons in Lisbon (from 330 to 360) and in Seville (from 285 to 345), where the optimal line is similar to the SVMRadial-Horizon approach. For rMAE, the relative improvement of blending models with respect to the optimal line goes from 3% (worst case) to 18% (best case, Jaen around 4 hours lead time). With respect to blending approaches, the best model for all stations is the SVMRadial-General approach. There is some overlap for some horizons with the SVMLinear-Horizon approach and with the SVMLinear-General approach in Lisbon station, but it still shows the best results. The relative improvement of the SVMRadial-General model with respect to the SVMLinear-Horizon approach is around 3%, 4% and 5% in many cases. It is also observed that the SVMRadial-Horizon approach is very poor for rMAE metric. It ends up being outperformed by all other machine learning models at far horizons in Jaen (from h=285 to h=360) and Lisbon (from h=225 to h=360) and at most horizons for Seville station. For short horizons, all models start at the same range and the rMAE grows as the horizon increases, but the growth amount depends on the station.

Unlike the rRMSE, observing the rMAE values for GHI, there is a clear best model for all stations, the SVMRadial-General approach. It is better for Jaen, Madrid and Seville than the rest of the machine learning approaches, and it is similar to the SVMLinear-General and SVMLinear-Horizon approaches for the Lisbon station.

Fig. 7.5 shows the rRMSE for DNI forecasting along the horizons. Similarly to GHI forecasting, the blending approaches outperform the optimal line, with relative improvements between 5% and 16%. There is also overlap between blending machine learning models, but as

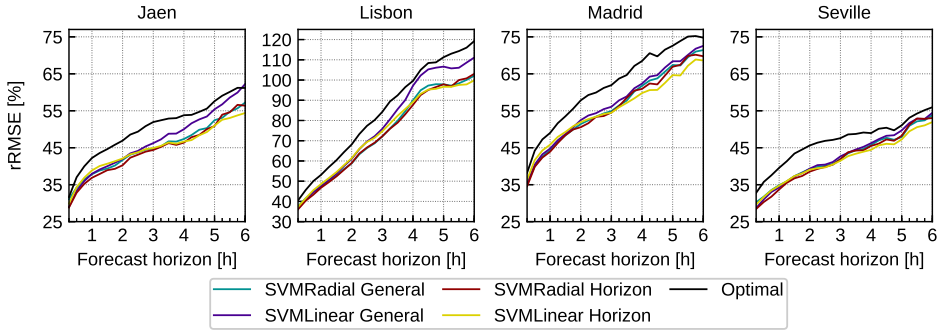

Figure 7.5: As in Fig. 7.3 but for DNI rRMSE forecasting errors.

the horizon increases, the SVMLinear-Horizon approach outperforms other models, where relative improvements of 2%, 3% and 4% are reached with respect to other approaches. At short horizons, the SVMRadial-Horizon model is slightly better, but the improvements are smaller. This pattern repeats for all stations, but it is more noticeable in Jaen, Madrid and Seville. Here the SVMLinear-General model is worse than the other models by a fair margin mainly for long horizons.

Fig. 7.6 shows the rMAE evolution for DNI forecasts. Unlike every other result presented before, here the optimal line outperforms or is similar to the blending models at short horizons (up to 60 minutes). As the horizon increases, the differences become greater and the SVMRadial-General approach appears to offer the best performance for long-term predictions, between 3% and 16% relative improvement over the optimal line. It is also observed that this approach overlaps with the SVMRadial-Horizon model in some stations and for short horizons. However, as the horizon increases, the advantage of the SVMRadial-General blending model becomes clearer, obtaining relative improvements up to 8%.

In summary, for GHI the four blending models perform similarly with respect to rRMSE, although for some horizons SVMLinear-Horizon is slightly better. For DNI and rRMSE, the SVMRadial-Horizon model seems the best model for short horizons and the SVMLinear-Horizon one for far horizons. For rMAE the SVMRadial-General model is the best for both GHI and DNI, although for DNI this shows more clearly at far horizons. Overall, the blending approaches outperform the optimal line for most of the forecasting horizons, at all stations and for both metrics, rRMSE and rMAE,

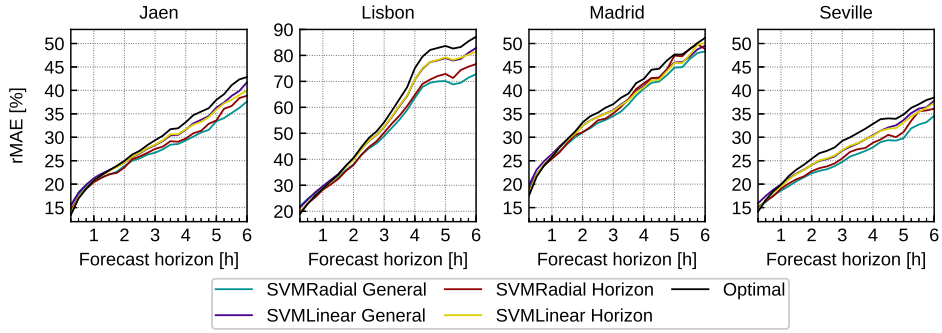

Figure 7.6: As in Fig. 7.5 but for DNI rMAE forecasting errors.

reaching relative improvements from 3% to 18%.

It can also be seen that the differences between the optimal DNI line and blending models increases with the forecasting horizon. Therefore, it can be concluded that the main added value of the blended model is attained at the end of the forecasting horizon and that the improvement statistics summarized in Table 7.2 can mainly be attributed to the longest forecasting horizons. For the case of the GHI, the differences among forecasting horizons are not so relevant, and the blending models provide a similar improvement along the whole forecasting period.

Comparisons of the improvements of the forecasting error attained here by model blending with those obtained in similar work are difficult due to, particularly, the differences in the dataset (location climatology, input models, and length of the dataset). A reference work is Wolff et al. (2016), that proposed and evaluated an optimal linear combination of three GHI short term forecasting models (persistence, satellite, and NWP). Evaluation was conducted in Germany for the period March to November. Model combination provided the lowest errors for all lead times, with improvements of a few percent of the installed solar power capacity. More recently, Dersch et al. (2019) evaluated an optimal combination of five DNI short term forecasting models (two satellite derived, two NWP derived and persistence). Evaluation was conducted in two stations located in south-eastern Spain and northern Africa. Results showed that the blended model outperformed all individual models for lead times below 300 minutes, in terms of rRMSE, when NWP models were the best performing models. Maximum relative improvements were found to be around 20% for one hour lead time, but differences rapidly decreased with

lead time. Results here presented are qualitatively similar, although climatology is more stringent here, regarding solar radiation forecasting, for the analyzed stations.

7.4.3 Analysis of models performance depending on the synoptic conditions

The role of the weather conditions on model blending is here assessed, based on the four WTs described in Rodríguez-Benítez et al. (2018). As has been mentioned in section 7.2.2, only the general approach has been used to study the influence of WTs in the blending model. Linear and non-linear approximations have been validated, but results are only displayed for the linear one (referred to as SVMLinear-WT) because it showed a better performance.

Fig. 7.7 breaks down SVMLinear-WT performance (rRMSE) by forecasting horizon, for both GHI and DNI. Performance is compared with the other blending approaches (General and Horizon) for the four stations. Results show that the inclusion of the weather types in the blending provides a relatively modest, but significant, improvement of the forecasts accuracy compared to the other blending approaches for some stations at certain forecasting horizons. Particularly, the SVMLinear-WT is shown to be the most accurate model for GHI at Seville for horizons larger than four hours, aproximately. The relative improvement is about 3% at the end of the forecasting period compared with SVMLinear-Horizon (overall the best approach for rRMSE). For the DNI and Jaen, the SVMLinear-WT shows better forecast for time horizons below 3 hours. But the best performance is observed for DNI and Seville, where a relative improvement between 1% and 3% is found for time horizons between 1 and 3 hours (compared to SVMRadial-Horizon, overall the best approach for rRMSE for short horizons). On the other hand, weather type information does not improve the forecast for Lisbon and Madrid.

The use of WT in model blending regarding short term solar radiation forecasting has not received much attention in the literature. Lu et al. (2015) reported an absolute improvement of the forecasting errors of about 30% (in terms of RMSE) when using a model blending approach based of WTs. But comparison is difficult due to, mainly, the different model inputs (they only used different NWP models as

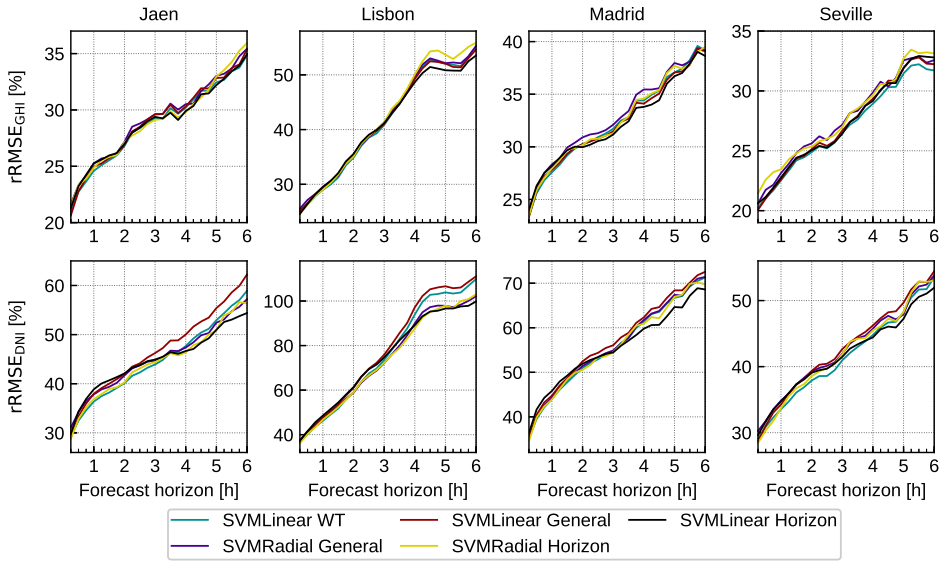

Figure 7.7: Performance of SVMLinear-WT model for rRMSE (GHI and DNI) compared with Global and Horizon approaches in all stations.

input of the blending approach) and the forecasting horizon (they assessed day ahead forecast).

7.4.4 Regional model approach

Next, experimental results for the two regional approaches described in section 7.2.3 are presented. In this case, SVM with radial kernel and a general approach (a single model for all horizons) have been used to construct the two regional blending models. They are referred to as SVMRadial-General-Mean (the first one: mean of local SVMRadial-General models) and SVMRadial-General-Regional (the second one: a model constructed with all available predictors of all local stations). Other approaches (General Linear, Horizon Linear or Horizon with Radial Kernel) might have been used, but only the general non-linear one has been selected here for empirical validation because it has shown good performance overall.

The results for both approaches are shown in Fig. 7.8, where the metrics (rRMSE and rMAE) have been calculated at each horizon and for both GHI and DNI. For comparison purposes, the figure also includes, for each predictor, its mean over the four locations.

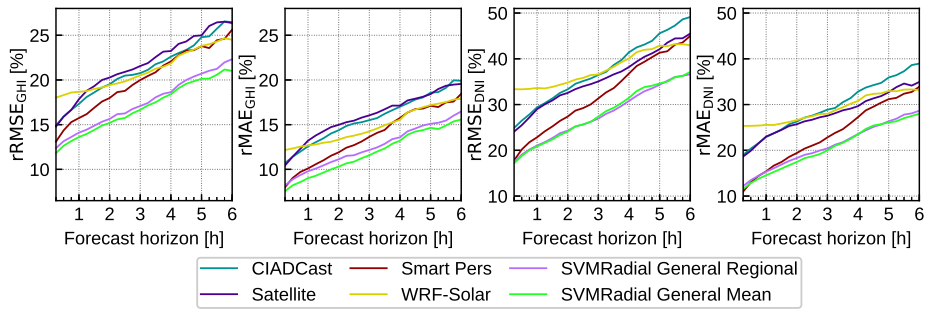

Figure 7.8: Performance of the two machine learning regional blending models (SVMRadial-General-Regional and SVMRadial-General-Mean). As reference, the performance of the four forecasting models (Satellite, WRF-Solar, Smart Persistence and CIADCast) and their average (mean) are also shown. Values are displayed both for the GHI and DNI and for rRMSE and rMAE metrics. The range of error values (in %) are different for GHI and DNI.

For instance, CIADCast in Fig. 7.8 is the mean of the CIADCast predictor over Jaen, Seville, Madrid, and Lisbon.

As may be expected, regional forecasts derived from each model (Satellite, WRF-Solar, Smart Persistence and CIADCast) show considerably smaller errors than forecasts of single stations (Figures 7.3 to 7.6), because of the spatial decorrelation of the forecasts errors. The Smart Persistence model performance is particularly outstanding, since only WRF-Solar, for GHI and for lead times greater than 4 hours, is competitive. It is also observed that both regional blending approaches display similar performance and they outperform the predictors for all horizons and for both GHI and DNI. Error reduction attained with these approaches increases with the forecasting horizon (Fig. 7.8), and is higher for rRMSE than for rMAE. For instance, an absolute improvement of about 8% is observed for DNI and rRMSE at the end of the forecasting period, when comparing the blending model against the best performing one (WRF-Solar).

As a summary of the regional forecasts results, Table 7.4 shows the average rRMSE and rMAE for all horizons and evaluated regional models. It can be seen that, indeed, the two regional approaches are very close, but SVMRadial-General-Mean displays smaller errors (except for rRMSE DNI). The relative improvement in rRMSE obtained by model blending is about 15% for GHI and 13% for DNI, compared

Table 7.4: Summary for average GHI and DNI errors of the four predictors means (top) and the two regional blending approaches (bottom). The average is computed over all forecasting horizons. Values are in %.

	GHI		DNI	
	rRMSE	rMAE	rRMSE	rMAE
Satellite	21.76	16.06	35.64	27.87
WRF-Solar	21.10	14.83	37.78	29.09
SmartPersistence	20.03	13.78	32.29	23.44
CIADCast	21.43	15.89	38.03	30.00
SVMRadial General Mean	**16.94**	**11.87**	28.01	**20.57**
SVMRadial General Regional	17.48	12.51	**27.84**	21.05

to the best performing input model. In rMAE, the improvements are 14% (GHI) and 12% (DNI). Comparing Tables 7.1, 7.2 and 7.3, it is observed, both for GHI and DNI, that the rRMSE values of the regional forecast are about one half of those of the individual stations. Previous work (Wolff et al., 2016) has reported that GHI forecasting errors of regional forecasts are reduced to about one third of the single stations forecasts. Differences may be explained based on the number of stations involved in the analysis, here only 4. This limited number of stations, only accounts for part of the spatial variability of the errors, and does not allow a full evaluation of the compensation effects of the forecasting error associated with the spatial decorrelation of these errors. Particularly, a further reduction of the error can be anticipated if the number of stations analyzed in the study area would be greater. Regarding DNI, no previous work has assessed regional DNI forecasts.

7.5 Summary and conclusions

In this article, the integration of GHI and DNI forecasting models has been addressed, with the aim of improving the prediction accuracies of those models used individually. This is achieved by using the forecasting models as predictors and blending them by means of machine learning techniques. In particular, in this work, the four forecasting models analyzed in the companion article (Satellite-based model, WRF-Solar, Smart Persistence and CIADCast) are blended

using SVM's. Two approaches have been evaluated: Horizon and General. The first one constructs a model for each horizon and the second one trains a single model valid for all horizons.

The blending approaches have been validated at four locations (Jaen, Lisbon, Madrid, and Seville) for forecasting horizons from 15 minutes to 6 hours with steps of 15 minutes. The two commonly used SVM kernels (linear and radial) have been tested. Results show that all blending approaches are able to reduce significantly the error of the original four predictors. This occurs for both GHI and DNI, for all horizons, and locations, and the two evaluation metrics. In general, it is observed that the improvement provided by the blending approaches does not strongly depend on the forecasting horizon, except for DNI. In that case, the improvement tends to increase for long horizons.

Differences between the blending approaches themselves are small, although the best approach depends on the kind of error to be used: Horizon models work slightly better for reducing rRMSE, while the General non-linear (radial) model works better for rMAE (being that difference larger than for rRMSE). This is true for both GHI and DNI.

With regard to the blending approach that uses weather-type information, it has been shown that there can be some improvement depending on the location. In this study, knowledge about weather conditions is useful for GHI at Seville for horizons larger than four hours. For DNI the gain is obtained for Seville and Jaen for horizons below three hours. Improvements occur only for rRMSE. This is probably due to weather-type help predicting outliers better, which is more important for the rRMSE metric.

Finally, predictor blending has also been applied in a regional context, where the aim is to predict the mean radiation of several locations in the region of interest. In this work, the region is represented by the four locations studied. Two machine learning approaches have been analyzed: the mean of the local blending models, and a model that uses as inputs all available predictors at all locations. Regional forecasting using the individual predictors (by computing their mean over the four locations) has been used for comparison. Again, the two regional blending methods outperform the regional predictors. This conclusion is general for both GHI and DNI, all horizons, and both metrics. The differences between the two approaches are small, although the first one obtains the smallest errors.

Acknowledgements

The authors are supported by the Spanish Ministry of Economy and Competitiveness, projects ENE2014-56126-C2-1-R and ENE2014-56126-C2-2-R (`http://prosol.uc3m.es`). The University of Jaén team are also supported by FEDER funds and by the Junta de Andalucia (Research group TEP-220). 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.

CHAPTER 8

Summary, conclusions, and future work

8.1 Summary

This study aims at contributing in the penetration of solar energy in the power grid. Independently of using storage system or balancing techniques in order to paliate the intermittency of solar production, solar radiation forecasting let the operators to anticipate contingency actions. In addition, with the aim of ease the tasks of solar plant managing systems operators and also motivated by the increasingly short-term required predictions, the forecasting horizon here conducted ranges from evaluation time to up to 3 or 6 h ahead, which is called nowcasting and short-term predictions, respectively. Two image-based technologies are mainly studied with this purpose: all-sky imager and MSG imagery. The detailed description of three state-of-the-art image-based prediction models and their assessment is showed in Chapter 3. Additionally, a measure-based method is also analyzed in order to assess the relative performance of the computed forecasts: (smart) persistence model. Finally, in order to promote all of solar power technologies, both GHI and DNI forecasts are computed and assessed.

Despite the fact that the forecasting models development usually aims to be location-independent, their performance might vary depending on several features as latitude, geographic and/or topographic conditions. In addition, sometimes due to the own models approach, the performance of prediction algorithms also varies depending on cloudiness and, particularly, the cloud type. Therefore, the latter might be an interesting feature to be taken into account. There are several satellite-based service with the aim of evaluating the cloud type. Alternatively, ground estimates of sky conditions have been traditionally retrieved by humans. To this end, an unsupervised cloud type classification based on sky-imager and, optionally, ceilometer is proposed (Chapter 4).

While some cloudiness is generated locally, and therefore its spatial variability in these cases is high, other cloud regimes are promoted owing to the atmospheric circulation. An analysis about the solar resource variability and its associated weather patterns is addressed in Chapter 5. Some results here obtained are assessed in light of cloud types and/or weather patterns. In addition, feature such as the weather pattern might be an additional predictor of solar resource variability.

Once a detailed assessment of forecasting models under different conditions and in different locations is conducted in Chapter 6, the best performing short-term prediction algorithm, as a result of a machine learning approach merging procedure, is proposed in Chapter 7.

8.2 Conclusions

Due to the fact that the persistence approach uses the own measure at evaluation time and its results are adapted according to the solar diurnal cycle (smart persitence), this simple forecasting model is difficult to be outperformed during the first minutes. The performance of smart persistence model depends on a clear sky model. As a consequence, this algorithm might not be used as a reference model to be outperformed. To this end, just a performance algorithm may be used, and the results might promote other sophisticated models, especially in later forecasting horizons. Nevertheless, at early lead times, the performance of persistence model may be almost the same

of smart persistence algorithm. In addition, the smart persistence model is used in this study as a reference algorithm due to the fact that its development is easy affordable, its computer resources consumption is light and its time consumption is much lower than other prediction algorithms. Therefore, the more sophisticated models have to probe their advantages in comparison to this easy affordable approach.

However, along all of the proposed studies, the Smart Persistence approach is not beat by any other model neither at evaluation time nor during the firsts time horizons (minutes). This results may be expected, particularly at lead-time 0, but also under conditions of clear skies or stratiform clouds. In other words, the more cloudiness intermittency, the less performance of persistence approaches.

As a result of the forecasting model comparison conducted in Chapter 3, the Camera mean model approach shows to be suitable at early horizons. Moreover, the Camera mean model outperforms all the individual camera ones. Nevertheless, the spread of camera prediction models reveals that slighly different sky conditions are observed from each camera. This fact is anticipated if geometrical relations are taken into account. From this point of view, cloud types tipically of lower altitudes would produces the higher spreads. However, cumulus clouds case-studies show low spread. Therefore, other reason might cause the miscoordination of the set of cameras, maybe the cloud detection method. On the other hand, the satellite-based models are showed to be suitable for nowcasting purposes. However, the Satellite HRV model, which involves a heavier preocessing than standard Satellite algorithm, does not offer skill over the latter, overall. Once again, an implicit "aggregation" step, applied when comparing standard images with HRV ones, might cause the better performance of Satellite model.

The sky-camera-based prediction model involves several procedures in order to compute the nowcast. However, just a part of these methods could be used with different purposes, as showed in Chapter 4. There, an sky-camera and, optionally, a ceilometer are used to compute an unsupervised classification of sky captures into different cloud types through machine learning techniques. Thus, a feasible alternative to traditionally human-based ground observation is proposed. However, the accuracy of the method is strongly dependent on the use of the ceilometer. Thus, without the proper cloud measures,

the accuracy of the method decreases, as well as when the number of cloud types increases. The latter conclusion could has been expected due to, intuitively, even for a human observator, the more categories to classify in, the more difficult. The whole proposed method is customized to the camera used and it is not directly extrapolated to another device.

As mentioned, the cloudiness, as the main cause of solar resource variability, could be related with the atmospheric circulation. In Chapter 5, the solar resource variability is associated with different weather patterns in terms of SLP. As a result of the proposed classification method, 4 main modes of variability are selected over the Iberian Peninsula in the annual case-study: a low-pressure pattern, a transient pattern coming from the Atlantic Ocean, a pattern that let the local conditions arise and, finally, a high-pressure pattern. Thus, given one of this main modes, the solar resource variability might be anticipated. Under this approach, a reduced number of modes ease the classification of a newly SLP observation. The proposed classification method relies on a nearly-objective algorithm. Thus, traditionally, but subjective, clustering metrics have been avoided. Moreover, when a newly SLP observation was classified into an elevated number of categories, a reduction in the classificator accuracy might be expected, as in case of Chapter 4.

In Chapter 6, a thorough assessment of a set of four short-term prediction models is conducted in light of, between others, the weather pattern. In addition to the results obtained by the Satellite (using standard visible channels 1 and 2) and Smart Persistence model, those from the WRF-Solar NWP model and the hybrid approach called CIADCast are also evaluated. As a conclusion, once again, the Smart Persistence model is difficult to be outperfomed, particularly under clear-sky conditions. However, intermittent cloudiness conditions let the other models to show some skill over the reference method. The conclusions are station-dependant but, particularly, similarly as in Chapter 5, the Lisbon station shows remarkable and distinctive features. In this case, this station is showed to be notably challenging to the forecasting models, probably due to the sea-land breeze effect.

In Chapter 7, the forecasting models previously described in Chapter 6 have been used as predictors in order to compute blended models. To this end, machine learning approaches, particularly SVMs

techniques, are used. All blending approaches definitely outperform each original prediction model individually. The best performing merging approach depends on the evaluation score.

8.3 Future work

The sky-camera-based nowcasting model involves a lot of methods. However, the highest impact on the results might be produced by the cloud detection algorithm due to two main reasons: 1) it is allocated in the firsts steps of the whole procedure and its inaccuracies are driven to the remaining methods, and 2) it is strongly dependent on the image features. There are a lot of issues that might be improved in the cloud detection algorithm, e.g. the unmanaged cases of reddish/bluish cloudy pixels or the traditional approach based on binary results (cloudy/cloud-free pixel). However, the wide casuistry derived from the brighness and the effect of aerosol in the sky make the cloud detector method a challenging issue, even in case of camera-customized developments. A future study at this regard may involve a different approach that includes non-binary-based results, with the aim of paliating the inaccuracies of misclassifications.

On the other hand, despite the fact that the streamline approach within the satellite-based procedure is showed to be suitable, it is expensive in terms of computation costs. Moreover, the single-pixel prediction approach makes these costs to enhance when the prediction for several pixels is requested. Alternatively, the prediction for a wide area (i.e. several neighbour pixels) might be preferable. Some attempts have been already conducted at this regard, also with hight computation costs. In those attempts, a pixel-by-pixel advection of Heliosat-2 cloud index is computed. While the intersection of pixels might be performed somehow, the divergence of pixels is, at the moment, unavailable. As a consequence, cloudiness coming from different pixels are aggregated at the same pixel and then, advected. Several improvements might be proposed but the resulting model performance, its sophistication and its computation costs have to be compared to those resulting from assimilating this feature into a NWP model. However, some future studies might be conducted at this regard.

CHAPTER 9

Resumen, conclusiones y trabajo futuro

9.1 Resumen

El objetivo de este estudio es contribuir a la penetración de la energía solar en la red eléctrica. Con independencia de si se emplean sistemas de almacenamiento o técnicas de compensación para paliar la intermitencia de la producción solar, la predicción de la radiación solar permite a los operadores anticipar acciones de contingencia. Además, con el fin de facilitar las tareas de los operadores de sistemas de gestión de plantas solares y también motivado por las predicciones cada vez a más corto plazo que se demandan, los horizontes de predicción aquí practicados se mueven en el rango desde el tiempo de evaluación hasta 3 o 6 horas en adelante, lo que se denomina predicción a muy corto plazo y a corto plazo, respectivamente. Para ello, se estudian principalmente dos tecnologías basadas en imágenes: cámara de cielo e imágenes del MSG. La descripción detallada de tres modelos de predicción basados en imágenes en el estado del arte y su evaluación se muestra en el Capítulo 3. Adicionalmente, también se analiza un método basado en medidas para valorar el rendimiento relativo de las predicciones calculadas: modelo de persistencia (inteligente). Finalmente, con el objetivo de promover todas las tec-

nologías de generación solar, ambas componentes de la radiación solar GHI y DNI son calculadas y evaluadas.

A pesar del hecho de que el desarrollo de modelos de predicción pretende ser independiente de la localización, sus rendimientos podrían variar dependiendo de varias características como la latitud, geografía y/o topografía. Además, algunas veces debido al propio enfoque de los modelos, el rendimiento de los algoritmos de predicción también varía dependiendo de nubosidad y, en particular, el tipo de nube. Por lo tanto, este último podría ser una característica interesante a tener en cuenta. Existen varios servicios basados en satélite para evaluar el tipo de nube. De manera alternativa, tradicionalmente se han adquirido estimaciones de tierra de las condiciones del cielo por humanos. Con este fin, se propone una clasificación no supervisada del tipo de nube basada en cámara de cielo y, opcionalmente, ceilómetro (Capítulo 4).

Mientras que alguna nubosidad se genera localmente, y por lo tanto la variabilidad espacial en estos casos es alta, otros regímenes de nube son desarrollados debido a la circulación atmosférica. Un análisis sobre la variabilidad del recurso solar y sus patrones de tiempo asociados se lleva a cabo en el Capítulo 5. Algunos de los resultados aquí obtenidos se valoran en virtud del tipo de nube y/o patrones de tiempo. Además, un característico como el patrón de tiempo podría ser un indicador adicional de la variabilidad del recurso solar.

Una vez llevada a cabo una valoración en detalle de los modelos de predicción bajo diferentes condiciones y en diferentes localidades en el Capítulo 6, en el Capítulo 7 se propone el algoritmo de predicción a corto plazo de mejor rendimiento, como resultado de un procedimiento de combinación desde el enfoque del *machine learning*.

9.2 Conclusiones

Dado que el enfoque de persistencia emplea la propia medida en el tiempo de evaluación y sus resultados son adaptados de acuerdo al ciclo solar diurno (persistencia inteligente), este simple modelo de persistencia es difícil de superar durante los primeros minutos. El rendimiento del algoritmo de persistencia inteligente depende a su vez

de un modelo de cielo despejado. Como consecuencia, este algoritmo podría no ser propuesto como modelo de referencia a ser superado. Para este fin, se puede utilizar simplemente un algoritmo de persistencia, y los resultados podrían favorecer otros modelos sofisticados, especialmente en horizontes de predicción lejanos. Sin embargo, en horizontes cercanos, el rendimiento del modelo de persistencia puede ser casi el mismo que el del algoritmo de persistencia inteligente. Además, el modelo de persistencia inteligente se usa en este estudio como algoritmo de referencia debido al hecho de que su desarrollo es fácilmente asequible, su consumo de recursos computacionales es ligero y su consumo de tiempo es mucho menor que el de otros algoritmos de predicción. Por lo tanto, los modelos más sofisticados tienen que probar sus ventajas en comparación a este enfoque fácilmente asequible.

Sin embargo, a lo largo de todos los estudios propuestos, el enfoque de Persistencia Inteligente no es superado por ningún otro modelo ni en el tiempo de evaluación ni durante los primeros horizontes temporales (minutos). Este resultado puede ser anticipado, especialmente a tiempo 0, pero también bajo condiciones de cielo despejado o nubes estratiformes. En otras palabras, a mayor intermitencia de la nubosidad, menor rendimiento de los enfoques de persistencia.

Como resultado de la comparativa de modelos de predicción llevada a cabo en el Capítulo 3, el enfoque del modelo de Cámara promedio muestra ser adecuado en los horizontes tempranos. Además, este modelo mejora a todos los de cámara individuales. Sin embargo, la dispersión de los modelos de predicción de cámara revela que se observan condiciones de cielo ligeramente diferentes desde cada cámara. Este hecho es esperado si se tienen en cuenta las relaciones geométricas. Desde este punto de vista, los tipos de nube típicamente a menores alturas producirían mayores dispersiones. Sin embargo, el caso de estudio de las nubes cúmulo muestran menor dispersión. Por lo tanto, otra razón podría causar la descoordinación del conjunto de cámaras, quizás el método de detección de nube. Por otro lado, los modelos basados en satélite muestran ser adecuados para propósitos de predicción a muy corto plazo. Sin embargo, el modelo Satélite HRV, que involucra un procesamiento más pesado que el algoritmo de Satélite típico, no ofrece mejoría sobre este último, en general. Una vez más, un paso de "agregación" implícito, aplicado cuando se compara las imágenes comunes con las HRV, podría causar el mejor comportamiento del modelo de Satélite.

El modelo de predicción basado en cámara de cielo involucra varios procedimientos para calcular la predicción a muy corto plazo. Sin embargo, solo una parte de esos métodos pueden ser utilizados con diferentes propósitos, como se muestra en el Capítulo 4. En este, una cámara de cielo y, opcionalmente, un ceilómetro se usan para realizar una clasificación no supervisada de capturas del cielo en diferentes tipos de nube a través de técnicas de *machine learning*. Así, se propone una alternativa factible a las tradicionales observaciones de tierra basadas en humanos. Sin embargo, la precisión del método es fuertemente dependiente del uso del ceilómetro. Así, sin las apropiadas medidas de la nube, la precisión del método decrece, del mismo modo que cuando el número de tipos aumenta. La última conclusión era de esperar debido a que, intuitivamente, incluso para un observador humano, a más categorías para clasificar, mayor dificultad. Todo el método propuesto está hecho a medida para la cámara empleada y no es directamente extrapolable a otro dispositivo.

Como se ha mencionado, la nubosidad, como la mayor causa de variabilidad del recurso solar, puede ser relacionada con la circulación atmosférica. En el Capítulo 5, la variabilidad del recurso solar se asocia con diferentes patrones de tiempo en términos de SLP. Como resultado del método propuesto de clasificación, 4 modos principales de variabilidad son seleccionados sobre la Península Ibérica en el caso de estudio anual: un patrón de bajas presiones, un patrón transitorio proveniente del Océano Atlántico, un patrón que permite que las condiciones locales aparezcan y, finalmente, un patrón de altas presiones. Así, dado uno de estos modos principales, la variabilidad del recurso solar podría ser anticipada. Bajo este enfoque, un número reducido de modos facilita la clasificación de una reciente observación de SLP. El método de clasificación propuesto está basado en un algoritmo cercano a ser objetivo. Así, las tradicionales, aunque subjetivas, métricas de clasificación han sido evitadas. Además, cuando una reciente observación de SLP fuese clasificada en un elevado número de categorías, podría esperarse una reducción de la precisión del clasificador, como en el caso del Capítulo 4.

En el capítulo 6, se lleva a cabo una evaluación en profundidad de un conjunto de cuatro modelos de predicción a corto plazo en virtud de, entre otros, el patrón de tiempo. Además de los resultados obtenidos por el modelo de Satélite (usando los canales visibles estándar 1 y 2) y Persistencia Inteligente, también se evalúan aquellos del modelo NWP WRF-Solar y del enfoque híbrido llamado

CIADCast. Como conclusión, una vez más, el modelo de Persistencia Inteligente es difícil de ser superado, particularmente bajo condiciones de cielo despejado. Sin embargo, las condiciones de nubosidad intermitente permiten a otros modelos mostrar algo de mejoría sobre el método de referencia. Las conclusiones son dependientes de la estación pero, en particular, de forma similar al Capítulo 5, la estación de Lisboa muestra características remarcables y distintivas. En este caso, esta estación muestra ser un reto notable para los modelos de predicción, probablemente debido al efecto de la brisa mar-tierra.

En el Capítulo 7, los modelos de predicción previamente descritos en el Capítulo 6 han sido usados como indicadores para elaborar modelos combinados. Con esta finalidad, se utilizan enfoques de *machine learning*, en particular, técnicas de SVM. Todos los enfoques de combinación mejoran definitivamente cada modelo de predicción original individual. El enfoque de combinación que mejor se comporta depende de la métrica de evaluación.

9.3 Trabajo futuro

El modelo de predicción a muy corto plazo basado en cámara de cielo involucra a muchos métodos. Sin embargo, el mayor impacto en los resultados podría ser producido por el algoritmo de detección de nube debido a dos razones principales: 1) está situado en los primeros pasos del procedimiento general y sus imprecisiones son transportadas al resto de métodos posteriores, y 2) es muy dependiente de las características de la imagen. Existen muchos problemas que podrían ser mejorados en el algoritmo de detección de nube, e.g. los casos no controlados de píxeles nubosos rojizos/azulados o el tradicional enfoque basado en resultados binarios (pixel nuboso/sin nube). Sin embargo, la amplia casuística derivada del brillo y los efectos del aerosol en el cielo hacen del método de detección de nube un problema exigente, incluso en el caso de desarrollos a la medida de una cámara. Un estudio futuro en este aspecto puede involucrar un enfoque diferente que incluye resultados no binarios, con el propósito de paliar las imprecisiones de los errores de clasificación.

Por otro lado, a pesar del hecho de que el enfoque de líneas de corriente dentro del procedimiento basado en satélite muestra ser

adecuado, es caro en términos de costes computacionales. Además, el enfoque de predicción para un único pixel hace que estos costes aumenten cuando se requiere la predicción para varios píxeles. Alternativamente, la predicción para una área amplia (i.e. varios píxeles vecinos) podría ser preferible. Algunos intentos han sido ya llevados a cabo a este respecto, también con grandes costes computacionales. En ellos, se realiza una advección pixel a pixel del índice nuboso del Heliosat-2. Mientras que la intersección de píxeles podría ser realizada de algún modo, la divergencia de píxeles no está disponible, en este momento. Como consecuencia, la nubosidad proveniente de diferentes píxeles es agregada en el mismo píxel y luego, advectada. Varias mejoras podrían ser propuestas pero el comportamiento del modelo resultante, su sofisticación y sus costes computacionales tienen que ser comparados con aquellos resultantes de asimilar este característico en un modelo NWP. Sin embargo, algunos estudios futuros podrían ser llevados a cabo a este respecto.

Bibliography

(2013). *Automated cloud classification using a ground based infra-red camera and texture analysis techniques*, volume 8890.

Aguiar, L. M., Pereira, B., Lauret, P., Díaz, F., and David, M. (2016). Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting. *Renewable Energy*, 97:599 – 610.

Alonso-Montesinos, J. and Batlles, F. (2015). The use of a sky camera for solar radiation estimation based on digital image processing. *Energy*, 90:377 – 386.

Alonso-Montesinos, J., Polo, J., Ballestrín, J., Batlles, F., and Portillo, C. (2019). Impact of DNI forecasting on CSP tower plant power production. *Renewable Energy*, 138:368 – 377.

Antonanzas, J., Pozo-Vázquez, D., Fernandez-Jimenez, L., and de Pison, F. M. (2017). The value of day-ahead forecasting for photovoltaics in the Spanish electricity market. *Solar Energy*, 158:140 – 146.

Arbizu-Barrena, C., Pozo-Vázquez, D., Ruiz-Arias, J. A., and Tovar-Pescador, J. (2015). Macroscopic cloud properties in the WRF NWP model: An assessment using sky camera and ceilometer data. *Journal of Geophysical Research: Atmospheres*, 120(19):10,297–10,312.

Arbizu-Barrena, C., Ruiz-Arias, J. A., Rodríguez-Benítez, F. J., Pozo-Vázquez, D., and Tovar-Pescador, J. (2017). Short-term solar radiation forecasting by advecting and diffusing MSG cloud index. *Solar Energy*, 155:1092 – 1103.

Avolio, E., Federico, S., Miglietta, M., Feudo, T. L., Calidonna, C., and Sempreviva, A. (2017). Sensitivity analysis of WRF model PBL schemes in simulating boundary-layer variables in southern Italy: An experimental campaign. *Atmospheric Research*, 192:58 – 71.

Barnett, T. P., Ritchie, J., Foat, J., and Stokes, G. (1998). On the Space–Time Scales of the Surface Solar Radiation Field. *Journal of Climate*, 11(1):88–96.

Beyer, H., Costanzo, C., Heinemann, D., and Reise, C. (1994). Short range forecast of PV energy production using satellite image analysis. In *Proc. 12th European Photovoltaic Solar Energy Conference*, volume 11, page 15. Amsterdam.

Blanc, P., Remund, J., and Vallance, L. (2017). 6 - Short-term solar power forecasting based on satellite images. In Kariniotakis, G., editor, *Renewable Energy Forecasting*, Woodhead Publishing Series in Energy, pages 179 – 198. Woodhead Publishing.

Boers, R., de Haij, M. J., Wauben, W. M. F., Baltink, H. K., van Ulft, L. H., Savenije, M., and Long, C. N. (2010). Optimized fractional cloudiness determination from five ground-based remote sensing techniques. *Journal of Geophysical Research: Atmospheres*, 115(D24).

Boland, J. and Grantham, A. (2018). Nonparametric Conditional Heteroscedastic Hourly Probabilistic Forecasting of Solar Radiation. *J – Multidisciplinary Scientific Journal*, 1(1):174–191.

BP (2019). BP Statistical Review of World Energy. techreport 68$^{\text{th}}$ edition, BP.

Brancucci Martinez-Anido, C., Botor, B., Florita, A. R., Draxl, C., Lu, S., Hamann, H. F., and Hodge, B.-M. (2016). The value of day-ahead solar power forecasting improvement. *Solar Energy*, 129:192 – 203.

Brayshaw, D. J., Troccoli, A., Fordham, R., and Methven, J. (2011). The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: A case study over the uk. *Renewable Energy*, 36(8):2087 – 2096.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.

Brouwer, A. S., van den Broek, M., Seebregts, A., and Faaij, A. (2014). Impacts of large-scale Intermittent Renewable Energy Sources on electricity systems, and how these can be modeled. *Renewable and Sustainable Energy Reviews*, 33:443 – 466.

Brox, T. and Malik, J. (2011). Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):500–513.

Calbó, J. and Sabburg, J. (2008). Feature Extraction from Whole-Sky Ground-Based Images for Cloud-Type Recognition. *Journal of Atmospheric and Oceanic Technology*, 25(1):3–14.

Cano, D., Monget, J., Albuisson, M., Guillard, H., Regas, N., and Wald, L. (1986). A method for the determination of the global solar radiation from meteorological satellite data. *Solar Energy*, 37(1):31 – 39.

Caruana, R., Karampatziakis, N., and Yessenalina, A. (2008). An Empirical Evaluation of Supervised Learning in High Dimensions. In *Proceedings of the 25th International Conference on Machine Learning*, ICML '08, pages 96–103, New York, NY, USA. ACM.

Caruana, R. and Niculescu-Mizil, A. (2006). An Empirical Comparison of Supervised Learning Algorithms. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, pages 161–168, New York, NY, USA. ACM.

Castro-Díez, Y., Pozo-Vázquez, D., Rodrigo, F. S., and Esteban-Parra, M. J. (2002). NAO and winter temperature variability in southern Europe. *Geophysical Research Letters*, 29(8):1–1–1–4.

Cazorla, A., Olmo, F. J., and Alados-Arboledas, L. (2008). Development of a sky imager for cloud cover assessment. *J. Opt. Soc. Am. A*, 25(1):29–39.

Cheng, H.-Y. and Lin, C.-L. (2017). Cloud detection in all-sky images via multi-scale neighborhood features and multiple supervised learning techniques. *Atmospheric Measurement Techniques*, 10(1):199–208.

Cheng, H.-Y. and Yu, C.-C. (2015). Block-based cloud classification with statistical features and distribution of local texture features. *Atmospheric Measurement Techniques*, 8(3):1173–1182.

Chow, C. W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., and Washom, B. (2011). Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy*, 85(11):2881 – 2893.

Coimbra, C. F., Kleissl, J., and Marquez, R. (2013). Chapter 8 - Overview of Solar-Forecasting Methods and a Metric for Accuracy Evaluation. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 171 – 194. Academic Press, Boston.

Coker, P., Barlow, J., Cockerill, T., and Shipworth, D. (2013). Measuring significant variability characteristics: An assessment of three UK renewables. *Renewable Energy*, 53:111 – 120.

Conover, W. J. and Iman, R. L. (1979). On multiple-comparisons procedures. Technical report, Los Alamos Scientific Laboratory. LA-7677-MS.

Correia, J., Bastos, A., Brito, M., and Trigo, R. (2017). The influence of the main large-scale circulation patterns on wind power production in portugal. *Renewable Energy*, 102:214 – 223.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Costa-Surós, M., Calbó, J., González, J. A., and Long, C. N. (2014). Comparing the cloud vertical structure derived from several methods based on radiosonde profiles and ground-based remote sensing measurements. *Atmospheric Measurement Techniques*, 7(8):2757–2773.

Crammer, K. and Singer, Y. (2001). On the Algorithmic Implementation of Multiclass Kernel-based Vector Machines. *Journal of machine learning research*, 2(Dec):265–292.

Deng, A., Gaudet, B., Dudhia, J., and 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*.

Derrien, M. and Gléau, H. L. (2005). Msg/seviri cloud mask and type from safnwc. *International Journal of Remote Sensing*, 26(21):4707–4732.

Dersch, J., Schroedter-Homscheidt, M., Gairaa, K., Hanrieder, N., Landelius, T., Lindskog, M., Müller, S., Ramirez Santigosa, L., Sirch, T., and Wilbert, S. (2019). Impact of DNI nowcasting on annual revenues of CSP plants for a time of delivery based feed in tariff. *Meteorologische Zeitschrift*, pages –.

Descombes, G., Auligne, T. D., Lin, H.-C., Xu, D., Schwartz, C. S., and 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.

Diagne, M., David, M., Lauret, P., Boland, J., and Schmutz, N. (2013). Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews*, 27:65 – 76.

Ela, E., Milligan, M., and Kirby, B. (2011). Operating Reserves and Variable Generation.

Ela, E., Tuohy, A., Entriken, R., Lannoye, E., and Philbrick, R. (2017). Using Probabilistic Renewable Forecasts to Determine Reserve Requirements. In 7$^{\text{th}}$ *Solar Integration Workshop. International Workshop on Integration of Solar Power into Power Systems*. EPRI. Electric Power Research Institute.

Engeland, K., Borga, M., Creutin, J.-D., François, B., Ramos, M.-H., and Vidal, J.-P. (2017). Space-time variability of climate variables and intermittent renewable electricity production – a review. *Renewable and Sustainable Energy Reviews*, 79:600 – 617.

ESIOS (2018). [dataset] measured generation balance. measurement system (simel). system operator information system (esios). red eléctrica de españa.

European Commission, Secretariat-General (2016). Communication from the commission to the european parliament, the council, the european economic and social comittee, the comittee of the regions and the european investment bank. clean energy for all europeans. Technical Report COM(2016), European Commision. 52016DC0860.

García-Bustamante, E., González-Rouco, J. F., Navarro, J., Xoplaki, E., Luterbacher, J., Jiménez, P. A., Montávez, J. P., Hidalgo, A., and Lucio-Eceiza, E. E. (2013). Relationship between wind power production and north atlantic atmospheric circulation over the northeastern iberian peninsula. *Climate Dynamics*, 40(3):935–949.

Gauchet, C., Blanc, P., Espinar, B., Charbonnier, B., and Demengel, D. (2012). Surface solar irradiance estimation with low-cost fisheye camera. In *Workshop on "Remote Sensing Measurements for Renewable Energy"*, Risoe, Denmark.

General Court (2015). Communication from the Commission to the European Parliament and the Council. Achieving the 10% electricity interconnection target. Making Europe's electricity grid fit for 2020. Technical Report COM(2015), European Commision. 52015DC0082.

Gueymard, C. (2004). High performance model for clear-sky irradiance and illuminance. In *PROCEEDINGS OF THE SOLAR CONFERENCE*, pages 251–258. AMERICAN SOLAR ENERGY SOCIETY; AMERICAN INSTITUTE OF ARCHITECTS.

Gueymard, C. A. and Wilcox, S. M. (2011). Assessment of spatial and temporal variability in the US solar resource from radiometric measurements and predictions from models using ground-based or satellite data. *Solar Energy*, 85(5):1068 – 1084.

Guillot, E. M., Vonder Haar, T. H., Forsythe, J. M., and Fletcher, S. J. (2012). Evaluating Satellite-Based Cloud Persistence and Displacement Nowcasting Techniques over Complex Terrain. *Weather and Forecasting*, 27(2):502–514.

Gutiérrez, C., Gaertner, M. Á., Perpiñán, O., Gallardo, C., and Sánchez, E. (2017). A multi-step scheme for spatial analysis of solar and photovoltaic production variability and complementarity. *Solar Energy*, 158:100 – 116.

Haiden, T., Forbes, R., Ahlgrimm, M., and Bozzo, A. (2015). The skill of ECMWF cloudiness forecasts. *ECMWF Newsletter*, 143:14–19.

Hamann, H. F. (2017). A Multi-scale, Multi-Model, Machine-Learning Solar Forecasting Technology.

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., and Wind, G. (2014). Remote sensing of cloud top pressure/height from SEVIRI: analysis of ten current retrieval algorithms. *Atmospheric Measurement Techniques*, 7(9):2839–2867.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621.

Haupt, S. E. (2018). *Short-Range Forecasting for Energy*, pages 97–107. Springer International Publishing, Cham.

Haupt, S. E., Kosović, B., Jensen, T., Lazo, J. K., Lee, J. A., Jiménez, P. A., Cowie, J., Wiener, G., McCandless, T. C., Rogers, M., Miller, S., Sengupta, M., Xie, Y., Hinkelman, L., Kalb, P., and Heiser, J. (2018). Building the Sun4Cast System: Improvements in Solar Power Forecasting. *Bulletin of the American Meteorological Society*, 99(1):121–136.

Heinle, A., Macke, A., and Srivastav, A. (2010). Automatic cloud classification of whole sky images. *Atmospheric Measurement Techniques*, 3(3):557–567.

Hoff, T. E. and Perez, R. (2012). Modeling PV fleet output variability. *Solar Energy*, 86(8):2177 – 2189. Progress in Solar Energy 3.

Horn, B. K. and Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17(1):185 – 203.

Houze Jr, R. A. (2014). *Cloud Dynamics*, volume 104. Academic press.

Huertas-Tato, J., Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Aler-Mur, R., Galván-León, I., and Pozo-Vázquez, D. (2017a). Automatic Cloud-Type Classification Based On the Combined Use of a Sky Camera and a Ceilometer. *Journal of Geophysical Research: Atmospheres*, 122(20):11,045–11,061.

Huertas-Tato, J., Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Aler-Mur, R., Galvan-Leon, I., and Pozo-Vázquez, D. (2017b). Automatic cloud-type classification based on the combined use of a sky camera and a ceilometer. *Journal of Geophysical Research: Atmospheres*, 122(20):11,045–11,061. 2017JD027131.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, 113(D13).

Illingworth, A. J., Hogan, R. J., O'Connor, E., Bouniol, D., Brooks, M. E., Delanoé, J., Donovan, D. P., Eastment, J. D., Gaussiat, N., Goddard, J. W. F., Haeffelin, M., Baltink, H. K., Krasnov, O. A., Pelon, J., Piriou, J.-M., Protat, A., Russchenberg, H. W. J., Seifert, A., Tompkins, A. M., van Zadelhoff, G.-J., Vinit, F., Willén, U., Wilson, D. R., and Wrench, C. L. (2007). Cloudnet: Continuous Evaluation of Cloud Profiles in Seven Operational Models Using Ground-Based Observations. *Bulletin of the American Meteorological Society*, 88(6):883–898.

Ineichen, P. (2008). A broadband simplified version of the Solis clear sky model. *Solar Energy*, 82(8):758 – 762.

Ineichen, P. (2014). Long term satellite global, beam and diffuse irradiance validation. *Energy Procedia*, 48:1586 – 1596. Proceedings of the 2nd International Conference on Solar Heating and Cooling for Buildings and Industry (SHC 2013).

Inman, R. H., Pedro, H. T., and Coimbra, C. F. (2013). Solar forecasting methods for renewable energy integration. *Progress in Energy and Combustion Science*, 39(6):535 – 576.

International Energy Agency (2018). *Market Report Series: Renewables 2018. Analysis and Forecasts to 2023.*

Jacobson, M. Z. and Delucchi, M. A. (2011). Providing all global energy with wind, water, and solar power, Part I: Technologies, energy resources, quantities and areas of infrastructure, and materials. *Energy Policy*, 39(3):1154 – 1169.

Jayadevan, V. T., Rodriguez, J. J., and Cronin, A. D. (2015). A New Contrast-Enhancing Feature for Cloud Detection in Ground-Based Sky Images. *Journal of Atmospheric and Oceanic Technology*, 32(2):209–219.

Jimenez, P. A., Hacker, J. P., Dudhia, J., Haupt, S. E., Ruiz-Arias, J. A., Gueymard, C. A., Thompson, G., Eidhammer, T., and Deng, A. (2016). WRF-Solar: Description and Clear-Sky Assessment of an Augmented NWP Model for Solar Power Prediction. *Bulletin of the American Meteorological Society*, 97(7):1249–1264.

Kassianov, E., Long, C. N., and Christy, J. (2005). Cloud-Base-Height Estimation from Paired Ground-Based Hemispherical Observations. *Journal of Applied Meteorology*, 44(8):1221–1233.

Kasten, F. and Czeplak, G. (1980). Solar and terrestrial radiation dependent on the amount and type of cloud. *Solar Energy*, 24(2):177 – 189.

Kazantzidis, A., Tzoumanikas, P., Bais, A., Fotopoulos, S., and Economou, G. (2012). Cloud detection and classification with the use of whole-sky ground-based images. *Atmospheric Research*, 113:80 – 88.

Kazantzidis, A., Tzoumanikas, P., Blanc, P., Massip, P., Wilbert, S., and Ramirez-Santigosa, L. (2017). 5 - Short-term forecasting based on all-sky cameras. In Kariniotakis, G., editor, *Renewable Energy Forecasting*, Woodhead Publishing Series in Energy, pages 153 – 178. Woodhead Publishing.

Kleissl, J. (2013). *Solar energy forecasting and resource assessment*. Academic Press.

Kliangsuwan, T. and Heednacram, A. (2015). Feature extraction techniques for ground-based cloud type classification. *Expert Systems with Applications*, 42(21):8294 – 8303.

Kniffka, A., Stengel, M., and Hollmann, R. (2014). [dataset] SEVIRI cloud mask dataset - Edition 1 - 15 minutes resolution, Satellite Application Facility on Climate Monitoring.

Köhler, C., Steiner, A., Saint-Drenan, Y.-M., Ernst, D., Bergmann-Dick, A., Zirkelbach, M., Bouallègue, Z. B., Metzinger, I., and Ritter, B. (2017). Critical weather situations for renewable energies – part b: Low stratus risk for solar power. *Renewable Energy*, 101:794 – 803.

Kokhanovsky, A. A. (2006). *Cloud Optics (Atmospheric and Oceanographic Sciences Library Book 34)*. Springer.

Kreuter, A., Zangerl, M., Schwarzmann, M., and Blumthaler, M. (2009). All-sky imaging: a simple, versatile system for atmospheric research. *Appl. Opt.*, 48(6):1091–1097.

Kriebel, K. T., Gesell, G., Kastner, M., and Mannstein, H. (2003). The cloud analysis tool APOLLO: Improvements and validations. *International Journal of Remote Sensing*, 24(12):2389–2408.

Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260):583–621.

Kuhn, M. et al. (2008). Building Predictive Models in R Using the caret Package. *Journal of statistical software*, 28(5):1–26.

Kuhn, P., Nouri, B., Wilbert, S., Prahl, C., Kozonek, N., Schmidt, T., Yasser, Z., Ramirez, L., Zarzalejo, L., Meyer, A., Vuilleumier, L., Heinemann, D., Blanc, P., and Pitz-Paal, R. (2018). Validation of an all-sky imager–based nowcasting system for industrial PV plants. *Progress in Photovoltaics: Research and Applications*, 26(8):608–621.

Kuhn, P., Wilbert, S., Prahl, C., Schüler, D., Haase, T., Hirsch, T., Wittmann, M., Ramirez, L., Zarzalejo, L., Meyer, A., Vuilleumier, L., Blanc, P., and Pitz-Paal, R. (2017). Shadow camera system for the generation of solar irradiance maps. *Solar Energy*, 157:157 – 170.

Kühnert, J. (2016). *Development of a photovoltaic power prediction system for forecast horizons of several hours*. PhD thesis, Carl von Ossietzky Universität Oldenburg.

Kühnert, J., Lorenz, E., and Heinemann, D. (2013). Chapter 11 - Satellite-Based Irradiance and Power Forecasting for the German Energy Market. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 267 – 297. Academic Press, Boston.

Lara-Fanego, V., Ruiz-Arias, J., Pozo-Vázquez, D., Santos-Alamillos, F., and Tovar-Pescador, J. (2012). Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain). *Solar Energy*, 86(8):2200 – 2217. Progress in Solar Energy 3.

Lauret, P., Perez, R., Aguiar, L. M., Tapachès, E., Diagne, H. M., and David, M. (2016). Characterization of the intraday variability regime of solar irradiation of climatically distinct locations. *Solar Energy*, 125:99 – 110.

Lave, M., Broderick, R. J., and Reno, M. J. (2017). Solar variability zones: Satellite-derived zones that represent high-frequency ground variability. *Solar Energy*, 151:119 – 128.

Lave, M., Kleissl, J., and Arias-Castro, E. (2012). High-frequency irradiance fluctuations and geographic smoothing. *Solar Energy*, 86(8):2190 – 2199. Progress in Solar Energy 3.

Law, E. W., Prasad, A. A., Kay, M., and Taylor, R. A. (2014). Direct normal irradiance forecasting and its application to concentrated solar thermal output forecasting – A review. *Solar Energy*, 108:287 – 307.

Lee, J., Weger, R. C., Sengupta, S. K., and Welch, R. M. (1990). A neural network approach to cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, 28(5):846–855.

Lee, J. A., Haupt, S. E., Jiménez, P. A., Rogers, M. A., Miller, S. D., and McCandless, T. C. (2017). Solar Irradiance Nowcasting Case Studies near Sacramento. *Journal of Applied Meteorology and Climatology*, 56(1):85–108.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Schroedter Homscheidt, M., Hoyer-Klick, C., Arola, A., Benedetti, A., Kaiser, J. W., and Morcrette, J.-J. (2013). McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmospheric Measurement Techniques*, 6:2403–2418.

Lew, D., Brinkman, G., Ibanez, E., Hummon, M., Hodge, B. M., Heaney, M., and King, J. (2012). Sub-Hourly Impacts of High Solar Penetrations in the Western United States: Preprint. United States.

Li, Q., Lu, W., and Yang, J. (2011). A hybrid thresholding algorithm for cloud detection on ground-based color images. *Journal of Atmospheric and Oceanic Technology*, 28(10):1286–1296.

Li, Q., Zhang, Z., Lu, W., Yang, J., Ma, Y., and Yao, W. (2016). From pixels to patches: a cloud classification method based on a bag of micro-structures. *Atmospheric Measurement Techniques*, 9(2):753–764.

Li, Y., Thompson, D. W. J., Stephens, G. L., and Bony, S. (2014). A global survey of the instantaneous linkages between cloud vertical structure and large-scale climate. *Journal of Geophysical Research: Atmospheres*, 119(7):3770–3792.

Liaw, A., Wiener, M., et al. (2002). Classification and Regression by randomForest. *R news*, 2(3):18–22.

Liberzon, A., Gurka, R., and Taylor, Z. (2009). Openpiv home page.

Lohmann, S., Schillings, C., Mayer, B., and Meyer, R. (2006). Long-term variability of solar direct and global radiation derived from ISCCP data and comparison with reanalysis data. *Solar Energy*, 80(11):1390 – 1401. European Solar Conference (EuroSun 2004).

Long, C. N. and DeLuisi, J. J. (1998). Development of an automated hemispheric sky imager for cloud fraction retrievals. In *Proc. 10th Symp. on meteorological observations and instrumentation*, pages 171–174.

Long, C. N. and Dutton, E. G. (2002). Bsrn global network recommended qc tests, v2.0. Technical report, Baseline Surface Radiation Network (BSRN).

Long, C. N. and Dutton, E. G. (2010). BSRN Global Network recommended QC tests, V2.0.

Long, C. N., Sabburg, J. M., Calbó, J., and Pagès, D. (2006). Retrieving Cloud Characteristics from Ground-Based Daytime Color All-Sky Images. *Journal of Atmospheric and Oceanic Technology*, 23(5):633–652.

Lorenz, E. and Heinemann, D. (2012). 1.13 - Prediction of Solar Irradiance and Photovoltaic Power. In Sayigh, A., editor, *Comprehensive Renewable Energy*, pages 239 – 292. Elsevier, Oxford.

Lorenz, E., Kühnert, J., and Heinemann, D. (2012a). 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*, volume 2428, page 44014405.

Lorenz, E., Kühnert, J., and Heinemann, D. (2012b). Short Term Forecasting of Solar Irradiance by Combining Satellite Data and Numerical Weather Predictions. In *27th European Photovoltaic Solar Energy Conference and Exhibition*, pages 4401 – 4405.

Lorenz, E., Kühnert, J., Heinemann, D., Nielsen, K. P., Remund, J., and Müller, S. C. (2016). Comparison of global horizontal irradiance forecasts based on numerical weather prediction models with different spatio-temporal resolutions. *Progress in Photovoltaics: Research and Applications*, 24(12):1626–1640.

Lorenz, E., Remund, J., Müller, S. C., Traunmüller, W., Steinmaurer, G., Pozo, D., Ruiz-Arias, J. A., Fanego, V. L., Ramirez, L., Romeo, M. G., et al. (2009). Benchmarking of different approaches to forecast solar irradiance. In *24th European photovoltaic solar energy conference*, pages 21–25. Hamburg, Germany.

Lu, S., Youngdeok Hwang, Khabibrakhmanov, I., Marianno, F. J., Xiaoyan Shao, Zhang, J., Hodge, B., and Hamann, H. F. (2015). Machine learning based multi-physical-model blending for enhancing renewable energy forecast - improvement via situation dependent error correction. In *2015 European Control Conference (ECC)*, pages 283–290.

Lucas, B. D., Kanade, T., et al. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision. *IJCAI*, pages 674–679.

Mannstein, H., Brömser, A., and Bugliaro, L. (2010). Ground-based observations for the validation of contrails and cirrus detection in satellite imagery. *Atmospheric Measurement Techniques*, 3(3):655–669.

Marcos, J., Marroyo, L., Lorenzo, E., and García, M. (2012). Smoothing of pv power fluctuations by geographical dispersion. *Progress in Photovoltaics: Research and Applications*, 20(2):226–237.

Marquez, R. and Coimbra, C. F. (2013). Intra-hour DNI forecasting based on cloud tracking image analysis. *Solar Energy*, 91:327 – 336.

Martínez-Chico, M., Batlles, F., and Bosch, J. (2011). Cloud classification in a mediterranean location using radiation data and sky images. *Energy*, 36(7):4055 – 4062.

Martucci, G., Milroy, C., and O'Dowd, C. D. (2010). Detection of Cloud-Base Height Using Jenoptik CHM15K and Vaisala CL31 Ceilometers. *Journal of Atmospheric and Oceanic Technology*, 27(2):305–318.

Mateos, D., Antón, M., Valenzuela, A., Cazorla, A., Olmo, F., and Alados-Arboledas, L. (2014). Efficiency of clouds on shortwave radiation using experimental data. *Applied Energy*, 113:1216 – 1219.

Mathiesen, P. and Kleissl, J. (2011). Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States. *Solar Energy*, 85(5):967 – 977.

Matos, M., Bessa, R., Botterud, A., and Zhou, Z. (2017). 11 - forecasting and setting power system operating reserves. In Kariniotakis, G., editor, *Renewable Energy Forecasting*, Woodhead Publishing Series in Energy, pages 279 – 308. Woodhead Publishing.

Matuszko, D. (2012). Influence of the extent and genera of cloud cover on solar radiation intensity. *International Journal of Climatology*, 32(15):2403–2414.

McCandless, T., Haupt, S., and Young, G. (2015). A model tree approach to forecasting solar irradiance variability. *Solar Energy*, 120:514 – 524.

McCandless, T., Haupt, S., and Young, G. (2016a). A regime-dependent artificial neural network technique for short-range solar irradiance forecasting. *Renewable Energy*, 89:351 – 359.

McCandless, T. C., Young, G. S., Haupt, S. E., and Hinkelman, L. M. (2016b). Regime-Dependent Short-Range Solar Irradiance Forecasting. *Journal of Applied Meteorology and Climatology*, 55(7):1599–1613.

Mellit, A. (2008). Artificial Intelligence technique for modelling and forecasting of solar radiation data: a review. *International Journal of Artificial Intelligence and Soft Computing*, 1(1):52 – 76.

Miller, S. D., Rogers, M. A., Haynes, J. M., Sengupta, M., and Heidinger, A. K. (2018). Short-term solar irradiance forecasting via satellite/model coupling. *Solar Energy*, 168:102 – 117. Advances in Solar Resource Assessment and Forecasting.

Mills, A. and Wiser, R. (2010). Implications of Wide-Area Geographic Diversity for Short-Term Variability of Solar Power. Technical report, Ernest Orlando Lawrence Berkeley National Laboratory. LBNL-3884E.

Moreno-Tejera, S., Silva-Pérez, M., Lillo-Bravo, I., and Ramírez-Santigosa, L. (2016). Solar resource assessment in seville, spain. statistical characterisation of solar radiation at different time resolutions. *Solar Energy*, 132:430 – 441.

Moreno-Tejera, S., Silva-Pérez, M., Ramírez-Santigosa, L., and Lillo-Bravo, I. (2017). Classification of days according to dni profiles using clustering techniques. *Solar Energy*, 146:319 – 333.

Mori, N. and 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. and Coimbra, C. F. (2014). Streamline-based method for intra-day solar forecasting through remote sensing. *Solar Energy*, 108:447 – 459.

Nouri, B., Wilbert, S., Kuhn, P., Hanrieder, N., Schroedter-Homscheidt, M., Kazantzidis, A., Zarzalejo, L., Blanc, P., Kumar, S., Goswami, N., Shankar, R., Affolter, R., and Pitz-Paal, R. (2019). Real-Time Uncertainty Specification of All Sky Imager Derived Irradiance Nowcasts. *Remote Sensing*, 11(9).

OECD/IEA (2018). World Energy Outlook 2018. techreport, International Energy Agency (I.E.A.) Secure Sustainable Together.

Ohba, M., Kadokura, S., and Nohara, D. (2016). Impacts of synoptic circulation patterns on wind power ramp events in east japan. *Renewable Energy*, 96:591 – 602.

Peng, Z., Yu, D., Huang, D., Heiser, J., and Kalb, P. (2016). A hybrid approach to estimate the complex motions of clouds in sky images. *Solar Energy*, 138:10 – 25.

Peng, Z., Yu, D., Huang, D., Heiser, J., Yoo, S., and Kalb, P. (2015). 3D cloud detection and tracking system for solar forecast using multiple sky imagers. *Solar Energy*, 118:496 – 519.

Perez, R., Cebecauer, T., and Šúri, M. (2013). Chapter 2 - Semi-Empirical Satellite Models. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 21 – 48. Academic Press, Boston.

Perez, R., David, M., Hoff, T. E., Jamaly, M., Kivalov, S., Kleissl, J., Lauret, P., and Perez, M. (2016). Spatial and Temporal Variability of Solar Energy. *Foundations and Trends® in Renewable Energy*, 1(1):1–44.

Perez, R. and Hoff, T. E. (2013a). Chapter 6 - Solar Resource Variability. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 133 – 148. Academic Press, Boston.

Perez, R. and Hoff, T. E. (2013b). *Solar Resource Variability*, chapter 6, pages 133 – 148. Academic Press, first edition. eBook ISBN: 9780123977724.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., and Vignola, F. (2002). A new operational model for satellite-derived irradiances: description and validation. *Solar Energy*, 73(5):307 – 317.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., and Hoff, T. (2011). Parameterization of site-specific short-term irradiance variability. *Solar Energy*, 85(7):1343 – 1353.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., and Hoff, T. E. (2012). Short-term irradiance variability: Preliminary estimation of station pair correlation as a function of distance. *Solar Energy*, 86(8):2170 – 2176. Progress in Solar Energy 3.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., and Hoff, T. E. (2010). Validation of short and medium term operational solar radiation forecasts in the US. *Solar Energy*, 84(12):2161 – 2172.

Pierro, M., Felice, M. D., Maggioni, E., Moser, D., Perotto, A., Spada, F., and Cornaro, C. (2017). Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data. *Solar Energy*, 158:1026 – 1038.

Pincus, R., Patrick Hofmann, R. J., Anderson, J. L., Raeder, K., Collins, N., and Whitaker, J. S. (2011). Can Fully Accounting for Clouds in Data Assimilation Improve Short-Term Forecasts by Global Models? *Monthly Weather Review*, 139(3):946–957.

Pozo-Vazquez, D., Santos-Alamillos, F. J., Lara-Fanego, V., Ruiz-Arias, J. A., and Tovar-Pescador, J. (2011). *The Impact of the NAO on the Solar and Wind Energy Resources in the Mediterranean Area*, chapter 15, pages 213–231. Springer Netherlands, Dordrecht.

Pozo-Vázquez, D., Tovar-Pescador, J., Gámiz-Fortis, S. R., Esteban-Parra, M. J., and Castro-Díez, Y. (2004). Nao and solar radiation variability in the european north atlantic region. *Geophysical Research Letters*, 31(5):n/a–n/a. L05201.

Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., Klüser, L., Lefèvre, M., Saboret, L., Schroedter-Homscheidt, M., and Wald, L. (2017). Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method. *Meteorologische Zeitschrift*, 26(1):33–57.

Quesada-Ruiz, S., Chu, Y., Tovar-Pescador, J., Pedro, H., and Coimbra, C. (2014). Cloud-tracking methodology for intra-hour DNI forecasting. *Solar Energy*, 102:267 – 275.

Ramírez, L. and Vindel, J. (2017). 13 - Forecasting and nowcasting of DNI for concentrating solar thermal systems. In Blanco, M. J. and Santigosa, L. R., editors, *Advances in Concentrating Solar Thermal Research and Technology*, Woodhead Publishing Series in Energy, pages 293 – 310. Woodhead Publishing.

Red Eléctrica de España (2017). The Spanish Electricity System. Preliminary Report 2017. Technical report, Red Eléctrica de España.

Redes Energéticas Nacionais (2018). Technical Data 2018. techreport, Redes Energéticas Nacionais.

Remund, J., Wald, L., Lefèvre, M., Ranchin, T., and Page, J. H. (2003). Worldwide Linke turbidity information. In *ISES Solar World Congress 2003*, volume CD-ROM, page 13 p, Göteborg, Sweden. International Solar Energy Society (ISES).

Renné, D. S. (2014). *Emerging Meteorological Requirements to Support High Penetrations of Variable Renewable Energy Sources: Solar Energy*, pages 257–273. Springer New York, New York, NY.

Reno, M. and Stein, J. (2013). Using cloud classification to model solar variability. In *Proceedings of the National Solar Conference for 2013 (SOLAR 2013), At Baltimore, MD*.

Rigollier, C., Bauer, O., and Wald, L. (2000). On the clear sky model of the ESRA – European Solar Radiation Atlas – with respect to the heliosat method. *Solar Energy*, 68(1):33 – 48.

Rigollier, C., Lefèvre, M., and Wald, L. (2004). The method Heliosat-2 for deriving shortwave solar radiation from satellite images. *Solar Energy*, 77(2):159 – 169.

Rodríguez-Benítez, F. J., Arbizu-Barrena, C., Santos-Alamillos, F. J., Tovar-Pescador, J., and Pozo-Vázquez, D. (2018). Analysis of the intra-day solar resource variability in the Iberian Peninsula. *Solar Energy*, 171:374 – 387.

Roesch, A., Wild, M., Ohmura, A., Dutton, E. G., Long, C. N., and Zhang, T. (2011). Assessment of bsrn radiation records for the computation of monthly means. *Atmospheric Measurement Techniques*, 4(2):339–354.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53 – 65.

Ruiz-Arias, J. A., Gueymard, C. A., Santos-Alamillos, F. J., and Pozo-Vázquez, D. (2016). Worldwide impact of aerosol's time scale on the predicted long-term concentrating solar power potential. *Scientific Reports*, 6.

Salazar, E., Sansó, B., Finley, A. O., Hammerling, D., Steinsland, I., Wang, X., and Delamater, P. (2011). Comparing and Blending Regional Climate Model Predictions for the American Southwest. *Journal of Agricultural, Biological, and Environmental Statistics*, 16(4):586–605.

Salgueiro, V., Costa, M. J., Silva, A. M., and Bortoli, D. (2016). Effects of clouds on the surface shortwave radiation at a rural inland mid-latitude site. *Atmospheric Research*, 178-179:95 – 101.

Sancho Ávila, J. M., Riesco Martín, J., Jiménez Alonso, C., Sánchez de Cos, M. d. C., Montero Cadalso, J., and López Bartolomé, M. (2012). Atlas de radiación solar en españa utilizando datos del saf de clima de eumetsat. *Arcimís*, pages 272–275.

Santos-Alamillos, F., Pozo-Vázquez, D., Ruiz-Arias, J., Lara-Fanego, V., and 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. *Renewable Energy*, 69:147 – 156.

Santos-Alamillos, F. J., Brayshaw, D. J., Methven, J., Thomaidis, N. S., Ruiz-Arias, J. A., and Pozo-Vázquez, D. (2017a). Exploring the meteorological potential for planning a high performance European electricity super-grid: optimal power capacity distribution among countries. *Environmental Research Letters*, 12(11):114030.

Santos-Alamillos, F. J., Brayshaw, D. J., Methven, J., Thomaidis, N. S., Ruiz-Arias, J. A., and Pozo-Vázquez, D. (2017b). Exploring the meteorological potential for planning a high performance european electricity super-grid: optimal power capacity distribution among countries. *Environmental Research Letters*, 12(11):114030.

Santos-Alamillos, F. J., Pozo-Vázquez, D., Ruiz-Arias, J. A., Lara-Fanego, V., and Tovar-Pescador, J. (2012). Analysis of Spatiotemporal Balancing between Wind and Solar Energy Resources in the Southern Iberian Peninsula. *Journal of Applied Meteorology and Climatology*, 51(11):2005–2024.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A. (2002). AN INTRODUCTION TO METEOSAT SECOND GENERATION (MSG). *Bulletin of the American Meteorological Society*, 83(7):977–992.

Schmidt, T., Kalisch, J., Lorenz, E., and Heinemann, D. (2016). Evaluating the spatio-temporal performance of sky-imager-based solar irradiance analysis and forecasts. *Atmospheric Chemistry and Physics*, 16(5):3399–3412.

Schmidt, Thomas, Calais, Martina, Roy, Eric, Burton, Ashton, Heinemann, Detlev, Kilper, Thilo, and Carter, Craig (2017). Short-term solar forecasting based on sky images to enable higher PV generation in remote electricity networks. *Renew. Energy Environ. Sustain.*, 2:23.

Scholkopf, B. and Smola, A. J. (2001). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, USA.

Schroedter-Homscheidt, M., Hoyer-Klick, C., Rikos, E., Tselepsis, S., and Pulvermüller, B. (2009). Nowcasting and forecasting of solar irradiance for energy electricity generation. In *SolarPACES Conf.*

Schroedter-Homscheidt, M., Killius, N., Guevara, D. M., Sirch, T., Hanrieder, N., Wilbert, S., and Yasser, Z. (2018). Satellite-based DNI nowcasting based on a sectoral atmospheric motion approach. *AIP Conference Proceedings*, 2033(1):190015.

Scolari, E., Sossan, F., Haure-Touzé, M., and Paolone, M. (2018). Local estimation of the global horizontal irradiance using an all-sky camera. *Solar Energy*, 173:1225 – 1235.

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., and 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. This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Singh, M. and Glennen, M. (2005). Automated ground-based cloud recognition. *Pattern Analysis and Applications*, 8(3):258–271.

Sirch, T., Bugliaro, L., Zinner, T., Möhrlein, M., and Vazquez-Navarro, M. (2017). Cloud and DNI nowcasting with MSG/SEVIRI for the optimized operation of concentrating solar

power plants. *Atmospheric Measurement Techniques*, 10(2):409–429.

Skamarock, W. C. and Coauthors (2008). A Description of the Advanced Research WRF Version 3. Technical report, NCAR Technical Note NCAR/TN-475+STR.

Smith, A. R. (1978). Color Gamut Transform Pairs. *SIGGRAPH Comput. Graph.*, 12(3):12–19.

Steiner, A., Köhler, C., Metzinger, I., Braun, A., Zirkelbach, M., Ernst, D., Tran, P., and Ritter, B. (2017). Critical weather situations for renewable energies – part a: Cyclone detection for wind power. *Renewable Energy*, 101:41 – 50.

Taravat, A., Del Frate, F., Cornaro, C., and Vergari, S. (2015). Neural Networks and Support Vector Machine Algorithms for Automatic Cloud Classification of Whole-Sky Ground-Based Images. *IEEE Geoscience and Remote Sensing Letters*, 12(3):666–670.

Tascikaraoglu, A. and Uzunoglu, M. (2014). A review of combined approaches for prediction of short-term wind speed and power. *Renewable and Sustainable Energy Reviews*, 34:243 – 254.

Thompson, G. and Eidhammer, T. (2014). A Study of Aerosol Impacts on Clouds and Precipitation Development in a Large Winter Cyclone. *Journal of the Atmospheric Sciences*, 71(10):3636–3658.

Trigo, R. M., Osborn, T. J., and Corte-Real, J. M. (2002a). The north atlantic oscillation influence on europe: climate impacts and associated physical mechanisms. *Climate Research*, 20(1):9–17.

Trigo, R. M., Osborn, T. J., and Corte-Real, J. M. (2002b). The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms. *Clim Res*, 20(1):9–17.

Trigo, R. M., Pozo-Vázquez, D., Osborn, T. J., Castro-Díez, Y., Gámiz-Fortis, S., and Esteban-Parra, M. J. (2004). North atlantic oscillation influence on precipitation, river flow and water resources in the iberian peninsula. *International Journal of Climatology*, 24(8):925–944.

Tuohy, A., Zack, J., Haupt, S. E., Sharp, J., Ahlstrom, M., Dise, S., Grimit, E., Mohrlen, C., Lange, M., Casado, M. G., Black, J.,

Marquis, M., and Collier, C. (2015). Solar Forecasting: Methods, Challenges, and Performance. *IEEE Power and Energy Magazine*, 13(6):50–59.

Tzoumanikas, P., Nikitidou, E., Bais, A., and Kazantzidis, A. (2016). The effect of clouds on surface solar irradiance, based on data from an all-sky imaging system. *Renewable Energy*, 95:314 – 322.

Urquhart, B., Ghonima, M., Nguyen, D. A., Kurtz, B., Chow, C. W., and Kleissl, J. (2013). Chapter 9 - Sky-Imaging Systems for Short-Term Forecasting. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 195 – 232. Academic Press, Boston.

Vislocky, R. L. and Fritsch, J. M. (1995). Improved Model Output and Statistics through Model Consensus. *Bulletin of the American Meteorological Society*, 76(7):1157–1164.

Viúdez-Mora, A., Costa-Surós, M., Calbó, J., and González, J. A. (2015). Modeling atmospheric longwave radiation at the surface during overcast skies: The role of cloud base height. *Journal of Geophysical Research: Atmospheres*, 120(1):199–214.

Voyant, C., Muselli, M., Paoli, C., and Nivet, M.-L. (2012). Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation. *Energy*, 39(1):341 – 355. Sustainable Energy and Environmental Protection 2010.

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., and Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105:569 – 582.

Wacker, S., Gröbner, J., Zysset, C., Diener, L., Tzoumanikas, P., Kazantzidis, A., Vuilleumier, L., Stöckli, R., Nyeki, S., and Kämpfer, N. (2015). Cloud observations in Switzerland using hemispherical sky cameras. *Journal of Geophysical Research: Atmospheres*, 120(2):695–707.

Wang, P., van Westrhenen, R., Meirink, J. F., van der Veen, S., and Knap, W. (2019). Surface solar radiation forecasts by advecting cloud physical properties derived from Meteosat Second Generation observations. *Solar Energy*, 177:47 – 58.

Weinberger, K. Q. and Saul, L. K. (2009). Distance Metric Learning for Large Margin Nearest Neighbor Classification. *Journal of Machine Learning Research*, 10(Feb):207–244.

Weinzaepfel, P., Revaud, J., Harchaoui, Z., and Schmid, C. (2013). DeepFlow: Large displacement optical flow with deep matching. In *ICCV - IEEE International Conference on Computer Vision*, pages 1385–1392, Sydney, Australia. IEEE.

Wess, S., Althoff, K.-D., and Derwand, G. (1994). Using k-d trees to improve the retrieval step in case-based reasoning. In Wess, S., Althoff, K.-D., and Richter, M. M., editors, *Topics in Case-Based Reasoning*, pages 167–181, Berlin, Heidelberg. Springer Berlin Heidelberg.

West, S. R., Rowe, D., Sayeef, S., and Berry, A. (2014). Short-term irradiance forecasting using skycams: Motivation and development. *Solar Energy*, 110:188 – 207.

Wey, E. and Schroedter-Homscheidt, M. (2014). APOLLO Cloud Product Statistics. *Energy Procedia*, 49:2414 – 2421. Proceedings of the SolarPACES 2013 International Conference.

Wiegner, M., Madonna, F., Binietoglou, I., Forkel, R., Gasteiger, J., Geiß, A., Pappalardo, G., Schäfer, K., and Thomas, W. (2014). What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET. *Atmospheric Measurement Techniques*, 7(7):1979–1997.

Wiemken, E., Beyer, H., Heydenreich, W., and Kiefer, K. (2001). Power characteristics of pv ensembles: experiences from the combined power production of 100 grid connected pv systems distributed over the area of germany. *Solar Energy*, 70(6):513 – 518. Grid-Connected PV.

Wild, M., Folini, D., Schär, C., Loeb, N., Dutton, E. G., and König-Langlo, G. (2013). The global energy balance from a surface perspective. *Climate Dynamics*, 40(11):3107–3134.

Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences*, volume 91 of *International Geophysics Series*. Academic Press, second edition. ISBN: 978-0-12-751966-1.

Wolff, B., Kühnert, J., Lorenz, E., Kramer, O., and 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. *Solar Energy*, 135:197 – 208.

World Meteorological Organization. International Cloud Atlas (2017). Preface to the 2017 edition. Web.

World Weather Research Programme. WWRP/WGNE Joint Working Group on Forecast Verification Research (JWGFVR) (2012). Recommended Methods for Evaluating Cloud and Related Parameters. Technical report, World Meteorological Organization, 7 bis, avenue de la Paix - P.O. Box 2300 - CH 1211 Geneva 2 - Switzerland.

www.ree.es (2018). The Spanish Electricity System. End of year forecast 2018. techreport, Red Eléctrica de España.

Xiao, L., Wang, J., Dong, Y., and Wu, J. (2015). Combined forecasting models for wind energy forecasting: A case study in China. *Renewable and Sustainable Energy Reviews*, 44:271 – 288.

Xie, P. and Arkin, P. A. (1996). Analyses of Global Monthly Precipitation Using Gauge Observations, Satellite Estimates, and Numerical Model Predictions. *Journal of Climate*, 9(4):840–858.

Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T., and Coimbra, C. F. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy*, 168:60 – 101. Advances in Solar Resource Assessment and Forecasting.

Zagouras, A., Pedro, H. T., and Coimbra, C. F. (2014). Clustering the solar resource for grid management in island mode. *Solar Energy*, 110:507 – 518.

Zamo, M., Mestre, O., Arbogast, P., and Pannekoucke, O. (2014a). A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production. *Solar Energy*, 105:792 – 803.

Zamo, M., Mestre, O., Arbogast, P., and Pannekoucke, O. (2014b). A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production. Part II: Probabilistic forecast of daily production. *Solar Energy*, 105:804 – 816.

Zhang, J., Florita, A., Hodge, B.-M., Lu, S., Hamann, H. F., Banunarayanan, V., and Brockway, A. M. (2015a). A suite of metrics for assessing the performance of solar power forecasting. *Solar Energy*, 111:157 – 175.

Zhang, J., Hodge, B.-M., Lu, S., Hamann, H. F., Lehman, B., Simmons, J., Campos, E., Banunarayanan, V., Black, J., and Tedesco, J. (2015b). Baseline and target values for regional and point PV power forecasts: Toward improved solar forecasting. *Solar Energy*, 122:804 – 819.

Zhen, Z., Wang, F., Sun, Y., Mi, Z., Liu, C., Wang, B., and Lu, J. (2015). SVM based cloud classification model using total sky images for PV power forecasting. In *2015 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5.