UNIVERSITY OF JAÉN

Polytechnic School
DEPARTMENT OF PHYSICS

PhD Thesis

●

ASSESSMENT AND FORECASTING OF SOLAR RESOURCE: APPLICATIONS TO THE SOLAR ENERGY INDUSTRY

By:
VICENTE LARA FANEGO

Thesis Advisors:
A. DAVID POZO VÁZQUEZ
JOAQUÍN TOVAR PESCADOR
JOSÉ A. RUIZ ARIAS

Jaén, June 2017

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

El **Dr. Joaquín Tovar Pescador**, Catedrático de Universidad del Departamento de Física de la Universidad de Jaén, España

Y el **Dr. José Antonio Ruiz Arias**, Experto en Radiación Solar y Simulación Meteorológica del Departamento de I+D de Solargis s.r.o., Eslovaquia

CERTIFICAN:

Que la presente memoria, titulada "**Assessment and forecasting of solar resource: applications to the solar energy industry**", ha sido realizada bajo su dirección. Y considerando que representa trabajo de Tesis, autorizan su presentación y defensa para optar al grado de Doctor con mención de Doctor Internacional.

Dr. A. David Pozo Vázquez
Departamento de Física
Universidad de Jaén

Dr. Joaquín Tovar Pescador
Departamento de Física
Universidad de Jaén

Dr. José Antonio Ruiz Arias
Departamento de I+D
Solargis s.r.o.

Memoria presentada para optar al grado de:
Doctor, con mención de Doctor Internacional

Vicente Lara Fanego
Licenciado en Física

Jaén, junio de 2017

A María,
mi vida

A mis padres,
que siempre me lo han dedicado todo

No se trata sólo de prever el futuro, sino de hacerlo posible.
Antoine de Saint-Exupéry (escritor francés)

It is important not merely to foresee the future, but to bring it about.
Antoine de Saint-Exupéry (French writer)

Agradecimientos

"Ayer se fue; mañana no ha llegado; hoy se está yendo sin parar un punto. Soy un fue y un será y un es cansado". Al escribir estas últimas frases al finalizar la elaboración de este texto, no puedo evitar pensar en estas palabras de Quevedo, si bien con una gran sensación de alegría al mismo tiempo. Ciertamente la elaboración de la tesis es un esfuerzo formidable y, sin duda, casi imposible sin la ayuda y el apoyo de muchas personas con las que tengo la inmensa fortuna de entrelazar mi vida. A ellas, a todas ellas, les debo de alguna forma haber podido llegar hasta aquí, este instante en el que me detengo al fin a darles las gracias. Para ellos dedico estas líneas con el mayor afecto.

En primer lugar quiero darles las gracias muy sinceramente a mis directores de tesis, el Dr. A. David Pozo Vázquez, el Dr. Joaquín Tovar Pescador y el Dr. José Antonio Ruiz Arias. Ellos me abrieron la puerta al mundo de la investigación que tan felizmente ha determinado mi vida todos estos años. Es mucho más que sus enseñanzas, su apoyo y su ánimo lo que les debo. Les debo para siempre una sincera amistad.

Dentro de esta familia que es el grupo de investigación MATRAS, (familia en el sentido figurado más real que exista) también he tenido la gran suerte de encontrar grandes amigos y compañeros. José Antonio, Paco, Samuel, Álvaro, Pepe, y más recientemente Clara y Javi. También quiero dedicarle un recuerdo al bueno de Husain, allá donde esté en Palestina. Con todos he compartido mucho y he aprendido mucho. Y sobre todo he sido feliz. Si los méritos para el doctorado se alcanzasen a base de reir, nosotros podríamos ser ya honoris causa varias veces. Gracias, muchas gracias, a todos.

Es también una gran alegría poder dedicar en esta hoja unas palabras de agradecimiento a una persona que es para mí más que un gran amigo y mentor. Gracias Paco Pepe, por estar siempre ahí.

También quiero aprovechar un pedacito de este papel para darle las gracias a una persona de la que aprendo mucho profesionalmente, y que

además me ha enseñado una nueva perspectiva de la bondad y la naturaleza de ayudar a los demás. Muchas gracias Chris.

Quiero también acordarme de Rafa, mi amigo y hermano (porque en este caso son casi sinónimos). Y por supuesto de Cris, que también es para siempre. El poder contar con ambos es algo esencial en mi vida, y en parte terminar esta etapa también se ha logrado gracias a eso.

Igualmente quiero agradecer el ánimo sincero de mis amigos David y David, de nuestro Pelícano. Siempre me resulta muy vitalizante teneros cerca y sentir vuestro apoyo.

Este renglón va dedicado a dos personas especiales. A mi cuñada la mayor, Monti, que tan bien me cuida cuando nos vemos, a pesar de las largas jornadas mías delante del ordenador. Muchas gracias por la ayuda, el ánimo y el cariño, Monti. Y por supuesto a Heraclio, al que echo de menos verlo reír y mirarme con cara de extrañeza porque no entiende mi acento mal pronunciado. Sé que estarás muy feliz viéndonos llegar hasta aquí.

Quiero también dar las gracias a mis hermanas, Tuti y Laura, que a pesar de la distancia siempre me ha sabido hacer llegar su ánimo y cariño. Otra de esas cosas que son para siempre. ¡Qué bien sienta! Muchas gracias a las dos.

Expresar la infinita gratitud que siento hacia mis padres me es imposible. Ellos son para mí el significado pleno de la confianza y el cariño: Los he sentido en cada instante de mi vida. Si estoy escribiendo estas palabras y todas las demás, es gracias a vosotros, y a vosotros os las dedico. Gracias de corazón.

Mi pensamiento final, junto con todo este trabajo y toda mi vida, lo dedico a mi mujer, María. No sabría encontrar en la eternidad suficientes palabras para agradecerte de corazón tanto, por todo. Sin ti no habría logrado esto. Sin ti no habría nada. Gracias, mi vida.

Abstract

In this thesis a study of the assessment and forecasting of the solar resource for its application in the solar industry is carried out. The main objective is to improve the knowledge about various aspects of solar radiation as primary energy source. The aim is to contribute to the development of this renewable energy within the current framework of transformation of the energy sector worldwide, strongly conditioned by climate change and the need for sustainable development. However, despite the relentless technological development and the considerable cost reductions, its degree of introduction at large-scale into power systems is still far from its real potential. This is due mainly to the fact that, although solar radiation is the most abundant primary energy source in the planet, it naturally presents a great spatial and temporal variability. This characteristic constitutes the major source of uncertainty in the development of solar projects, both in the initial phase of feasibility study and during the phase of operation. In order to contribute to the reduction of this uncertainty, the research work carried out in this thesis has developed and evaluated methods for the characterization and estimation of surface solar irradiance, both components: global (GHI) and direct (DNI).

Firstly, concerning solar resource assessment a novel method has been developed to obtain a representative year for the characterization of solar irradiance at multi-year scales in a particular location of interest. This method, named EVA, allows the generation of so-called typical solar years (TSY). These artificial years are widely used within the solar industry, mainly in the bankability analysis of solar projects. In particular, the use of TSYs has become a standard for estimating the production of solar plants at different conditions of solar resource availability. However, nowadays there is no scientific consensus on obtaining TSYs. This leads to greater uncertainty, since the application of different methodologies on the same dataset can result in different TSYs. Therefore the standardization of the method for the generation of the TSYs is a demand of the industry. In this sense, the method developed in this thesis aims to contribute to such standardization. The

EVA method has a statistical basis and its formulation is completely analytical. The most important feature of the method is that it allows generate TSYs for any scenario of annual solar energy availability. This includes situations of low annual values of solar irradiation, which are fundamental in the bankability studies of solar projects. The results of the evaluation show that the method provides consistent results for the two components (GHI and DNI), with low errors for any probability of exceedance. In addition, the method preserves long-term statistics at high temporal resolutions.

On the other hand, regarding solar resource forecasting, the use of the numerical weather prediction (NWP) model Weather Research and Forecasting (WRF) for the prediction of surface solar irradiance has been in-depth analyzed. The solar resource forecast is essential for the integration of solar energy into the power supply structures, which must always operate under the principles of safety and stability. It is also essential for the operation and management of plants, as well as for energy trading. Therefore, it is decisive for improving the level of competitiveness of solar energy. Among the tools for solar radiation forecasting, NWP models stand out as the single most powerful tools. In particular, the WRF model is one of the most advanced that currently exists. However, NWP models are not specifically devised for solar energy applications. Thus, in this thesis is presented a comprehensive evaluation of GHI and DNI forecasts reliability provided by the WRF model. The analysis is carried out by distinguishing different sky conditions, seasons and forecast horizons. In a second step, the predictions are evaluated within the frame of reference of a benchmarking exercise with other NWP models. Finally, based on the results obtained in the previous works, a study is carried out to analyze the role played by the horizontal spatial resolution with respect to the reliability of the solar radiation forecasts. The performance is also evaluated after the application of a post-processing based on a spatial averaging to the model raw outputs. In general the results show that WRF tends to overestimate the surface solar irradiance. In addition, the quality of the predictions, measured in terms of the usual statistical errors, is better for lower spatial resolutions. On the other hand, the

results of the benchmarking show that the forecasts from global models present errors lower than those of the regional models, like WRF. However, it is concluded that the convenience of using one type of model or another depends on the final application.

Contents

List of acronyms

List of acronyms

AEMet	Agencia Estatal de Meteorología
AI	Artificial intelligence
ANN	Artificial neural networks
AOD	Aerosol optical depth
ARIMA	Autoregressive integrated moving average model
ARPS	Advanced Multiscale Regional Prediction System
BSRN	Baseline Surface Radiation Network
CC	Climate Central
CDFmeas	Cumulative measured frequency distribution
CDFpred	Cumulative predicted frequency distribution
CENER	Centro Nacional de Energías Renovables
CFSR	Climate Forecast System Reanalysis
CIEMAT	Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
CMC	Canadian Meteorological Centre
CMV	Cloud motion vectors
CPV	Concentrator photovoltaics technology
CSP	Concentrating solar power technology
CSRT	Clear-sky radiative transfer models
DNI	Direct Normal Irradiance
DSWRF	Downward shortwave radiation flux at the surface
ECMWF	European Centre of Medium-Range Weather Forecast
EEA	European Environment Agency
EMC	Environmental Modeling Center
EU	European Union
GCM	Global circulation models
GDPS	Global Deterministic Prediction System
GEM	Global Environmental Multiscale

GEOS-5	Goddard Earth Observing System Model Version 5
GFS	Global Forecast System
GHG	Greenhouse gases
GHI	Global Horizontal Irradiance
HIRLAM	High Resolution Limited Area Model
IEA	International Energy Agency
IEA SHC	International Energy Agency Solar Heating and Cooling Programme
IFS	Integrated Forecast System
Imas	maximum possible irradiance value
Imeas	measured Irradiance
IPCC	Intergovernmental Panel on Climate Change
Ipred	predicted irriadiance
KSI	Kolmogorov–Smirnov test integral
MAE	Mean absolute error
MASS	Mesoscale Atmospheric Simulation System
MBE	Mean bias error
MERRA2	Modern Era Retrospective Re-Analysis 2
MOS	Model output statistics
MSE	Mean square error
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NDFD	National Digital Forecast Database
NOAA	National Oceanic and Atmospheric Administration
NREL	National Renewable Energy Laboratory
NWP	Numerical weather prediction
PV	Photovoltaic technology
REE	Red Eléctrica de España (Spanish transmission system operator)

RMSE	Root mean square error
SURFRAD	Surface Radiation Network NOAA
TDY	Typical direct (normal irradiance) year
TGY	Typical global (horizontal irradiance) year
TMY	Typical meteorological year
TOA	Top of the atmosphere
TSI	Total Solar Irradiance
TSO	Transmission system operator
TSY	Typical solar year
UNFCCC	United Nations Framework Convention on Climate Change
USA	United States of America
VRE	Variable renewable energy
WMO	World Meteorological Organization
WRF	Weather Research and Forecasting model
WRF-ASRC	WRF-Model from Atmospheric Sciences
WRF-AWS	WRF Model from AWSTruepower
WRF-Meteotest	WRF Model from Meteotest
WRF-UJAEN	WRF Model from University of Jaen

Chapter 1
Background and research framework

1.1 Introduction

Nowadays the world is witnessing a revolution in the energy sector. This is driven by several factors among which it should be specially highlighted the continuous strong increment in the energy demand, particularly of electrical energy, and the social awareness on climate change due to the increment of greenhouse gases as a direct consequence of human activity (IPCC, 2014a). Both elements are partially interrelated –the primary energy transformation is one of the most important processes of anthropogenic emissions of greenhouse gases-. Nevertheless, the perception of the importance of these two aspects is unequal, as the increase in the energy demand directly concerns the economy and its consequences are perceived more rapidly. Meanwhile the climate change has effects that appear more gradually, but they are constant and could be catastrophic. Therefore, these two principal motivations, along with the relentless advance in knowledge and technology, are definitively favoring the current transformation of the energy paradigms of the last half century.

The climate change

Climate change is one of the greatest challenges facing the future of humankind. In order to study its causes and to foresee its possible consequences, the international community established the Intergovernmental Panel on Climate Change (IPCC) to carry out those studies and to fulfill the objective of proposing the measures to mitigate its effects. Within this organization and outside it, the scientific community has been also carrying out an important effort to achieve this goal. The IPCC considers climate change as any change in the mean and/or the variability of the properties of the state of climate over

an extended period due to any cause. The consideration of the United Nations Framework Convention on Climate Change (UNFCCC) restricts the causes only to those related, directly or indirectly, to the human activity. In both cases the fact is that the evidences demonstrate that the Earth's lower atmosphere is suffering an average global warming; in words of the IPCC: "scientific evidence for warming of the climate system is unequivocal". This is measured worldwide as a constant increment of global average temperatures of air and oceans. This average global warming together with its effects modifies the state of climate and constitutes what is called climate change. The immediately first consequences of it are currently visible, such as the melting of vast extensions of snow and ice, the rising of the average sea level and the alterations in the vital patterns of some animals and plants. The IPCC describes that impacts of climate change are also affecting agriculture, land and oceans ecosystems and water supplies, and highlights that this is occurring widespread all over the world (IPCC, 2014b). Furthermore, the World Health Organization (WHO) asserts that climate change has an impact in human health, because it affects the air quality, safe drinking water and agriculture. The WHO says that "climate change is expected to cause approximately 250,000 additional deaths per year, from malnutrition, diseases and heat stress" between 2030 and 2050. In Europe a recent report of the European Environment Agency (EEA) warns that is expected an augmentation of the meteorological extreme events in the continent due to climate change: heat waves, floods, droughts and more frequent and intense storms (EEA, 2017). It points out that currently there have been negative impacts in the economy and health. In the southern Europe, as it is the case of Spain, the EEA anticipates an increase of maximum temperatures, droughts, floods and wild fires, as well as less precipitations. It is expected also that the higher temperatures favor the propagation of insects that causes the redistribution and increment of diseases. In order to face this reality on December, 2015, the Conference of the Parties of the UNFCCC approved the Paris Agreement -signed by 195 nations- to upgrade the Kyoto Protocol of 1997. The aim of the agreement is to keep "a global temperature rise this century well below 2 degrees Celsius above pre-industrial levels

and to pursue efforts to limit the temperature increase even further to 1.5 degrees Celsius". The goal is to maintain the relatively stable climate conditions and minimize the impacts of the climate change. The 2°C threshold is the value agreed by the scientific community to preserve such conditions. In order to assure a more secure scenario the 1.5°C value was proposed.

There are different factors that could contribute, individually or added together, to the alteration of the average global temperature. In this sense, for instance there have been seven cycles of glaciation in the last 650,000 years. These cycles are produced by changes in the total amount of energy received on Earth from the Sun due to slightly variations in the terrestrial orbit. There are also other natural factors that can contribute to changes in the climate system, such as the activity of the Sun, volcanos and natural fluctuations of the concentration of greenhouse gases (GHG) in the atmosphere. In a natural way, part of the energy of the incoming solar radiation –approximately 30%- is reflected back to the outer space by clouds, aerosols and Earth's surface. Without the greenhouse effect of the atmosphere the temperature in the surface would be approximately -18°C, instead of the actual 15°C. Nevertheless, nowadays human activities have emitted sufficient carbon dioxide and other GHG to enhance the atmospheric greenhouse effect producing a global warming beyond these 15°C, evidencing a determinant anthropogenic cause of the climate change (IPCC, 2014a). In particular, carbon dioxide (CO_2) –one of the GHG that contributes more to the global warming- has increased remarkably its presence in the atmosphere; and it continuous to increase. At the beginning of the industrial era the concentration was 278 ppm, which is considered a balanced reference value of the climate system. In 2016, the World Meteorological Organization (WMO) announced that in 2015 the milestone of a globally averaged concentration of 400ppm of CO_2 has been reached, and maintained even after the end of the strong El Niño effect. Currently, at the beginning of 2017, reference observatories, such as those of Manua Loa in Hawai (NOAA, EEUU) and Izaña in Tenerife (AEMET, Spain), are registering values close to 410ppm, establishing a new record. According to the National Oceanic

and Atmospheric Administration (NOAA) of the United States, the annual rate of increase is 1.92 ppm. At the same time, recent observations of temperature anomalies measured by different scientific organisms show that the 1.5°C safety value could be closer than it was expected. Currently global temperature increment over the reference pre-industrial period is around 1°C. According to Pachauri *et al*. (2014), in the current situation the projected scenarios show that it is possible that the global average temperature of the surface will be 4°C higher than the preindustrial period by the end of 21st century. Furthermore, a recent publication of Climate Central (CC) shows that if current trends of the GHG emissions continue, it is possible that the 1.5°C threshold will be crossed between the years 2025-2030 (CC, 2016). Moreover, a recent research work (Crowther *et al*., 2016) has also concluded that the climate change could be happening much faster than it was thought, since the release of CO_2 stored in the soil is enhanced by the climate warming. The authors concluded that, in a conservative scenario, the increase of CO_2 emissions from soil stock due to climate warming could be around 12 to 17% of the expected anthropogenic emissions by 2050. Although there is a significant uncertainty, this work evidences that this additional contribution to the total CO_2 emitted to the atmosphere will stimulate a positive feedback warming and can decisively contribute to accelerate the climate change. This result compels to review current projections of climate change to incorporate these new emission values not taken into account previously. Crowther *et al*. (2016) also stated it is possible that the enhancement of the greenhouse effect could be now irreversible, as the system could have passed the point of no return. In any case, the IPCC warns that "without additional mitigation, and even with adaptation, warming by the end of the 21st century will lead to very high risk of severe, widespread and irreversible impacts globally". Therefore, to carry out decisive actions to dampen the effects of climate change is imperative.

According to the scenario described above, the international community agrees to promote coordinated politics to face this reality. In this sense, the effective strategies are based on reducing the impacts

of the climate change and also on carrying out specific efforts for adapting to it. To this end, the premise is that the world has to intensely accelerate the reduction of the GHG emissions. Among the particular policies planned by each country, the Paris Agreement is promoting the adoption of the so-called carbon pricing systems (such as emissions trading plans and carbon taxes) by the signing countries, particularly by the world's largest economies. This proposal is expected to strongly reduce the global emissions and thus mitigate the climate change. In addition, it is a favorable framework for renewable energies. Nowadays, the majority of the anthropogenic emissions are related to the industry activity and the transformation of energy. In particular, the 25% of the anthropogenic GHG emissions are generated by the processes of electricity and heat production (IPCC, 2014a). Therefore, the renewable energies –like solar energy- are eminent active reducers of GHG emissions. Then, the promotion of these sources of energy is one of the most important assets for mitigating the impacts of the climate change through a sustainable development.

Energy scenario

Despite that the impacts of climate change shall be strongly enough to motivate the evolution of the energy sector, the economic factor associated to this filed has also a decisive weight, as necessary determinant, because it is the engine that powers this transformation. In this sense, the current framework of renewable energies and its future projection is of key importance. In this very respect, the International Energy Agency (IEA) in its World Energy Outlook 2016 report (IEA, 2016a) predicts a global primary energy demand increment of 30% -mostly from developing countries- by 2040. Within this increment, the demand for electricity will increase more strongly than any other end-use energy. In particular, in the reference case the 37% of power generation will be from renewables compared to 23% today. The expected rate of increase for non-hydropower renewables –principally wind and solar- is in average 2.9% per year from 2012 to 2040, being the fastest-growing source of energy for electricity generation. In this way, non-hydropower renewables go from 5% of total world generation in 2012 to 40% in 2040. They will represent almost the 50% of the total

power capacity installed in the period 2015-2020 (IEA, 2015). At the same time, it is expected a strong growth of natural gas and –much less pronounced- oil consumption. In the reference case the energy carbon emissions due to the energy sector is estimated to reduce its annual average rate from 2.4% since 2000 to 0.5% in 2040. Nevertheless, the IEA recognizes that, despite it is a significant achievement, this rate does not allow reaching the objective of 2°C scenario of Paris Agreement. It will only limit the rise in average global temperatures to 2.7°C by 2100. Therefore, to meet the climate goals will be a very difficult challenge in reducing emissions and enhancing efficiency. Notwithstanding, the IEA stated that, while it would require bigger efforts, it is possible to reach the goal of the Paris Agreement by means of the appropriate policies that accelerate the growth of low carbon technologies and the energy efficiency. In this sense, the IEA stresses that it is very important to try to expand the use of the renewable energies to other key sectors from the standpoint of energy consumption, such as industry, building and transportation. For its part, in the European Union (EU) the commitment to the development of clean energies is very important. A common goal, binding for all member states, has been set to achieve a 20% energy consumption from renewable sources by 2020, as well as an agreement to achieve a reduction in greenhouse gas emissions from at least 20% compared to 1990 (European Directives 2009/28 / EC and 2009/29 / EC). According to Eurostat the share of the renewables in the total primary energy production was 25.4 % in 2014, an increment of 73.1% since 2004 (average increase of 5.6 % per year). The contribution to the total electricity generated from renewable energy sources was 27.5%. In this regard, Spain is one of the leading countries in the implementation of renewable energies. It is the fourth largest producer of renewable energy within the EU-28, with a 9.4% share of the total production. The contribution of these energy sources to its electric mix is already very important. According to the transmission system operator (TSO) of Spain (Red Eléctrica de España, REE) the percentage of electricity generation from renewable sources over the total generation in the peninsular Spain was 41.1% (REE, 2016). Furthermore, installed renewable power has allowed the net balance of electricity exchange -

via high-voltage interconnection-, between Spain and its neighboring countries (Morocco, Portugal and France), to be exporter (REE, 2015).

The framework described above highlights the resolute momentum that renewable energies are receiving globally, and particularly in Spain and the rest of Europe, by governments and the energy sector. Thus, the IEA points out that electricity generation is entering a period of transformation towards a safe and sustainable development environment in which renewable energies play a major role, most specially wind and solar energies. On the one hand this type of energies is progressively increasing its level of competitiveness thanks to the development of technology, which translates into a gradual lowering of their costs of generation and investment. On the other hand, it has allowed to establish a new map of the distribution of energy resources by introducing abundant and geographically distributed sources, such as solar radiation and wind. This allows to supply part of the energy demand through own resources and finally reduce the cost of the energy bill of the countries and their dependence on the outside. This is of particular interest in a country like Spain, whose external dependence on energy supply is around 80% (20 points above the European average). Hence, taking into account the expectations of a growing consumption, a suitable deployment of renewables is a convenient strategy and a key element in the economic projections of any country. In this respect, the IEA underlines that, despite the current lower prices of oil –some experts point out that prices will never rebound beyond $100/barrel-, power production from renewables has expanded at its fastest-ever rate in 2015 (IEA, 2016b). This has been possible thanks in part to favorable energy policies by governments. But also largely thanks to an important technological development and, consequently, a sharp reduction of costs in the processes of manufacture of capital goods. For instance the costs of production of solar technology have decreased 80% in the period 2008-2015. Therefore the technological development of renewable energies makes them more economically profitable, increasing their competitiveness against other more mature energy sources, such as fossil fuels. However, the IEA emphasizes that there are risks to consider.

Financing is a key parameter to provide a sustained investment in the renewable energy sector. Therefore, it is crucial to provide a safe and stable regulatory framework for the long-term and reduce policy uncertainties. Unreliable policies can compromise the pace of deployment of renewable energies by undermining investor confidence. In addition, IEA stresses that the problems associated to the integration of the variable renewables into the electric systems and the market become a critical priority for energy policy. In particular, IEA highlights as a strategic action to improve the advanced forecasting of renewable resources as part of the operating strategies.

Solar energy

In this scenario, solar energy will play a fundamental role. Within renewable energies it is probably the only one with sufficient potential to sustainably cover the planet's long-term energy expectations (Perez, 2008). The radiant energy reaching the Earth from the Sun is major driver of the climate system which powers the atmospheric circulation (Pozo-Vázquez *et al*., 2004). This primary energy source exceeds the current world energy requirement by about 1,500 times (Perez, 2008), being the most abundant energy resource on the planet. For this reason, an important part of the effort dedicated to the development of renewable energies focuses on the exploitation of this resource (Szuromi *et al*., 2007). Solar power generation has increased enormously in recent years thanks to improved energy conversion efficiency and lower production and installation costs. It is expected that renewable energy based on technologies that exploit the surface solar irradiance will be a central element in the future electricity generation. For instance this is the case in the EU. According to Eurostat the solar production in the EU remains relatively low, accounting for a 6.1 % share of total primary energy production of renewable energy in 2015. Nevertheless, the growth of solar energy has been the greatest of all renewables; its contribution to the total electric generation of renewables rose from 0.1% to 10.0%, in the period 2004-2014. Thus, for example, Europe has recently reached the value of 100GW of installed power of photovoltaic technology (PV). For its part, Spain is a benchmark in the use of solar energy, as it is one of the

pioneer leading countries in solar electric production worldwide (IEA, 2016a). In particular, Spain is the world leader in the development and operation of plants based on so-called concentrating solar power technology (CSP), with approximately 2300MWe in operation (Fernández-García *et al*., 2010), followed by the United States of America with 1700MWe. Regarding PV systems, the leading country is China, with an installed capacity of 78GWe. Spain is the tenth in the ranking with an installed capacity of 5.5GWe. To achieve the goal agreed by the EU countries on energy consumption, Spain has developed a new National Renewable Energy Plan for the next decade (2011-2020). This plan estimates a contribution of renewable energies to energy consumption in Spain of 22.7% by 2020, with a contribution to electricity production of 42.3%. Both percentages far exceed the targets set by the EU. An important part of the estimated increase is expected to be of solar origin.

Solar irradiance reaching the earth surface is highly variable in space and time, essentially due to geometrical factors of the Sun-Earth position and weather conditions. From the standpoint of the capability of certain technology of providing electricity on demand, solar energy is considered a variable renewable energy (VRE), like wind power, wave and tidal power and run-of-river hydropower (IEA, 2008). In contrast, there are renewable energies that can be classified as firm technologies, namely: reservoir hydropower, biomass, geothermal, and to a lesser degree some CSP technologies that have molten salt thermal storage. This natural feature of solar radiation, along with its relative early-stage of development, reduces its competitiveness against other sources of energy. As mentioned above the increasing technological development has notably reduced the costs of the solar technology. Nevertheless, the problems associated with the uncertainty of the primary energy source remain as a determinant factor in the final cost of solar projects. In this regard, a fundamental concept concerning solar projects is bankability. In short, a project is bankable if investors – public or private- consider that there are sufficient investment guarantees that they are willing to finance it. This means that a financeable project is likely to ensure financial success with the highest

degree of reliability possible. In a solar project one of the major issues to assure the success of the project is the capacity or level of production –there are others like the final energy price-, which is directly dependent on the solar resource. In this way, there are two aspects to consider concerning the solar resource: the assessment carried out during the early stages to evaluate the project feasibility and the forecasting during the operational phase.

Regarding solar resource assessment, the expected solar energy in the location of interest during the life time of the facility is carefully evaluated, because it is the highest source of uncertainty –this parameter is translated into risk and this into financial interest, expressed in a simplified form-. To this end, long-term time series of historical data are examined, and different scenarios of energy availability –based on statistical considerations- are contemplated. Nowadays this is made by means of the generation of artificial time series of solar resource data that are aimed to gather all the statistical information of the historical long-term time series into a single year period, called typical meteorological years (*vide infra* section 1.2 *State-of-the-art review*). Therefore, the importance of such data and the generated scenarios of being reliable is unquestionable.

On the other hand, meanwhile the costs of production are close to zero and very stable –surface solar irradiance is much more abundant than what is usable and freely available-, the variable character of solar energy significantly complicates its integration into the large scale power supply systems, increasing the associated costs. The origin of the problem lies in the fact that the electric power cannot be stored on a large scale. Hence, generation, transportation and consumption of electricity must be coordinated –by TSOs- and made at the same time. In the case of a VRE, like solar energy, the intermittence in the generation constitutes a notably challenge. The successful integration of solar energy must be approached in two ways. The first one concerns the electricity supply system itself, which must be adapted to gain flexibility in order to be able to respond reliably and rapidly to fluctuations in supply and demand (IEA, 2008). In addition they should

continue evolving from isolated grids to national and international markets building more sustainable and secure energy system (IEA, 2016c). The second way is to know in advance –with the maximum certainty possible- the solar energy production, so as to the TSOs can program and control the operation of the system safeguarding the maximum requirements of stability and security. To this end, a reliable forecast of the solar resource availability in short and medium range (hours to days) is the key, not only for the TSOs system operations, but also for solar energy producers, for operation and management of the solar facilities, as well as for energy trading.

Concerning solar energy conversion systems, today there are essentially two main technologies with different characteristics that operate integrated in the power supply systems (Quaschning, 2004): photovoltaic plants (PV) (IEA, 2014a) and concentrating solar power (CSP) plants -medium and high temperature- (IEA, 2014b). PV systems directly transform global solar irradiance into electricity through semiconductor devices -photovoltaic cells-. CSP plants transform the energy of the direct normal irradiance component (DNI) into heat through the use of devices that focus solar irradiance on receivers, thus initiating a thermodynamic cycle that transform heat into mechanical energy and then into electricity –solar thermal electricity (STE)-. The use of heat has an advantage from the standpoint of energy storage. Some CSP plants allow having a thermal energy storage system based on molten salts, which is capable to produce up to around 10 hours of generating capacity at full load. The technologies employed could be: parabolic trough collectors, central receiver, linear Fresnel reflector and dish Stirling. Today the largest CSP plant in operation is Ivanpah Solar Power Facility in USA, but installed powers of 50MW and 100MW are usual. On the other hand, the largest solar PV parks currently achieve nominal installed powers far over 500MW, being Kurnool Ultra Mega Solar Park in India currently the largest PV plant in the world, with 900MW in operation. Finally it should be mention a technology that has much lesser deployment but reaches the highest efficiency in PV energy transformation, the concentrator photovoltaics (CPV). This technology uses DNI as primary energy source but unlike CSP, it

focuses the sunlight into a PV cell. This technology seems to have a promising future. In addition to the high efficiency values it yields, the need of smaller PV arrays reduces the system costs, which improves its competitiveness. However, CPV still has to solve important technological challenges to become competitive against PV or CSP.

To summarize, the applied research about the solar irradiance as a primary energy source for solar energy applications is of maximum importance to meet the objective of increasing its competitiveness and favor further deployment. To this end, two aspects stand out: i) the reduction of the uncertainty in the solar resource assessment and ii) to improve the solar resource predictability in order to facilitate the integration into the large-scale power supply structures. Thus, in addition to the policies favorable to its development within a stable legal framework and the adaptation of the energy supply structures to assimilate the VRE, the study of the solar resource and its application to the solar industry are undoubtedly fundamental to favor the expansion of this energy source, since the exploitation of its enormous potential will be decisive in achieving the objectives of the Paris Agreement in the fight against climate change and the achievement of more sustainable development.

1.2 State-of-the-art review

The research in the field of solar radiation and its interaction with the components of climate system is under continuous development. Despite there are plenty of solid concepts, the development of the solar applications has brought new improvements, upgrades, supplementations and the creation of new ones, covering all the different aspects concerning the two dimensions of the solar resource from the standpoint of solar applications: assessment and forecasting. Thus, basic fundamentals of solar resource investigations cover different interrelated aspects, such as instruments, measurements, data quality control, modelling and post-processing. There are others more directly connected with the solar industry, such as the development of solar products to the end-users and their bankability. Following a

general description of the current state-of-the-art of the essential aspects of this field in continuous progress is presented.

1.2.1 General considerations

Solar radiation at the earth surface

Sunlight at the earth surface presents a variable spectral distribution of energy that covers wavelengths broadly ranging from 0.3 to 4 μm. This range contains most of the total electromagnetic energy from the Sun that reaches the earth surface, and it is customarily referred to shortwave solar radiation or, simply, solar radiation. Most usually, solar radiation is described in terms of its projection over the horizontal surface, which is the parameter known as global horizontal irradiance (GHI). It accounts for the total energy flux received per unit area. GHI is used by PV technologies, but projected over the surface of the panel; in that case it is usual to name it as global tilted irradiance (GTI) (Perez *et al*., 1990a; Gueymard, 2009a; Ineichen, 2011a). The beam irradiance is the solar radiation coming from the solar disc. In practice, however, its measurement also contains the solar radiation coming from a solid angle centered at the solar disc, called circumsolar radiation (Blanc *et al*., 2014). Beam irradiance is generally described over a surface normal to the direction of propagation, and thus named direct normal irradiance and denoted DNI. This is the primary energy source of interest for concentrating technologies, namely: CSP and CPV. A third component is the diffuse irradiance, which is the irradiance coming from the scattered sunlight by the atmosphere constituents. It is usually named in terms of its horizontal projection, that is, the diffuse horizontal irradiance. DNI along with the diffuse irradiance are related with GHI by the closure equation (Eq. 1.1), where Z is the solar zenith angle and *dhi* is the diffuse horizontal irradiance. Ultimately, it should be mentioned that spectral considerations, which are of significant importance for technologies such as CPV, are not included here as they are out of the scope of this research. Therefore, in the context of this work solar resource is defined in terms of the broadband spectral shortwave solar irradiance of interest for solar energy applications, namely: GHI and DNI.

$$GHI = DNI \cdot cosZ + dhi \quad (1.1)$$

Solar irradiance reaching the earth surface is highly variable in space and time. Its distribution is a function of several factors. Primarily, there are deterministic factors, which depend on the geometric features of the relative position of the Earth orbiting the Sun. The variations of these parameters modify the insolation received by the Earth at any time, being responsible of the day and night, seasons and other long and very long-term variations occurring along thousands of years –Earth gravitational interaction with other bodies of the solar system causes changes in its orbital eccentricity, obliquity, and precession-. Over geological short periods the irradiation received at the top of the atmosphere (TOA) –the so-called extraterrestrial irradiance- remains roughly constant, determined by the solar spectrum that changes with the Sun activity. It has a recently upgraded value of 1361.2Wm^{-2} -the integral of the solar irradiance over its whole spectrum-, referred to as the Total Solar Irradiance (TSI) or also formerly known as the solar constant (Gueymard, 2006; Gueymard, 2012b). Without considering the atmosphere, the distribution of this energy over lands and oceans takes place according to the latitude and the local topographic effects (Ruiz-Arias *et al*., 2010a). Nevertheless, the existence of the atmosphere additionally introduces a set of elements that interacts with the solar radiation like a kind of dynamical filter, which drastically alters the initial homogeneous radiation field. These interactions take place naturally by means of physical processes of scattering and absorption of the solar radiation by the atmospheric constituents, namely: gas molecules, aerosol particles and clouds droplets and particles. From the gases present in the atmosphere 98% corresponds to N_2 and O_2 and the remaining 2% is constituted by Ar, water vapor and trace gases: CO_2, CH_4, O_3, N_2O, CO and chlorofluorocarbon compounds (CFCs). The effect in solar radiation of this 2% of constituent gases is very important. Aerosol particles have a wide range of sizes -from less than 0.1 μm up to more than 20 μm- and shape distributions (Olmo, 2008). Clouds are mostly formed by water droplets and ice crystals. The natural fluctuations in the amount and

distribution of these elements in the atmosphere are responsible of the strong variability in space and time of the solar irradiance at the earth surface. In particular, among the extinction processes –absorption and scattering- produced by these atmospheric constituents there is no one most relevant for solar energy applications than that produced by clouds. This is due to the rapid spatial and temporal variability of clouds, which is not equaled by any other atmospheric specie (Mayer, 2009). Thus, clouds are the largest source of uncertainty for solar resource assessment and forecasting (Kim and Ramanathan, 2008; Ruiz-Arias *et al*., 2016). Then aerosols are the second largest one, being the first under cloudless sky conditions; in addition aerosols acts as condensation nuclei for cloud droplets, which can enhance the amount of cloud cover and also influence the lifetime of clouds. Among the solar radiation components, DNI is largely the most affected by interactions with clouds and aerosols, and therefore it is more difficult to estimate and typically presents greater uncertainty values than GHI (Bellouin *et al*., 2005; Ruiz-Arias *et al*., 2015a). Consequently, this affects more to the technologies based on concentrated solar energy, which uses DNI as primary energy source. In summary, the atmospheric constituents, particularly clouds, introduce an extremely high variability in the solar radiation field that finally reaches the earth surface, both in space and time, with variations that may already be significant for solar applications purposes in ranges of meters and minutes. The final contribution of all these deterministic and stochastic factors is generally the attenuation of the initially available TSI; although sometimes it may even be a temporary enhancement effect due to bright cloud reflections.

Solar radiation modelling

Solar radiation modelling is essential for solar energy applications. There is an ample variety of methods that sweep from simple statistical approaches to sophisticated physically founded techniques. The suitability of the method depends on the concrete application and the available information (Ruiz-Arias and Gueymard, 2015b). In this sense, for instance, the solar position is a major issue in solar radiation applications. For solar tracking systems the precise location of the Sun

throughout its deterministic daily path in the sky –defined by its azimuth and zenith angles- should always be known with high accuracy; in particular CPV technology requires an extremely high accuracy. To this end, different algorithms are currently available, which differ in the level of accuracy and complexity –usually translated into higher computational cost- (Michalsky 1988; Blanco-Muriel *et al*. 2001; Reda and Andreas 2004, 2008; Grena 2008; Blanc and Wald, 2012). From these, Reda and Andreas 2008 is the most accurate, but it is also the most computationally expensive. Other key geometric parameter for solar radiation calculations is the so-called air mass, which accounts for the path-length through the atmosphere that is traversed by the sunrays with respect to a vertical trajectory. For high Sun elevations, the approximation of a plane-parallel atmosphere could be enough. However, for large solar zenith angles the approximation is not appropriate and then empirical corrections should be applied to take into account the refraction of the sunrays (Kasten and Young, 1989; Gueymard, 2003).

Two of the most useful approaches in solar radiation modelling for solar applications are the clear-sky models and the all-sky separation models. The former group of models –also called clear-sky radiative transfer (CSRT) models- aims to faithfully represent the daily curve of the solar radiation components in absence of clouds at any location over the earth surface. To this end, several approaches are currently available. One of the main differences between them is the input parameters needed to compute the solar irradiance, which in all cases is a reduced number of total column quantities such as the precipitable water, aerosol optical depth (AOD) at certain wavelengths, Ångström exponent, surface pressure or the Linke turbidity coefficient (TL). Remarkable examples of these broadband models are: ESRA model (Rigollier *et al*., 2000), Simplified Version of SOLIS (Ineichen, 2008a) and REST2 (Gueymard, 2008a). Exhaustive benchmarking studies have been carried out to compare the estimations of different models respect to ground measurements (Ineichen, 2006). More recent studies have shown that REST2 obtains the best results, closely followed by the Simplified Version of SOLIS (Gueymard, 2012a; Badescu *et al*., 2012).

The results of these models are extremely accurate, even in the range of instrument uncertainties, but always as long as the input parameters are correct. More recently a new model has been implemented, called McClear model (Lefèvre *et al*., 2013). Unlike the others broadband models, this is based on the interpolation of the solar irradiance from pre-computed tables. On the other hand, because of its prominence it should be mentioned also an example of spectral model, called SMARTS (Gueymard, 2001, 2005). The application of these models in solar resource assessment and forecasting is fundamental for the estimation of GHI and DNI under clear-sky conditions, when the solar facilities can reach the higher values of production. They are also quite useful for data quality check, time reference verification, data interpolation, generation of smart-persistence models and even for preliminary check of instrumentation performance, among many others applications.

Regarding all-sky separation models, it should be said that currently they are of key importance for solar energy applications. These models separate GHI into its direct and diffuse components. This is quite useful to take advantage of the much more abundant GHI measurements. In addition, currently the most settled methods to estimate surface solar irradiance from satellite measurements are developed to obtain GHI. Therefore all-sky separation models play a fundamental role to provide DNI satellite-based estimations. In fact, these models actually are the methods employed by some of the most reputed providers of solar irradiance data in the solar industry. Notwithstanding, these models are not the unique solution for estimation of DNI from satellite measurements (*vide infra* Section 1.2.2 *Solar radiation assessment*). In addition, they are also used to provide DNI forecasts when the operated models do not provided this variable among its outputs. Moreover, because they are based on empirical approaches, separation models are relatively easy to implement and run. In summary, these models, despite they are not indispensable, continue to be a fundamental modelling approach for solar energy applications. Therefore, to evaluate the performance of these models is of major interest. An important work was made firstly by Ineichen (2008b). Nevertheless,

most recently Gueymard and Ruiz-Arias (2016) have presented an enormous effort of evaluation in a published work where they deeply analyze the performance of 140 separation models and which is continued in Aler *et al*. (2017). The results show that this kind of models is strongly dependent of the particular conditions of the location where the evaluation is carried out. This is due to the fact that each separation model is adjusted with a particular set of data, and therefore they are conditioned by the information enclosed in such datasets. Thus, it is quite difficult to advance which would be the expected performance of any model in any location. Anyhow, the study shows that some models stand out over others. Such is the case of the models of Perez (1992, 2002) and Engerer (2015). Other examples of different approaches are: Reindl *et al*., 1990, Elminir *et al*., 2007, Boland *et al*., 2008 and Ruiz-Arias *et al*., 2010b.

Solar irradiance evaluation

Finally, both for solar radiation assessment and forecasting it is essential to analyze and contrast the solar radiation modelled values against ground observations. In essence, the evaluation procedure consists of a comparison of the data being examined against a counterpart reference obtained by appropriate measurements, which are considered as the "truth" and are commonly called the real values or the observations (Oreskes, 1994, 1998). Any evaluation should be done under the necessary considerations that assure the certainty of the results. This essentially means that the reference measurements should have the properties of reliability, accuracy and stability. Otherwise they degrade the strength of the results and, consequently, the validity of the conclusions. For instance, high quality measurements of GHI can be obtained by means of a *pyranometer*. However, this kind of instruments cannot avoid an error associated with the device design itself, which appears at large solar zenith angles –instrumental cosine error-. Thus, if the evaluation is not limited to solar elevation angles above certain threshold –which depends on the instrument-, the uncertainty in the observed GHI is added to the evaluation, affecting the reliability of the results. In that case, it is much more advisable to limit the scope of the evaluation in exchange for obtaining more reliable results. Therefore, it

is easily understood that solar radiation measurements are of key importance in solar energy applications. On the one hand, during the planning phase, data of high quality are indispensable to support the project. In this sense, typically at least one year of local measurements is required by financing institutions in order to correct the necessary long-term time series, which are usually derived with satellite data and can present high uncertainty values (*vide infra* Section 1.2.2 *Solar resource assessment*). On the other hand, quality databases gain a much greater value over time. Thus, during the operational phase of the plant these data are of great importance to carry out a continuous control of the plant performance. Consequently, the selection of the proper instrument is fundamental, as it is the selection of the best location as well (Stoffel *et al*., 2010a; Vignola *et al*., 2012). The best choice of the instrument depends on the characteristic of the application for which it is to be used. It is also subordinated to the constraints imposed by the conditions of installation and the possibility of carrying out the mandatory activities of operation and maintenance. In this respect, there are several specialized publications of comparative studies that analyze the performance of different types of radiometers (Badosa *et al*., 2014; Gueymard and Myers, 2009b; Michalsky *et al*., 2005, 2011; Vuilleumier *et al*., 2014; Geuder *et al*., 2014). Ideally, the instruments should be reliable, durable, easy to maintain and not expensive. The operation and maintenance of the high-quality instrumentation for measuring solar irradiance it is a very demanding task if the intention is to keep all the potential quality of the instruments. This should include a program of periodical calibrations, which are essential to ensure the quality of the data over the time (Gueymard and Myers, 2008b). Moreover, it is a recommended practice, when possible, to have redundant measurements from independent instruments in the same station. This is quite useful for data quality check and to recover missing data. Ultimately, the data obtained following the guidelines to ensure the quality over the time are extraordinary valuable. Nevertheless, despite all the efforts, measuring solar irradiance is a complex and delicate task that is not exempt from possible errors -soiling, loss of tracking, etc.-, which affect the final quality of the dataset. In this regard, data quality check should be always a standard

procedure to carry out when working with solar irradiance data, not only to analyze extended periods of data but also for checking in near real time. This set of procedures aims to detect possible errors in the data, setting the corresponding flag associated to each data, which describes a level of suspicion from completely alright to certainly erroneous (Wilcox *et al*., 2012). Possibly, the most popular set of quality checks is that proposed by the Baseline Surface Radiation Network (BSRN) (Long and Shi, 2008; Roesch *et al*., 2011). Nonetheless, there are other possibilities that add more or less restrictive limits to validate the data (Ineichen, 2014; Gueymard and Ruiz-Arias, 2016). Among these automatic checks, a visual inspection should be always carried out. Anyhow, for data quality check a very good practice is to record the data with a sampling of very high frequency, like 1 to 5 minutes.

Once the observations are correctly characterized, the evaluation procedure can be performanced. To this end, a set of objective parameters are established to quantitatively assess the differences between the estimated data and the real measurements. Among the statistical parameters available, there are three that are of standard use, namely: mean bias error (MBE), mean absolute error (MAE) and root mean squared error (RMSE). All of them describe an important characteristic of the errors in a single averaged value. Thus, MBE accounts for the systematic error, allowing to know if the data overall overestimate or underestimate the solar irradiance. MAE always add the difference between the estimated data and the observed ones, being the most useful to account for expected deviations in solar irradiance forecasting and, hence, for the possible consequent penalties. RMSE, related to MBE through the standard deviation, accounts for the deviations, but penalizing more the larger ones. Therefore, these parameters only provide an averaged view of the errors, but allows for an easy assessment. In addition, they provided a simple way to establish intercomparisons between different data sources against the same observations in a common framework, which allow knowing the performance of such sources in a relative way. In addition to these popular parameters, there is a comprehensive set of methods to evaluate

solar irradiance, which is devised to analyze the information regarding the examined dataset. Recently Gueymard (2014) and Jensen *et al.* (2016) have carried out a complete revision of these methods. Finally, it should be mentioned that it is a common practice for evaluating solar irradiance forecast to use a simple reference model, such as the persistence model or the smart persistence model. Since this kind of model is the most simple to use, it is expected that a more sophisticated method beats it.

1.2.2 Solar resource assessment

Solar resource assessment involves a collection of methods to characterize the solar radiation normally available in a location of interest. In this sense, it is not a simple quantification of the available solar irradiance, but an exhaustive description that includes a detailed statistical analysis and a thorough study of associated uncertainty. In addition, it is an interdisciplinary field in which several disciplines should work together, as solar irradiance modelling, radiometry, metrology, meteorology, climatology, geography, engineering, remote sensing, statistics, and financing. In this sense, a reliable assessment of the solar resource along with a correct analysis of the uncertainty is fundamental to favor the bankability of the project. Furthermore, the reduction of the uncertainty gives the promoter confidence in the project, since it provides more reliable estimations of the annual production, which facilitates to seek investors and to face bidding processes. Nevertheless, solar resource assessment has had a slow development until recent years, when the impetus of the industry to meet its needs has led to faster and more consistent progress.

To characterize the solar resource with the lowest uncertainty, on-site long-term high-quality solar irradiance measurements are required. However, these measurements are never available in practice. Thus, satellite-based models (Perez *et al*., 2013a; Miller *et al*., 2013) are extensively used today since they can provide both full spatial coverage and long-term historical time series of data. Alternatively, it is possible to use more sophisticated physically-based methods which make use of

radiative transfer and atmospheric modeling (Jones and Fletcher, 2013). This is the case, for instance, of atmospheric reanalyses, such as MERRA2 of NASA, ERA-Interim of ECMWF or CFSR of NOAA. However, these models are generally less accurate that the satellite-based ones (Boilley and Wald, 2015). But, even the satellite-based models present most often higher uncertainty than direct observations, with a large range of variation depending on local features of the site of interest. Thus, to better characterize the solar resource, a variety of methods is used, by combining both onsite measurements and long-term satellite-derived modeled data. Several factors influence in the selection of the methods to apply, such as the purpose of the study, requirements and the availability of data, among others. Ruiz-Arias and Gueymard (2015b) propose an ensemble of best practices for solar resource assessment. These are: the selection of a proper location for the installation of the radiometric station, the collection of at least 12 months of data, acquisition of long-term historical data of solar irradiance from satellite, filling of gaps due to erroneous and missing data, assessment of measured data quality and uncertainty, evaluation of satellite against ground observations, long-term bias correction of satellite data and derivation of useful information for solar resource characterization, as the statistical parameters of the long-term dataset, TMY and the total uncertainty.

Satellite-based estimations

To cover long-term periods (over 15 years) and almost any part of the globe, satellites are the best option to estimate the surface solar irradiance. They transport radiometric instruments that sense the multispectral radiation reflected and emitted by the Earth, both from the surface and the atmosphere, with high spatial and temporal resolutions. To derive the solar irradiance received by the earth surface from the reflectance measured by the satellite sensor, several approaches have been proposed along the years. Today, the most common methods can be classified according to their nature. Thus, there are physically founded methods and semi-empirical based methods –also known as cloud-index methods- (Ruiz-Arias and Gueymard, 2015b). The first type of methods has probably more potential for improvement in the

forthcoming years, although it is more complex and computational demanding. In addition, like CSRT models, they require precise information of the actual atmospheric state, which is not always available with the required accuracy (Habte *et al*., 2013; Miller *et al*., 2013). On the other hand, semi-empirical methods still are the most widely used for solar energy applications (Perez *et al*., 2002, 2013a; Polo *et al*., 2008; Rigollier *et al*., 2004; Lefèvre *et al*., 2007). In these methods, GHI is estimated superimposing the cloud optical amount derived from the satellite measured radiance on the clear-sky GHI obtained from a CSRT model. DNI is then derived by means of all-sky separation model. Usually these methods are additionally adapted by researchers and solar resource providers according to their own criterion in order to reduce the uncertainty of the estimations. Ineichen (2014) presented a benchmarking study of seven semi-empirical approaches from different databases. Anyhow, several sources of uncertainty still remain (Cebecauer *et al*., 2011). Most recently other approaches have been proposed, which are based on artificial intelligence (AI) techniques, such as artificial neural networks (ANN) (Quesada-Ruiz *et al*. 2015; Linares-Rodríguez *et al*., 2015). Unlike the semi-empirical methods, AI-based methods can provide directly both GHI and DNI, and are constrained by the amount of information available for carrying out the model training.

Site adaptation

Also referred to as dataset merging or measured record extension, the term site adaptation designates an ensemble of methods applied to correct the systematic error in the long-term dataset used to characterize the solar resource in a location of interest by means of short-term local ground measurements (Suri and Cebecauer, 2011; Bender *et al*., 2011; Thuman *et al*., 2012). The aim is to reduce the original uncertainty of the satellite estimation. To this end, the different methods calibrate the long-term satellite data against local observations during the overlapping period. This strategy reduces the random errors and, overall, the bias or systematic errors. Today this is a standard requirement for bankability for large solar projects.

Nowadays, there are several approaches to carry out site adaptation. They all are based on the use of short term solar irradiance ground observations datasets –generally at least one year is required-. These methods can be statistically or physically founded. Polo *et al*. (2016) present a review of the state-of-the-art of site adaptation methods.

Typical meteorological years

Typical meteorological years (TMYs) are artificial annual time series -normally of hourly or sub-hourly values- of useful variables for solar energy systems, which aims to condensate all the long-term information into a single year. Thus, it has the natural diurnal and seasonal variations and theoretically represents a year of typical climatic conditions in the location of interest. In the context of solar resource applications, the aim of TMY is to preserve the statistical features that characterized the solar resource in the location of interest, in order to describe the expected behavior of the solar system, ideally during the lifetime of the facility (Stoffel *et al*., 2010a; Sengupta *et al*., 2015). In this sense, TMY representativeness directly benefits from the use of improved long-term data. Originally it was introduced in 1978 by the Sandia National Lab and later by NREL, as a solution to the limitations of the computational capabilities that made difficult solar energy simulations. Since then it has become a tool of common use in the solar energy industry. It is necessary for complementing the bankability analysis of solar energy projects, such as CSP and PV. However, nowadays its use for plant design is not recommended.

From a different perspective, TMYs are fundamental in comprehensive solar resource assessment studies. The theoretical features of this synthetic year facilitate to obtain the objective quantitative estimation of the expected solar energy production in the location of interest. In this sense, this is a key property of TMYs, since they allow specialists to estimate the values of the expected annual solar irradiation at different probability scenarios of available solar energy amount. Such scenarios are usually measured in terms of the probability of exceedance -habitually also referred to by the name of its complementary concept, the percentile-, and the corresponding

estimated uncertainty, which includes the uncertainty associated to the estimation of the long-term data. These valuable properties are of major interest for the solar energy industry, because they are key information to assess the economic risk of the solar project, which finally is translated into financial interest. Thus, TMYs of several scenarios are demanded to carry out the feasibility analysis, usually the average -probability of exceedance of 50%- and low-energy or pessimistic cases -probability of exceedance of 75%, 90%, 95% or even 99% (Cebecauer and Suri, 2015).

TMY does not correspond to any particular year of a certain period, but it is a calendar year artificially built as a statistical-based weighted composition of twelve months selected from the historical long-term time series of the location of interest. Customarily, weights are used to take into account not only the main variable of interest for generating the TMY -GHI or DNI, depending on the application-, but also ancillary meteorological variables, such as temperature, wind speed, relative humidity, etc., which are of interest for production analysis. Nonetheless, the experts do not agree in the appropriateness of taking into account the meteorological variables and how to do it in order to adjust the composite of weights, since there are not conclusive evidences of such convenience. Thus, Habte *et al*. (2014) from NREL have proposed a new configuration of weights in which the main variables for PV or CSP applications, that is, GHI and DNI respectively, have the 100% of the weight. They name the resulting TMY as typical global (horizontal irradiance) year (TGY) and typical direct (normal irradiance) year (TDY). In the context of this research work, to refer to both at the same time the term typical solar year (TSY) is used. It should be noted that TMY may not have necessarily more information than TSYs, but slightly different, since the last can be completed simply adding the corresponding meteorological information. The difference is that TMY takes into account –in a non-consensual way- the meteorological variables to generate the year, while TSYs only consider the solar irradiance component of interest to form it.

Consequently, unfortunately there is no scientific consensus on a standard method to generate TMYs or TSYs. On the contrary, there is a varied set of methods and strategies followed by different authors and providers to generate their own TMYs. Thus, both methods and weights may result in different TMYs for the same location and long-term dataset, which is an unwanted situation. In this context, some research efforts analyze the performance of different TMY approaches in solar energy applications (Ineichen, 2011b; Realpe *et al*., 2016).

Uncertainty

Uncertainty is probably the most important individual concept in solar resource assessment. It is essential to evaluate the quality of the assessments and key for the bankability of solar projects. Uncertainty is, thus, indispensable for any rigorous analysis in solar energy applications in order to obtain comprehensive conclusions. The uncertainty in the solar resource is directly related to the uncertainty in the expected performance of the solar plant (Sengupta *et al*., 2015). Furthermore, the uncertainty in the estimation of the solar resource is the largest source of the overall energy production uncertainty.

Nevertheless, it is a parameter that sometimes is not completely clear in the solar energy industry. It should not be confused with the variability of the data. Actually the total uncertainty is a composition of individual contributions of all the elements involved in the characterization of the solar irradiance: measurements, modelling, interannual variability, spatial variability, representativeness of the period and site adaptation approach, (Meyer *et al*., 2009). In addition to all these, for TMY it should be also take into account the uncertainty derived from the method itself (Fernandez-Peruchena *et al*., 2016). The characteristics of each component of the uncertainty depend on the specific aspect of the solar resource assessment. To add all these contributions (u_i) to obtain the total uncertainty (U) it is common to use the Gaussian law of error propagation (equation 1.2).

$$U = \sqrt{\sum (u_i)^2} \tag{1.2}$$

1.2.3 Solar radiation forecasting

Solar radiation forecasting for solar energy applications is a relative recent topic. The progress since the early studies to the state-of-the -art of the present day has been considerable, currently covering both components -GHI and DNI- and the different spatial and temporal dimensions of the matter. Anyhow, it is an ongoing research of capital importance for the cost-efficient integration at large-scale of the solar energy into the power supply systems. Today, reliable predictions of solar power production are a basic tool for grid management in countries with substantial level of penetration, like it is the case of Spain. Spanish TSO currently uses power forecasts for both PV and CSP technologies in order to enhance the control of the system in terms of the grid stability. It also uses it to carry out the management of the complementarity with other variable energy sources. In general, TSOs are responsible for the secure balancing of the grid. In addition, during the operational phase of the solar plants, power forecasts are of key importance for efficient operation and management of the facilities, as well as for energy trading. Since the uncertainty of the solar power output is primarily due to the uncertainty in the foresight of the incoming solar irradiance, the importance of a reliable forecast of solar resource is unquestionable. In this sense, several studies have evaluated the worthiness of operational forecasting of solar irradiance for energy applications (Dumortier *et al*. 2009; Kleissl, 2013). Some of this research works focus on a particular technology, like the studies of Perez *et al*. (2007) and Marcos *et al*. (2013) for PV, and those of Wittmann *et al*. (2008), Kraas *et al*. (2013) and Law *et al*. (2016) for CSP.

A convenient way to basically describe solar radiation forecasting is taking into account its time dimension. In this regard, three main times should be considered with respect to the prediction, namely: forecast horizon, forecast time resolution and forecast update frequency. There is other time that should be taken into account in practical situations, which is the forecast delivery time. In operational use and for energy trading, the time at which the forecast should be available is an

important constraint to consider. Regarding forecast horizon, which is related to the application requirements in which the forecast is to be used, it decisively influences the most appropriate methodologies to be applied. In this way, along the time dimension, the forecast can be divided into three main parts, with fuzzy limits. The first 10 to 20 minutes, up to 4 to 6 hours, and up to several days. A review of the current state-of-the-art of forecasting methods can be found in Inman *et al*. (2013), Diagne *et al*. (2013) and Kleissl (2013). Antonanzas *et al*. (2016) carried out a review of forecasting methods for PV production.

Intra-hourly forecast

The intra-hourly forecast is used for solar plant operation. Usually this forecast is made by means of all-sky cameras and ground measurements. The forecast horizon actually depends on the type of clouds, velocity of displacement and the cloud-base height. These forecasts normally achieve very high spatial and temporal resolutions, of meters and minutes. Also the update is very high, in the order of 1 minute. This method is based on the assumption of predominance of horizontal advection of clouds. That is, during a very short-term period of minutes the cloud shape and its optical properties do not change significantly, while horizontal velocity, detected by intercomparisons of two consecutive sky images, is assumed constant and hence cloud motion is extrapolated ahead. Thus, the future position of clouds is estimated (Urquhart *et al*., 2013, 2015; Quesada-Ruiz *et al*., 2014). This method, based in determining the so-called cloud motion vectors (CMV), is also the basement of the next step in forecast time horizon.

Nowcasting

Short-term forecasting with time horizons up to several hours –also known as *nowcasting* to differentiate it from the third type of forecast horizon-, is also mainly used for plant operation. As it is based on geostationary satellite images, the nominal time resolution is that of the satellite, with updates every 15 to 30 minutes. 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 the NWP approach. This method also uses the same strategy of detecting

the cloud motion vectors (CMV), but in this case from satellite images. This difference is essential, logically, changing the features of the method, such as the spatial coverage, the resolutions and the time range. Cloud properties and irradiance estimation are obtained as it has been exposed before, usually by means of semi-empirical methods (*vide supra* Sub-section 1.2.2 *Solar radiation assessment*) (Hammer *et al*. 2003; Kühnert *et al*. 2013; Perez and Hoff, 2013b). A combined method of satellite and NWP is proposed by Müller and Remund (2013). They use the cloud shape and position information from the satellite and use the wind information provided by the NWP model. A similar approach is proposed by Miller *et al*. (2013), but in this case they use a physical method to retrieve the cloud properties and solar irradiance. Finally, recent approaches take advantages of available improved cloud information that can be advected and diffused by a NWP. This hybrid method presents better results than the traditional ones based on CMV (Arbizu-Barrena *et al*., 2017).

Forecasting

Short and medium-range forecasts –simply referred to as *forecasting*- are usually carried out by means of NWP models. Comparative exercises have shown that this technique is the most effective for forecast horizons beyond 4 to 6 hours, when the correlation of the actual state of the atmosphere and the future state is low. Then, a physically founded method such NWP model performs better.

Arguably, NWP models are singly the most powerful tool for solar irradiance forecasting, because they are built upon physical principles and are able to resolve the complete state of the atmosphere at each time step in the future. They are a very complex instrument, whose constraints are continuously transcended thanks to the research efforts for improving the physical modelling and an increasing computational power. Nowadays, their potential seems to be unrivaled. For solar energy applications, not only they can provide forecasts of the main solar irradiance components –GHI directly and DNI may be derived-, but also for a comprehensive ensemble of meteorological variables of

interest for solar energy production, such as temperature, relative humidity or wind speed and direction. Moreover, as they cover the largest time horizons, they are the suitable tool for plant management and energy trading.

Basically, NWP models are a software implementation of the physical processes that take place in the atmosphere and the rest of elements of the climate system, as well as the interaction between them. For instance, this includes the formation, advection, diffusion and dissipation of clouds and the transfer of solar irradiance from the top of the atmosphere to the earth surface. The differential equations of NWP models are able to resolve the state of the atmosphere for determined spatial –or grid- and temporal resolutions. However, there are many key processes that occur at subgrid scales. For instance, some cloud structures can easily have lower sizes than typical grid resolutions. Then it is said that they cannot be seen by the model, or in other words, they cannot be resolved. Thus, it is expected that greater resolutions reduce the error in the representation of modelled processes. Nevertheless, there are still unknown details about the performance of the model respect to the spatial resolution for certain processes. Anyhow, to take into account for subgrid phenomena, models are constitutively complemented with the parameterizations schemes (Stensrud, 2007). These are software implementations that describe such physical phenomena that occur at higher resolutions than the nominal one of the model. For instance, solar radiation in NWP models is an example of parameterized process.

NWP models require knowing the initial state of the atmosphere and the boundary conditions in order to integrate the differential equations and thus to cast the atmosphere forward in the time dimension. This information is based on worldwide observations that are routinely gathered and processed to produce what is called the analysis. Today, this process typically takes between 6 to 12 hours to be completed. With this initial state of the atmosphere the great centers of global prediction run their own models –called global circulation models (GCM)- to generate the world weather forecast. Some examples of

GCMs are: the Global Forecast System (GFS) produced by the National Centers for Environmental Prediction (NCEP, USA), the Integrated Forecast System of the European Centre of Medium-Range Weather Forecast (ECMWF, UE), the Global Deterministic Prediction System (GDPS) of Canadian Meteorological Centre (CMC) and the Goddard Earth Observing System Model Version 5 (GEOS-5) developed jointly by NOAA, NCEP and EMC (USA). The temporal resolution at which their outputs are currently disseminated ranges from 1 to 3 hours. Typical spatial resolutions are 25 km, except for IFS that is approximately 12 km. These models produce reliable forecast of solar irradiance, with low biases. However the nominal time resolution, between 1 to 3 hours, is low for solar energy applications. Usually 1 hour is required for integration into the power supply systems, but intra-hourly resolutions are preferred by producers to better estimate the production, particularly in CSP plants. Then temporal interpolation is required when using GCM forecast for solar energy applications (Lorenz *et al*., 2009a). In addition, usually this kind of models does not provide DNI as output. Therefore, it should be derived by means of external model, for example all-sky separation models. This could be one of the reasons underneath the fact that this component has been less studied than GHI. Other important reason is that it is much more difficult to deal with. Moreover, PV technology is much more extended than the concentrating-based ones. In addition, reliable DNI ground measurments are more difficult to obtain and are scarcer.

Using the analysis and the global forecasts provided by the GCMs as initial and boundary conditions, regional NWP models can be run to provide forecast of a complete set of meteorological variables at very high spatial and temporal resolutions, typically in the order of few kilometers and few minutes. Among all the regional models, the WRF model (Skamarock *et al*., 2008) is one of the most advanced and widely used worldwide. It is supported in a collaborative effort by NCAR and other institutions of the USA. It is also developed thanks to the contributions of the scientific community throughout the world. WRF has an extensive set of parameterizations that give the model with great

flexibility and adaptability to the specific meteorological and geophysical conditions of a particular region.

Nevertheless, NWP models are intended to produce weather forecast. Therefore, they are not devised specifically for solar energy applications. This is evident by the limitations exposed above regarding DNI. But there are also other aspects concerning solar irradiance that are not fully considered in NWPs. This is mainly due to two main circumstances. On the one hand, the lack of information of determined parameters, which are essential for solar radiation interactions, like aerosols. On the other hand, because of solar radiation in the model is very expensive in terms of computation and hence the radiative calculations tends to be simplified as much as possible. Nonetheless, the increasing importance of the solar resource has stimulated initiatives to progress in this field. In this way, the WRF model has been recently evolved with a set of updates directed to improve solar forecasting for energy applications. All together, these updates make up a particular configuration of the WRF model known as WRF-Solar (Jimenez *et al*., 2016). Specifically, WRF-Solar is designed to improve the representation of clouds (Thompson and Eidhammer, 2014; Deng *et al*., 2014) and aerosols (Ruiz-Arias *et al*., 2014), as well as the interactions of these key elements with solar radiation. The authors have evaluated the model under clear-sky conditions and they have obtained promising results, significantly improving those obtained for the three solar radiation components estimated with the standard version of WRF. They also demonstrate that it is necessary to incorporate the influence of aerosols to obtain accurate estimations of solar irradiance. This conclusion has been also found by Ruiz-Arias *et al*. (2013).

Currently, numerous studies have been carried out to evaluate the performance of several models under different conditions and regions. Some examples of these research works are: Gaston *et al*. (2009), Lorenz *et al*. (2011, 2012), Mathiesen and Kleissl (2011), Pelland *et al*. (2013), Isvoranu and Badescu (2015), Aryaputera *et al*. (2015),

Zempila *et al*. (2016), He *et al*. (2016), Lima *et al*. (2016) and Sosa-Tinoco *et al*. (2016).

There are also other possible methods for solar radiation forecasting. They are based essentially on statistical approaches. These methods are founded in the assumption that patterns in the historical data sets are repeated in the future. Therefore, they require historical time series of data to analyze and recover such patterns. Several techniques can be applied in order to take advantage of this persistence, like for example the autoregressive integrated moving average (ARIMA) models. More trendy methods are based on artificial intelligence techniques, such as artificial neural networks, k-nearest neighbors or support vector machines. A description of several statistical methods used for solar irradiance forecasting is presented by Coimbra and Pedro (2013) and Diagné *et al*. (2013). The combination of the statistical and physical forecasting methods is of special interest, since it provides powerful tools to be applied to the raw output of the physical models such as NWP. These methods allow reducing the possible systematic errors of the physical model. They can take into account local effects and also include extra information about external parameters. Additionally, they can be used to derive information that is not included in the original output of the model. The ensemble of techniques that combines both statistical methods and physical models is usually called post-processing techniques or model output statistics (MOS). These can include simple techniques such as smoothing filters by means of spatial averaging (Lorenz *et al*., 2009a).

Finally, it is important to mention the convenience of carrying out intercomparisons between the forecasts from different sources. This is tailored to establish a general reference frame where different methods can be evaluated against others. In this sense, benchmarking exercises are a convenience practice. Larson (2013) and Lorenz *et al*. (2015) have carried out a study of intercomparisons between several NWP models.

1.3 Motivations

Solar renewable energy is increasing its presence in the large scale electric power supply. However, nowadays its weighted importance within the energy mix is still far from its theoretical potential. Notwithstanding, as mentioned before, its expectations of growth for the next decades are the highest of all the renewable energy sources. This prospective growth will be driven by technological and economic dynamics, which are unfailingly interconnected. The rate of penetration will be associated with the capability of providing solutions that can align the technical knowledge advances with the economic benefits, beyond the initiatives of the different governments and public institutions. In this way, today many of the efforts in the field of solar energy are focused on reducing costs while protecting the security of the electricity supply. The present research work focuses on those concerning the solar resource.

Within the topic of solar energy, in addition to the technology of exploitation itself, there is an eminently essential aspect, namely the solar resource. This primary source of energy -the solar irradiance reaching the earth surface- and the meteorological factors that affect it, naturally determine in a direct way the most important aspects associated with solar energy, which are the development and integration of this renewable source in the power supply structures, both at generation and distribution levels. In addition, the solar resource also conditions the technology of exploitation, being determinant for all of them. Over the last recent years there has been an increase in the interest and, in parallel, an increase in the research associated with this matter in its two main dimensions: the assessment and the forecast of solar resource. Most of this research follows a distinctly practical pathway, providing solutions and answers to each particular problem posed by the solar industry. However, there are many questions that are far from being solved, despite the fact that several alternatives have been proposed. At the same time, there are solutions that need to be developed in greater depth so that they can reach the potential that is expected from them. In both circumstances, the elaboration of rigorous

works that are scientifically validated is of great importance, since they contribute to the development and integration of the solar renewable energy in a real way -in a practical sense– when assimilated and applied by the solar industry. In this respect, it is the industry itself which demands this type of high valuable works that help the growth of the sector. The elaboration of the research study presented in this thesis decidedly follows this orientation, contributing with new solutions and deepening in others, both in the scopes of solar resource assessment and solar resource forecast.

Regarding solar irradiance assessment, one of the major current issues for the industry concerns the need to have the most reliable information possible about the solar resource, in order to carry out the decision making in solar projects. In particular, promoters are required by financial institutions to conduct objective feasibility analysis of their projects, with high dependability of the expected performance of the solar energy based facilities during their life-time. These performance analyses are carried out based fundamentally upon time series of solar irradiance data. These time series should be able to objectively characterize the long term performance of the solar plants. These performance studies are made in terms of the estimated production capability of the facility, according to its expected behavior under certain stress scenarios that usually consider annual periods of low solar energy availability. The definition of such scenarios, together with the uncertainty parameter derived from the solar energy assessment, are of key importance, because they determines essential aspects of the projects feasibility, such as the return of investment and financial costs. Therefore the quality of the information contained in the datasets used and its later processing to obtain such valuable information are decisive. It is precisely the definition of such scenarios that represents one of the current difficulties in the field of solar resource assessment. As explained in the previous section (*vide supra* Section 1.2 *State-of-the-art*), these scenarios are established based on historical long-term time series of surface solar irradiance –and other meteorological variables of interest- and described in terms of the so-called typical meteorological year (TMY) –or the most recent concept of typical solar

year (TSY)-. The problem is founded on the fact that there is not a common unified method for generating TMYs that is widely accepted by the scientific community. Further more, the existence of different methodologies for TMY generation evidences the lack of scientific consensus about this question. Moreover, the current methods are scarce. From them not all allow for generating TMY for any probability scenario of solar resource availability (Cebecauer and Suri, 2015). In addition some of them are under discussion and some problems have been detected (Blanc, 2015). Finally not all of them are public, but they are part of the private knowledge of certain companies that provide this kind of consulting services. Nevertheless, from the standpoint of the final users of the solar industry the possibility of obtaining two different results for the same data source introduces an extra factor of uncertainty, which remains uncontrolled. Althougt TMY is a tool that does not account with the favor of the scientific experts, it is so used in the solar industry that it remains to be a key tool for solar projects. Hence, due to its practical and extensive use, an endeavor for a unified scientific answer is therefore ineluctable. The aim is to develop a standard method for TMY generation that can generate confidence by having a broad scientific consensus and contributes to enhance the bankability of the projects.

On the other hand, to integrate a variable source like solar energy into the energy supply structure on a large scale compromises the operability of the system, which must work under the principle of security and stability, while the production should be balanced with the expected demand. In addition, producers need to know the expected resource that will be available in order to program the plant operation and management, as well as to plan the better strategy for participating in the electricity market. In this sense, it is usual to apply a policy of penalties against the deviations announced in advance by the producers to the TSOs. These penalties reduce the potential revenues of the plant respect to the case of 100% of forecast reliability, that is, without any deviation. Furthermore, to know the solar irradiance with certain foresight would expand the possibilities for improving the profitability of the plant, for instance by means of its participation in special

markets that are more rewarded, like the so-called adjustment markets. Hence, solar radiation forecasting will be a key tool for effective scheduling, improving decisively the competitiveness of solar energy. Thus, it can be concluded that the final problem is not so much the variability of the solar radiation, but its ***predictability*** (IEA, 2008). Therefore, solar industry demands a specialized reliable forecast of solar resource for all useful periods in plant operational phase. Among the methodologies for solar radiation forecasting, NWP models stand out as the most powerful tool. They provide comprehensive forecasts on a physical base for short and medium ranges –for minutes up to days ahead- for both single locations and extended regions, with high resolution in space and time. In particular, the WRF regional model is one of the most advanced NWP currently in the world. It is widely supported and developed by the scientific community, being in the state-of-the-art of numerical weather modeling. Nevertheless, NWP models have not been specially designed for the application of solar irradiance forecasting to the solar energy sector. Even more, the knowledge about the skillfulness of the model to predict solar irradiance is far from being widely studied. Main issues remain unknown. It is necessary to understand the specific behavior of the model when forecasting solar irradiance. In particular, the forecast of DNI component remains little investigated. Thus, it is basic to know the model performance under different cloudiness situations, the stability of the predictions respect to the forecast horizon and how is the skill of the model predicting GHI and DNI. In addition, it is very important to benchmark the forecasts against other models, so as to establish a reference framework that allows understanding not only the performance of the model in absolute terms, but also in relative ones. Finally, it is essential to know if higher resolutions, which a priori mean a better representation of the physical phenomena, provide better results when evaluating solar radiation forecasts (Stensrud, 2007). This provides a better insight about the model performance to forecast solar irradiance. All in all, this research provides a better understanding of the capabilities of WRF model providing solar irradiance forecasts and its applicability to the solar industry regarding the integration of this source of energy into the power supply structures.

To summarize in a general way, the research work of this thesis is motivated for the need of answering several relevant questions that remain unknown or incomplete, but that are of key importance for the application of the state-of-the-art knowledge to the solar energy sector. In particular, the interest lies on enhancing the competitiveness of the solar energy through the improvement of the solar resource assessment and forecast in the specific aspects described above.

1.4 Objectives

The aim of this PhD dissertation is to develop and to evaluate methods for the characterization and estimation of the surface solar irradiance for its practical application in the field of solar energy. The scope of this goal addresses two dimensions of solar radiation as a primary source of energy: its assessment and its forecast. In this thesis, we deal with the two main solar irradiance components for solar energy conversion technologies, namely, global horizontal irradiance (GHI) and direct normal irradiance (DNI) (the latter much less investigated than the former in the literature so far).

The specific objectives covered in this thesis are itemized below:

1. To develop and to evaluate a method for the generation of representative years of solar irradiance for the characterization of the solar resource at multi-year scales in any location of interest.
2. To evaluate the reliability of solar irradiance forecasts provided by the regional Weather Research and Forecasting (WRF) mesoscale numerical weather prediction (NWP) model.
3. To benchmark the solar irradiance forecasts provided by the WRF model against other reference NWP models.
4. To analyze the effect of the spatial horizontal resolution of the WRF model in the reliability of the solar irradiance forecasts.
5. To study the effect of the spatial aggregation of the predicted solar irradiances in the forecast reliability.

The first objective has to do with the assessment of the solar resource. It deals with the requirement of the solar industry of a scientifically-sound method that becomes a standard for the generation of typical solar years (TSYs) particularly focused in the description of the solar resource. TSYs are conventionally used for analyzing key aspects of the financial feasibility of solar projects –usually within the realm of *bankability* studies-. The attainment of this objective may lead in a common tool widely used for the computation of TSYs. This would also allow unifying the part of the total uncertainty associated with the method itself in the projects of solar resource assessment. In addition, its open character and its attribute of being easily implemented in an algorithm could favor its adoption by the solar community.

The remaining objectives concern forecasting of solar radiation, which is essential for the integration of the solar energy in the power supply systems. In particular, the second objective analyzes the use of the WRF NWP model for its application in the prediction of the surface solar irradiance, in order to ascertain its suitability as a supporting tool for solar power integration. WRF is a state-of-the-art community mesoscale model, publicly available and with a broad community of users worldwide. The third objective deals with the need of evaluating the reliability of WRF against other reference NWP models. This allows to determine a scale of reference in a comparative framework and thus to know the model skill in relative terms. The fourth and fifth objectives intend to investigate thoroughly the influence of the spatial resolution in the WRF model when estimating surface solar irradiance. This is directly connected with the conclusions of the results obtained in the second and third objectives. It is important because, precisely, the capability of regional models like WRF to achieve high spatial resolutions, a priori, is assumed to be an advantage, since this quality is considered to provide a better representation of the physics phenomena. These objectives should conduct to a better comprehension of using the WRF model as a powerful tool for solar irradiance forecasting, allowing the competitiveness of the solar energy to enhance. Moreover, they should serve as a supportive work for further research in this area.

To summarize, these objectives outline a thesis whose general purpose has a strong influence by industry-related concerns. To this end, the research work attempts to align scientific advances with particular interests of the solar industry from the standpoint of the primary energy source. Thus, this thesis shall contribute to enhance the level of progress and integration of the current exploitation technologies of the solar renewable energy.

1.5 Thesis structure

This thesis is organized as follows. Chapter 1 introduces the background and the framework which the research is carried out in. Following, the core of the research work is presented. It is composed by a corpus of four published works –chapters 2 to 5-. The organization of the chapters does not follow the chronological order of the research, instead it has been structured pursuing the accustomed preferences in the solar community, which follow the normal chronology of development of solar projects. In this way, chapter 2, which concerns solar resource assessment, is placed before the chapters dedicated to the solar resource forecast –chapters 3 to 5-. A summary description of these chapters is presented below. Finally, the thesis ends with a discussion of the main conclusions derived from the research work, along with an explanation of the future research proposed to continue the research lines here initiated.

Chapter 2 - *A novel procedure for generating solar irradiance TSYs*

Typical Solar Years (TSYs) are key parameters for the solar energy industry. In particular, TSYs are mainly used for the design and bankability analysis of solar projects. In essence, a TSY intends to describe the expected long-term behavior of the solar resource (direct and/or global irradiance) into a condensed period of one year at the specific location of interest. A TSY differs from a conventional Typical Meteorological Year (TMY) by its absence of meteorological variables other than solar radiation. Concerning the probability of exceedance (Pe) needed for bankability, various scenarios are commonly used, with Pe90, Pe95 or even Pe99 being most usually required as unfavorable

scenarios, along with the most widely used median scenario (Pe50). There is no consensus in the scientific community regarding the methodology for generating TSYs for any Pe scenario. Furthermore, the application of two different construction methods to the same original dataset could produce differing TSYs. Within this framework, a group of experts has been established by the Spanish Association for Standardization and Certification (AENOR) in order to propose a method that can be standardized. The method developed by this working group, referred to as the EVA method, is presented in this contribution. Its evaluation shows that it provides reasonable results for the two main irradiance components (direct and global), with low errors in the annual estimations for any given Pe. The EVA method also preserves the long-term statistics when the computed TSYs for a specific Pe are expanded from the monthly basis used in the generation of the TSY to higher time resolutions, such as 1 hour, which are necessary for the precise energy simulation of solar systems.

Chapter 3 - *Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain)*

In this work, we evaluate the reliability of three-days-ahead global horizontal irradiance (GHI) and direct normal irradiance (DNI) forecasts provided by the WRF mesoscale atmospheric model for Andalusia (southern Spain). GHI forecasts were produced directly by the model, while DNI forecasts were obtained based on a physical post-processing procedure using the WRF outputs and satellite retrievals. Hourly time resolution and 3 km spatial resolution estimates were tested against ground measurements collected at four radiometric stations along the years 2007 and 2008. The evaluation was carried out independently for different forecast horizons (1, 2 and 3 days ahead), the different seasons of the year and three different sky conditions: clear, cloudy and overcast. Results showed that the WRF model presents considerable skill in forecasting both GHI and DNI, overall, better than a trivial persistence model. Nevertheless, both MBE and RMSE values presented a marked dependence on the sky conditions and season of the year. Particularly, for 24 h lead time, the MBE of the forecasted GHI was 2% for clear-skies and 18% for cloudy conditions.

However, the MBE of the forecasted DNI increased up to about 10% and 75% for clear and cloudy conditions, respectively. Regarding RMSE values, in the case of forecasted GHI, results ranged from below 10% under clear-skies to 50% for cloudy conditions. In the case of forecasted DNI, RMSE ranged from 20% to 100% for clear and cloudy skies, respectively. This proved the higher sensitivity of DNI to the sky conditions. In general, an increment of the MBE and RMSE values with the cloudiness was observed. This reflects a still limited ability of the WRF model to properly forecast cloudy conditions compared to clear skies. Nevertheless, the model was able to accurately forecast steep changes in the sky (cloudiness) conditions. Finally, WRF performed considerable better than the persistence model for clear skies both for GHI and DNI, with relative RMSE values about a half. However, for cloudy conditions, performance was similar.

Chapter 4 - *Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe*

This article combines and discusses three independent validations of global horizontal irradiance (GHI) multi-day forecast models that were conducted in the US, Canada and Europe. All forecast models are based directly or indirectly on numerical weather prediction (NWP). Two models are common to the three validation efforts – the ECMWF global model and the GFS-driven WRF mesoscale model – and allow general observations: (1) the GFS-based WRF- model forecasts do not perform as well as global forecast-based approaches such as ECMWF and (2) the simple averaging of models' output tends to perform better than individual models.

Chapter 5 - *Evaluation of DNI forecast based on the WRF mesoscale atmospheric model for CPV applications*

The integration of large-scale solar electricity production into the energy supply structures depends essentially on the precise advance knowledge of the available resource. Numerical weather prediction (NWP) models provide a reliable and comprehensive tool for short- and medium-range solar radiation forecasts. The methodology followed here is based on the WRF model. For CPV systems the primary energy

source is the direct normal irradiance (DNI), which is dramatically affected by the presence of clouds. Therefore, the reliability of DNI forecasts is directly related to the accuracy of cloud information. Two aspects of this issue are discussed here: (i) the effect of the model's horizontal spatial resolution; and (ii) the effect of the spatial aggregation of the predicted irradiance. Results show that there is no improvement in DNI forecast skill at high spatial resolutions, except under clear-sky conditions. Furthermore, the spatial averaging of the predicted irradiance noticeably reduces their initial error.

Chapter 2

A novel procedure for generating solar irradiance TSYs

***Vicente Lara Fanego**, Jesús Pulgar Rubio, Carlos M. Fernández Peruchena, Martín Gastón Romeo, Sara Moreno Tejera, Lourdes Ramírez Santigosa, Rita X. Valenzuela Balderrama, Luis F. Zarzalejo Tirado, Diego Bermejo Pantaleón, Manuel Silva Pérez, Manuel Pavón Contreras, Ana Bernardos García, Sergio Macías Anarte (2017). A Novel Procedure for Generating Solar Irradiance TSYs. Am Ins Phys Conf Proc, 1850, 140015 (2017).*

2.1 Introduction

A Typical Solar Year (TSY) is a commonly used tool for design and bankability analysis of solar energy projects, such as CSP, and in many other fields. A TSY is similar to the more ubiquitous Typical Meteorological Year (TMY), but only includes solar radiation data, such as direct normal irradiance (DNI) for CSP applications, whereas a TMY also contains information about many additional meteorological variables, such as temperature, humidity, or wind speed. Early developments of TSYs were made at National Renewable Energy Laboratory (NREL) for either DNI or global horizontal irradiance (GHI) (Habte *et al*., 2014). They were referred to as TDY and TGY, respectively. Although the use of TMY or TSY is not recommended by experts for the precise design of solar-based renewable energy conversion systems (Stoffel *et al*., 2010b; Sengupta *et al*., 2015), such as CSP or PV, it has nevertheless become standard practice in the evaluation of the economic feasibility of such projects. Similarly to TMYs, TSYs are representative annual time-series of typical solar

radiation conditions expected at a specific location over a long time period, ideally the lifetime of the facility. A TSY does not correspond to any particular year of a specific period, but is artificially built as a composition of twelve "typical" months selected from different natural years, themselves extracted from a long-term time series of solar radiation data at the location of interest. The selection of those months to be used in the generation of the TSY involves a statistical characterization of the long-term time series of the dataset. The theoretical features of this synthetic year facilitate the objective quantification of the expected solar irradiance for different annual scenarios of available energy amount, usually measured by the probability of exceedance (Pe), and the corresponding estimated uncertainty. These valuable properties allow for an objective risk balance and safer analysis of the economic feasibility of the projects. Therefore, reliable TSYs for different solar resource scenarios—usually low-energy or "near worst case" (Pe90 to Pe99) and average or median (Pe50)—are demanded by the industry. Unfortunately, there is no scientific consensus on a standard method to generate such TSYs outside of the conventional Pe50 scenario. In fact, most likely only one method seems to have been proposed so far for the construction of TSYs corresponding to any probability scenario (Cebecauer and Suri, 2015). Nevertheless, the different strategies followed by the existing or future methods could undesirably generate different TSYs and thus different financing risk factors for the same initial meteorological dataset.

Within this framework, the Spanish Association for Standardization and Certification (AENOR) has established a working group of experts with the goal to design and standardize a method to generate TSYs for any solar energy scenario, specifically for applications in solar thermal power plants. In a first step, Fernandez-Peruchena *et al*. (2016) recently presented a study that showed that any Pe (hereafter, Pexx) could be inferred by the estimation of the continuous cumulative distribution functions (CDF) evaluated from long-term annual series of data of GHI and DNI. The present contribution presents a novel procedure for the selection of the most appropriate individual months (among all those

available in the long-term time series) to generate a TSY for any particular Pexx.

2.2 Methodology

For the generation of TSYs, the complete methodology must comprise two parts. In the first part, the annual values of each variable in the long-term time series are calculated. With this discrete number of annual values the estimated continuous CDF is derived using a Weibull distribution. The estimation of the parameters of the Weibull annual distribution is performed here using the fitdistrplus package (Delignette-Müller *et al*., 2015) in R version 3.2.4 ("R: The R Project for Statistical Computing," 2003) with the maximum likelihood method. This procedure is well described and analyzed in Fernandez-Peruchena *et al*. (2016). This estimated CDF provides the values for any annual probability of exceedance (Pexx), hence for any desired scenario, corresponding to design, bankability, etc. It should be clarified here that there is a common acceptation of using the concepts of probability of exceedance and percentile (normally referred to as Pe and P, respectively) as if they were the same. Both are complementary but should not be confused: for a determined percentile value the probability of exceeding that value is the complementary to 100%; for instance, for a percentile 5, the probability of being exceeded is 95%. Therefore it would not be correct to use a percentile 95 to refer to a scenario of low energy. Conversely, a probability of exceedance of 95% (Pe95) more appropriately means that the value will be exceeded 95% of the time. In this work we used the concept of probability of exceedance instead of percentile, because it is more commonly used by the industry.

The second part of the method corresponds more specifically to the analysis presented in this work, and is referred to as the EVA method, which is an acronym constructed from the Spanish words for seasonality and variability. Once the annual target value for a specific Pexx of interest is obtained from the estimated continuous CDF in the first part, the next step is to concatenate a subset of twelve calendar

months from the long-term time series of the original data, which might be measured (preferably) or modeled. Ultimately, this ensemble constitutes a TSY for the scenario determined by the target Pexx value. Therefore, the objective of the method is to determine which subset of twelve months shall be extracted from the complete long-term dataset. The EVA method is composed of two stages. In the first stage, the aim is to find those monthly values that respect two requisites: (i) in combination they must be statistically representative of the desired Pexx scenario; and (ii) their annual sum must correspond to the Pexx target value. Those monthly values are called "monthly expected values" (MEV), and they do not have to be necessarily equal to any of the available monthly values of the long-term time series. With this definition, the sum of all MEVs is exactly equal to the annual target value of Pexx. To determine these monthly values the method uses the following conditions:

1. For each month, the MEV should be the particular value that minimizes its distance to the corresponding monthly median (least-squares equation).
2. The MEV annual sum must be equal to the annual target value of Pexx (binding).
3. The intra-annual statistics that describe the natural behavior of the solar irradiance at the location of study is introduced by a composition of weights that conveniently modifies the least-squares equation. These weights, noted w_i, are determined by the product of two factors. The first one ($f^1{}_i$) measures the variability of the irradiance during each month relative to the others. Prior to obtain these values the seasonality of the monthly time series must be removed by means of a clear-sky model. For this study, the Bird clear-sky model described by Iqbal (1983) as Model C (Iqbal, 1983), particularized to a clean and dry atmosphere version, is used. As a measure of the monthly variability the statistic named median absolute deviation (MAD) is computed for each month. The second factor ($f^2{}_i$) accounts for each individual monthly energy contribution relative to the total annual energy. It is calculated

as the mean of the monthly values of each month available in the original time series. The weights are combined by the product: $w_i = f^1{}_i \cdot f^2{}_i$, for i=1,…,12 .

From a mathematical standpoint, the combination of the three conditions above can be described as a minimization problem with constraint. In other words, the problem consists in the minimization of the function

$$f(x_1, \dots x_{12}) = \sum_{i=1}^{12} \left(\frac{\sum_{i=1}^{12} w_i}{w_i} \right) \left(x_i - Pe_i^{50} \right)^2 \qquad (2.1)$$

With the following constraint:

$$\sum_{i=1}^{12} x_i = Pe_y \qquad (2.2)$$

Where:

$i = 1, \dots, 12$; month of the year.
x_i: monthly expected value (MEV) for month i.
Pe_i^{50}: median of the available values of month *i* relative to the long-term time series.
Pe_y: annual probability of exceedance at the *y* level (Pexx).
w_i: weight.

This minimization problem is analytically resolved by the method of Lagrange multipliers. After application of the procedure an equation is finally obtained. The unknowns of this resulting equation are the 12 MEV(x_i) values [(1)].

The second stage simply consists in finding the available monthly values that are closest in distance to the corresponding MEV. These distances, called residuals, are obtained as the absolute value of the

[(1)] The analytical formulation of the method is described in detail in Annex A

difference between the MEV and the available monthly solar radiation values. Finally, the 12 selected months constitute the desired TSY for the specific Pexx of interest.

In summary, this method can be said to be statistically based and analytically resolved. Because of its general definition it can be applied for both components GHI and DNI. Moreover, it should be pointed out that the method makes no assumptions about the monthly distributions. This is a key factor that justifies the selection of MAD as a measure of variability, since it is a robust statistic. Finally, it should be noted that the normalized weights are introduced inverted in function (2.1) by means of the factor $\left(\sum_{i=1}^{12} w_i/w_i\right)$. Therefore, for the ith month, a low value of w_i (low variability and low energy) makes the ith MEV value closer to the median value of the corresponding monthly distribution, comparatively to MEVs whose w_i are higher.

2.3 Results

The EVA method has been applied to a wide sample of different climatic locations around the world (Fernandez-Peruchena *et al*., 2016). The selection of these stations is based on the availability of long-term time series (at least 20 years) of high-quality data of surface solar irradiance. The results presented below correspond to the evaluation that has been carried out for DNI and GHI at the Burns radiometric station (BRN, 43.52 ºN, 119.02 ºW) of the University of Oregon's Solar Radiation Monitoring Laboratory (UO-SLMR, http://solardat.uoregon.edu/). The high quality of the original data has been reinforced through additional quality checks (Long and Shi, 2008). Note that the irradiance measurements had differing time resolutions depending on period (starting at hourly in 1980 and finishing at 5-min from 1995 on). For consistency, the time series has been homogenized to the conventional hourly basis throughout. Hourly data were then integrated to obtain monthly and yearly values. For this study, the time series of both DNI and GHI covered the complete 33-year period (1980–2012). The procedure has been applied to derive TSYs for a wide range of Pexx values, including the most commonly

ones required by the industry, namely: Pe99, Pe95, Pe90 and Pe50. For validation purposes, the generated TSYs have been analyzed against the initial long-term time series in order to consider the representativeness of each single artificial year under each possible Pexx scenarios.

To obtain an intuitive perspective of the problem, a possible way is to show the distributions of the monthly irradiation values, since the calendar month is the working unit of time adopted by the method. In Fig. 1, the distribution of the monthly DNI values is shown by means of boxplot diagrams. Different curves linking the twelve calendar months of a year would enclose the area of the total annual energy amount corresponding to different Pexx scenarios. Hence, for any specific value of this amount of energy determined by a certain Pexx, there is a set of curves whose integral approaches the annual target value of that specific Pexx. The aim of the method is to determine the shape of the particular curve that meets the former condition of the integral and provides the most representative behavior of the long-term irradiation among all possibilities.

In Fig. 1, the boxplot diagrams show the set of twelve distributions of monthly DNI values at the Burns station. For each boxplot, the interquartile range (IQR) amounts to 50% of the data, between Q1 (25%) and Q3 (75%), and the whiskers represent the quantity 1.5·IQR. The mean and median values of each monthly distribution are highlighted. Interestingly, each monthly distribution is different from the others, and a normal distribution cannot always be assumed in all cases. For instance, in May and September at the Burns station, some values are outside of the whiskers interval and can be considered outliers. These circumstances justify the preference of using the robust MAD statistic rather than the usual standard deviation as a measure of the variability of the monthly distributions. Even though the standard deviation is also a measure of variability in data samples, it presents two important disadvantages when the distribution cannot be assumed Gaussian: outliers can strongly influence the standard deviation value, and the standard deviation can force a preference for lower vs. higher values, or vice versa. Figure 1 also shows how the natural seasonal

tendency of the time-series exhibits a strong pattern with higher energy values during summer months, which is typical of temperate climates. In order to properly compare the variability of the 12 monthly distributions, the monthly dataset has been seasonally adjusted by means of the clear-sky model. Figure 1 presents the MEV values calculated with the EVA method for the "near worst-case" scenarios of Pe99 and Pe90 along with the MEV values obtained with the simplified method that would only use the variability factor $f^1{}_i$ to configure the weights. When only this variability factor $f^1{}_i$ is taken into account, the winter months (December, January and February) have higher weights due to their greater variability (note that, as defined in the Methodology section, the higher the variability the higher the weight, and thus the farther the MEV is to the median value of the monthly distribution, because of the special way the weights are defined in function (1)) Because winter months are naturally less energetic than summer months, MEVs for winter months should have very low values to compensate for the contribution of the less variable and more energetic summer months (June, July and August). In other words, winter months have to contribute with extremely low energy values relative to the total annual energy amount in order to achieve the unfavorable low-energy scenarios Pe99 or Pe90. This occurs in detriment of the higher energetic (but less variable) summer months, which cannot contribute to the low annual energy amount with low monthly values, but with energy values close to the median. As can be seen in Fig. 1, this is more pronounced for the extreme Pe99 case than for the milder Pe90 case. Counting only on variability factor $f^1{}_i$ by ignoring $f^2{}_i$ can produce a misrepresentation of the possible contribution of the high-energy months to the total amount of annual energy in unfavorable cases, such as those determined by scenarios Pe99 or Pe90. The high-energy months (as dependent on seasonality) should be considered a major potential contributing factor when extreme years of DNI and GHI must be constructed. This is simply because low values of high-energy months can notably reduce the available energy of the whole year, and should therefore be taken into account. With the EVA method, this is done by means of the energy factor $f^2{}_i$. As shown in Fig. 1, this

produces a different distribution of the contribution of each month relative to the annual Pexx value, depending on the variability and energy for that month. Thus, winter months may reach low DNI values without being extreme, whereas summer months may reach lower values instead of being forced to be closer to the median. In particular, Fig. 1 shows that, for the extreme case represented by Pe99, the MEV values for all months are below their respective IQRs. In contrast, when the method is applied using only the first variability factor $f^1{}_i$, in July the MEV value is still almost within its IQR, whereas extreme values are reached in January and December.

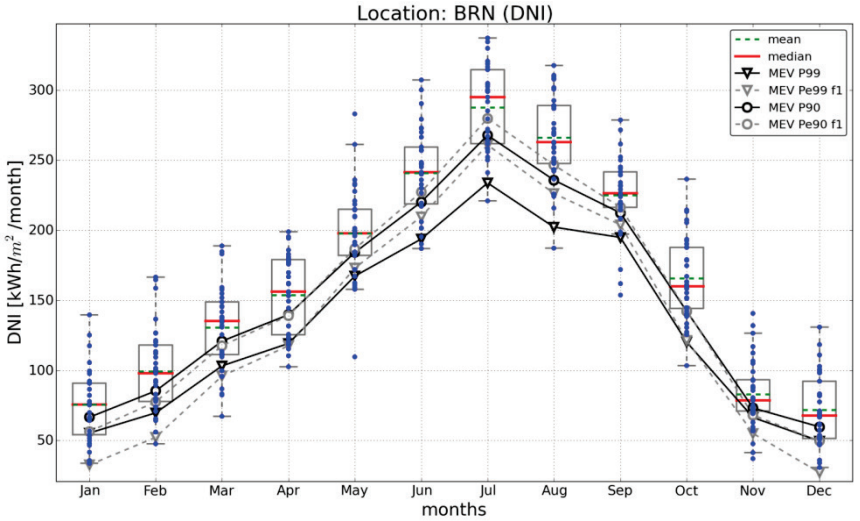

Figure 1. Boxplot diagrams of monthly DNI values for the complete dataset measured at the Burns (BRN) station. The IQR of the boxplots is the interquartile [Q1, Q3], and the whiskers equal to 1.5·IQR. The monthly mean and median values of the distribution are shown (green dashed and continuous red line, respectively). The MEV values for Pe99 and Pe90 annual values, along with those calculated from only the variability factor $f^1{}_i$ are also plotted.

By definition, the sum of the 12 MEVs is exactly equal to the target annual Pexx value. When minimizing the residuals, an error is inevitably introduced because the MEV values do not usually coincide with the available monthly values of the long-term time series, so that the sum of the residuals is not zero. These errors are presented in Figs.

2 and 3 along with the annual values of the constructed TSY over the estimated CDF curve, for DNI and GHI respectively. As shown in both figures, the value of the errors differs according to the specific target Pexx value. Errors are higher for DNI than for GHI, as could be expected since the former has more interannual variability than the latter (Gueymard, 2012c). Absolute values of the relative errors (in percent relative to the Pexx value) fall within the ranges 0.03–1.90 and 0.00–0.20 for DNI and GHI, respectively. The highest error is produced in the extreme case (Pe99) for the DNI variable (Fig. 2). This suggests that the more extreme Pexx is, the higher the error. However, this is not necessarily true in general.

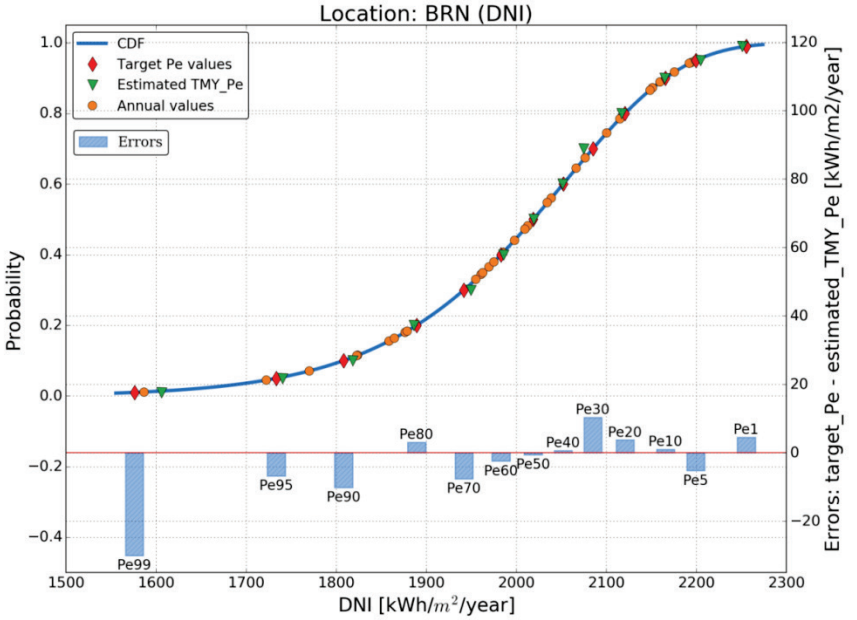

Figure 2. DNI annual values and estimated CDF at Burns (BRN) station, along with the annual target Pexx values and the estimated annual TSY-Pe values. Errors between the target and estimated TSY values for the different Pexx are also shown.

In Fig. 3, for instance, the errors for Pe60 and Pe10 are higher than those for Pe99. Furthermore, Figs. 1 and 2 show that the errors are not always of the same sign: they can indicate either overestimation (negative error) or underestimation (positive error). In general, it can be said that the magnitude of the error depends on the number of years available in the long-term time series to generate the TSY, because a longer time series implies more possibilities of finding actual monthly values that are closer to the MEV values. Finally, it should be highlighted that the errors are very low in all the cases presented in Figs. 2 and 3, and also for the pool of locations analyzed elsewhere (Fernandez-Peruchena *et al*., 2016) (not shown). The errors are below the usual standard limits that have been established to account for slight corrections –usually consisting in day substitutions- to obtain a better approach to the target Pexx.

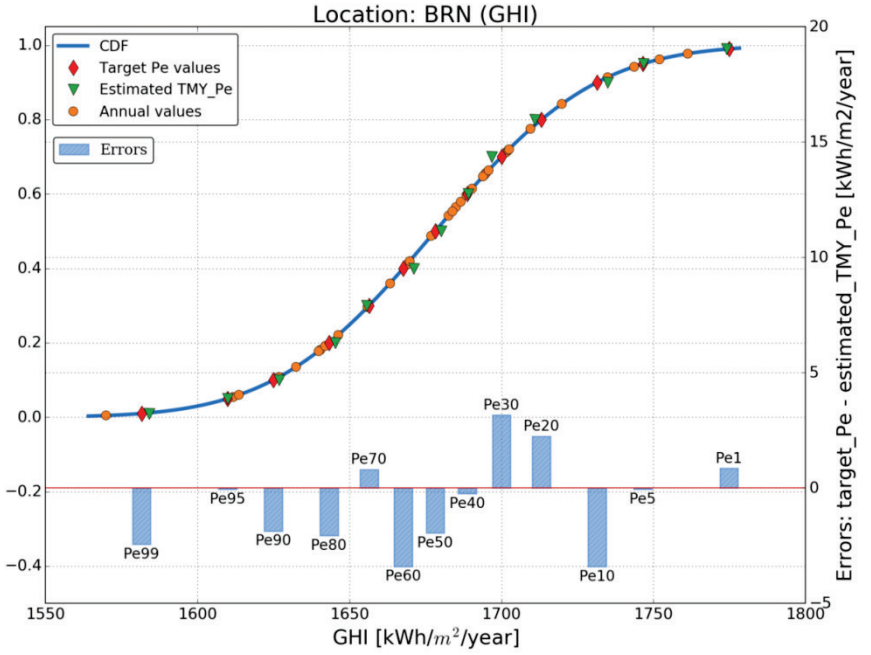

Figure 3. Same as Figure 2, but for GH.I

Finally, it is important to consider that, although the methodology to generate TSYs is based on a monthly time unit, the representative annual time series is most usually needed at higher time resolutions. In principle, since the TSY is constructed from blocks of monthly periods extracted from the original long-term time series, the original time resolution (e.g., hourly) should be conserved. In fact, this desirable feature of the TSY both restricts and determines the way it is defined. To evaluate the capability of the EVA method to generate representative years for higher time resolutions, the constructed TSYs are set to the original hourly resolution of their constituent months in order to compare the frequency distribution of the hourly TSYs with that of the long-term. Figure 4 shows the frequency distributions of TSY for Pe95 and Pe50 in the case of DNI at hourly time resolution, compared to that of the original long-term time series.

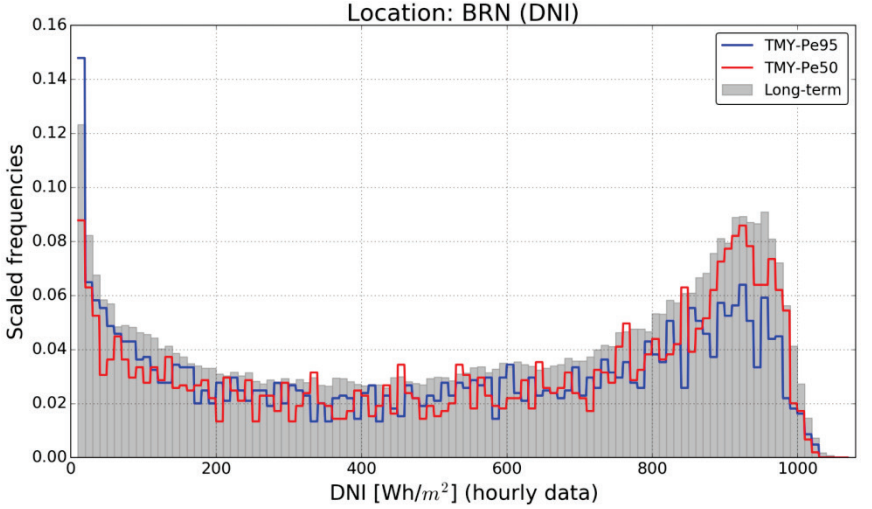

Figure 4. Frequency histograms of sun-up hourly values of DNI for generated TSYs at Pe95 and Pe50 and for the original long-term time series.

The figure shows that the shapes of the frequency distributions of the TSYs are quite similar to that of the long-term. The reduction in annual energy from the Pe50 to the Pe95 scenarios is mainly produced by the higher DNI values, which appears logical.

2.4 Conclusions

TMYs and TSYs are demanded by the solar energy industry because of their usefulness, mainly for the design and bankability analysis of solar projects. An issue, however, is that some temporal variability information is lost during the construction of such artificial years, with respect to the long-term time series of solar data they are based upon. Thus, the use of TMYs or TSYs is not recommended by experts for the precise design of solar energy conversion systems. Nevertheless, TMYs and TSYs have become a standard source of data for typifying the expected energy production at CSP or PV plants. Due to the lack of scientific consensus on how to define a method for generating TSYs, several initiatives have surfaced to propose methods to generate "standard" TSYs or TMYs. The design characteristic of representativeness and the constraint of being conformed by data of the historical long-term time series are the commonly elements used by the proposed methods. Whereas well-established methods exist to develop synthetic years for average or median conditions, the current challenge is to represent more extreme solar resource situations, such as what a financial institution would consider a worst-case scenario for interest repayment, which is indicated by the widely-used concept of probability of exceedance.

In this work, a novel procedure for generating TSYs of solar irradiance -both DNI and GHI components- for any probability scenario is presented. The method, referred to as the EVA method, is based on statistical criteria and it has an analytical definition. The objective of the method is the determination of the energy monthly values of each calendar month whose annual sum is equal to the annual target value of probability of exceedance Pexx. These 12 months are called monthly expected values (MEV). The TSY is comprised by the available 12 calendar months whose absolute difference respect to the corresponding MEV is minimal. The EVA method makes use of a composition of weights that accounts for: i) the variability of the seasonal adjusted monthly distributions, ii) the individual monthly

energy contribution of each calendar month respect to the total annual energy.

The results show that the method provides reasonable results, with low differences between the annual target value for a certain probability of exceedance and the annual value of the irradiance generated by the TSY. It also preserves the long-term statistics when the constructed TSY of a determined probability of exceedance is set to the original higher time resolution of its constituents months (1 hour). It has also other valuable properties such as its statistical base and analytical definition, flexibility and facility to be implemented in a software code.

Deeper research should be carried out to extend the pool of sites where the evaluation presented here is possible. Further work should also analyze other essential aspects of bankability, like uncertainty. In particular, it would be important to establish how to take the uncertainty in the EVA method into account with respect to the total uncertainty in the generation of a TSY. Such uncertainties include those related to the generation of the data in the original long-term time series, along with those due to the representativeness of the available dataset (usually of limited duration) with respect to actual long-term conditions. It would also be useful to compare results obtained from long-term time series of measured data to those derived from other sources, such as satellite-based modeled data. Finally, it would be important to carry out a study in which the solar plant's energy production would be analyzed in direct connection with the solar irradiance data. Such study would examine the relations between the solar resource and the expected vs. actual energy production, as affected by the use or not of a TSY at the design stage.

Acknowledgments

This work has been possible thanks to the availability of a long period of historical time series of the variables analyzed. The authors would like to give special thanks to the past and current staff of the UO SRML network that has been making the great and important effort of maintaining and operating the stations along all these years. Authors

would also like to thank to Dr Christian Gueymard for the valuable contributions that he has kindly provided with the revision of this work.

Chapter 3

Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain)

***Lara-Fanego V**, Ruiz-Arias J A, Pozo-Vázquez D, Santos-Alamillos F J, Tovar-Pescador J (2012). Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain). Sol Energy 86(8):2200–2217. Citations in WEB of Science: 65. Impact factor: 3.685*

3.1. Introduction

Much of the focus on sustainable energy is aimed at different ways of tapping the most abundant renewable resource: solar energy (Szuromi *et al*., 2007). There are mainly two different ways for solar electricity production: by Solar Thermal Power Plants (STPP) and Photovoltaic plants (PV). The STPPs use the direct normal solar irradiance (DNI) to convert solar energy through focusing receivers into heat, which is then used to drive a thermodynamic cycle and thereby produce electricity. PV systems enable direct conversion of global horizontal irradiance (GHI) into electricity through semi-conductor devices. Technical potential estimates for global STPP have been evaluated at several hundred of GWe by 2030 (IEA, 2006a, b), or about 2% of the global electricity demands. Similarly, PV electricity is estimated to have a technical potential of 205 GW by 2020 or about 2% of global electricity demand (EPRI, 2003).

A major challenge for the future will be the integration of these large scale solar yields into existing energy supply structures. The

problem arises from the fluctuating character of the solar resource, as compared to conventionally generated electricity, and its dependence on non-deterministic weather patterns. Particularly, governments support is necessary for these (and others) renewable energy sources in electricity because, although desirable from a social welfare perspective, their private costs are not competitive in the power generation systems because of its intermittent production creates negative externalities as those associated with grid integration costs. The experience with wind energy shows that accurate wind speed forecasts can substantially reduce grid integration costs (Saintcross *et al*., 2005; Tindal *et al*., 2006). Similarly, accurate information on expected solar irradiance can be used for the management of the electricity grids, for scheduling conventional power plants and also for decision making on the energy market. This should, finally, end reducing the solar energy yield integration costs (IEA, 2007).

Spain can be regarded as a test site since it is one of the world leading countries in electricity production from renewable sources (IEA, 2010). This is because the legal and economic instruments to promote the renewable energy have had success (Perez and Ramos-Real, 2009a). Particularly, the Royal Decree 661/2007 (Royal Decree, 2007) established the framework in Spain regarding the legal and economic aspects of the production of electricity in the special regime. It provides two options for the producers: either transfer the yield to the power distribution company with the electricity sale price stated as a single regulated tariff (tariff model), or sell on the free market at the going market price plus a premium (premium model), which, for solar electricity, is set at 250% of its average electricity tariff. Operators of installations are obliged to provide the distributor with a forecast of the electricity they intend to feed into the grid the next day by at least 11:00 h (local time) of the previous day. Penalties are established for deviations: the cost of deviation is (10% of the spot market prices applied to the forecast deviations) when the permitted tolerance (20% for solar and wind power) is exceeded. The premium model, therefore, includes incentives to the correct prediction of the solar yield for the next day, thus providing the grid operator the expected electricity

production for the day ahead and, then, allowing a reliable basis for the grid operation for the next day. The case of the wind power forecast in Spain can be considered as a reference. Today, there are about 20GWe of wind power installed capacity (20% of the total installed power), with 90% of the producers having taken the market-based retribution (Perez and Ramos Real, 2009b).

Concerning the solar energy systems, particularly the STPP, today Spain is the leading country in the world with about 500 MWe in operation. In addition, there are 13 plants under construction and more than 200 planning projects totalling about 5 GWe (Fernández-García *et al*. 2010); most of them are located in the study region object of this work. Regarding PV systems, the leading country in the world is Germany, with about 5 GWe installed capacity followed by Spain with about 4 GWe (Salas and Olias 2009). For the next decade the Directive 2009/28/CE of the European Parliament and Council, establish, for Spain, a goal of 20% of primary energy from renewable origin by 2020. To achieve this goal, the new National Renewable Energy Plan for the decade 2011-2020, now under discussion, aims to increase the PV and STPP installed capacity in 13 GWe by 2020. This will suppose about 20% of the total installed capacity of the Spanish electricity system. To manage the electricity grid with such amount of solar energy will require high-quality information on every aspect of solar power generation, particularly, the solar radiation forecasting, as it is already the case for wind power generation. But, unlike the wind power, solar yield forecast is still on an early state and very few works have dealt with the forecasting of the solar resources and its application for management of solar-based electricity power plants and grid integration strategies. Different approaches to forecast irradiance can be taken depending on the target forecasting time. For very short time forecasts (up to 6 hours, nowcasting) approaches based on extrapolating the solar radiation field from cloud motion have been proposed (Heineman *et al*., 2006). These forecasts are meant for solar field control in STPP and PV plants. In addition, statistical techniques have been proposed for forecasting solar irradiance with up to 24 hours (Mellit and Pavan, 2010). However, Numerical Weather Prediction (NWP) models are the

basis of solar yield forecasts (IEA 2007) with up 48 hours of time horizon, the time range useful for grid integration and decision making in the energy market. In addition, these models provide forecasts of other parameters as temperature, relative humidity or wind speed, also useful for plant operating strategies. NWP models use atmospheric reanalyses as initial and boundary conditions for the model run, which then realistically downscale (using physical equations) to a finer physical resolution. The NWP model that downscales reanalysis data is termed a mesoscale model. Because the mesoscale models run over a smaller area than global scale models, the physics can include additional details. Therefore, provided sufficient computing power, these models can be used to forecast solar irradiance over a wide area with high temporal and spatial resolution.

Earliest evaluation studies on the MM5 (Grell *et al*., 1995) mesoscale model reliability for estimating GHI were carried out by Zamora *et al*. (2003, 2005) in some locations in USA. Heinemann *et al*. (2006a) evaluated the MM5 model GHI forecasts in Germany for lead time up to 48 hours. Results showed relative RMSE values of about 50%. Lorenz *et al*. (2009b) evaluated several NWP-based GHI forecasts in Europe. Overall, results showed relative RMSE values of about 40% for Central Europe and about 30% for Spain. Lorenz *et al*. (2009a) evaluated hourly GHI forecasts, based on the European Centre for Medium-Range Weather Forecast (ECMWF) NWP model, for power prediction of PV systems in Germany. They reported relative RMSE values of about 35% for single stations for a 24 hours horizon forecasts. Remund *et al*. (2008) evaluated different NWP-based GHI forecasts in the USA, reporting relative RMSE values ranging from 20% to 40% for a 24 hours forecast horizon. Similar results were reported by Perez *et al*., (2009b), evaluating NWP-based irradiance forecasts in several places in the U.S.A.

Regarding the forecasts of DNI, an additional problem comes into the scene, since DNI is not provided by the NWP model. Consequently, and additional processing procedure is needed. Breitkreuz *et al*. (2009) proposed a model (called AFSOL) based on the combined used of

information provided by a NWP model, an air-quality model and remote sensing retrievals. The model substantially relied on correct forecasts of the aerosol load (Breitkreuz *et al*., 2007). Evaluation was carried out in Europe and Northern Africa, reporting relative RMSE values of around 20% for the DNI forecasts under clear sky conditions. Wittmann *et al*. (2008) analyzed the Spanish premium feed-in tariff model in a case study for a STPP in Andalusia, based on DNI forecasts provided by the AFSOL (Breitkreuz *et al*. 2009) model. Interestingly, results proved that the economic benefits than can be achieved based on these forecasts are strongly depended on the time of the day at which forecast deviations take place.

In this work, we present a comprehensive evaluation study of the reliability of GHI and DNI forecasts based on the Weather Research and Forecasting (WRF; Skamarock *et al*., 2008) mesoscale model. The study is carried out in Andalusia (southern Spain), with a temporal horizon of up to 72 hours. The aim is to establish the current performance of the WRF model for solar yield forecasting in the study region, one of the areas with larger solar electricity installed capacity in the world (Fernández-García *et al*. 2010). It should be emphasized that not statistical post processing, usually depending on local radiometric measurements, was performed upon the forecasts provided by the model. Therefore, although obtained for a particular region, results can be regarded as representatives for regions with a similar climate. WRF is the mesoscale most used model around the world and it has been extensively assessed. In this work, DNI forecasts were obtained based on a physical post processing (Ruiz-Arias *et al*, 2011) which uses WRF outputs and satellite retrievals. Since the methodology aims to be fully operational, only satellite retrievals readily available at the model integration time are considered. The evaluation of the GHI forecasts was conducted based on radiometric data collected, along the years 2007 and 2008, at four validation locations in Andalusia. For the DNI, only two of these four validation stations were used. The assessment was carried, independently, for GHI and DNI and each season of the year, as a function of the sky conditions and forecast horizon.

The work is organized as follows: in Section 3.2 the experiment design is presented. Particularly, this section includes the description of the study area and the observational data, the WRF model setup, the post-processing methodology to derive DNI and, finally, a description the forecasting evaluation procedure. The evaluation of the model forecasts based on the observation is presented in Section 3.3. Finally, a summary of the results and some conclusions are provided in Section 3.4.

3.2 Experiment design

3.2.1 Study area and observations

The study area (Fig. 1) covers the region of Andalusia, in the southern part of the Iberian Peninsula. The region extends over an area of roughly 87,000 km^2, with approximately 1,000 km of coast line, and a varied topography. Particularly, western part of the region is an almost homogeneous flat area with a mean elevation of around 100 m.a.s.l. This area extends around the lower Guadalquivir river basin, which flows into the Atlantic Ocean. On the other side, eastern part presents a very complex topography, with steep elevation gradients reaching altitudes over 3,000 m.a.s.l. in the Mulhacén Peak, the highest summit in the whole Iberian Peninsula. The whole region is located in the transition zone from temperate to subtropical climates, presenting a Mediterranean climate ruled by the Azores high. Precipitations occur primarily from autumn to spring, associated mainly with Atlantic frontal systems. Nevertheless, the topographic and geographic characteristics produce a wide range of weather and climate conditions. Particularly, in the region coexist one of the rainiest areas in the Iberian Peninsula at the west (Sierra de Grazalema) with the unique desert area in Europe at the very east (Desierto de Tabernas).

Results of the present study have been evaluated based on ground data collected at four radiometric stations (Fig. 1). These stations were selected, among the available, based on two criteria. Firstly, because of their location, close to areas where solar energy plants are now

operating or under construction. Secondly, because of their representativeness, at least partially, of the above commented climate variability within the study region. Particularly, station labelled as 1 corresponds to the airport of Córdoba (4.8506 º W, 37.8444 º N; 91 m.a.s.l.) situated 6 km outside the city. This station is in the very centre of the study region, at the middle of the Guadalquivir river basin, and presents a continental Mediterranean climate with Atlantic influence. Annual precipitation is about 500 mm. Station labelled as 2 is located at the Andasol STPP (3.1680 º W, 37.2180 º N; 1150 m.a.s.l), near Guadix (Granada) at the east of the study region. It is located in one of the highest plateaus of Spain, bounded at the south by the Sierra Nevada mountain range, that isolates this region from the influence of the Mediterranean Sea. As a consequence, the area presents a dry continental climate, with extreme temperatures in summer and winter. Annual precipitation is about 400 mm. The third station (labelled as 3) is sited in Huelva (6.9097 º W, 37.2800 º N; 19 m.a.s.l.), at the western part in the Atlantic coast. As a consequence, presents a mild climate with relatively high precipitation (about 700 mm/year). The last station (labelled as 4) is also in the western part of Andalusia and corresponds to the airport of Jerez de la Frontera (6.0633 º W, 36.7458 º N; 27 m.a.s.l.). The airport is 8 km far from the city and about 25 km from the Atlantic coast without topographic obstacles in between. To sum up, these two last stations (3 and 4) are representative of the climate of the western area of the study region, while station 2 is more representative of the eastern area. Station 1, located in the centre of the region is, probably, the most representative of the region climate as a whole.

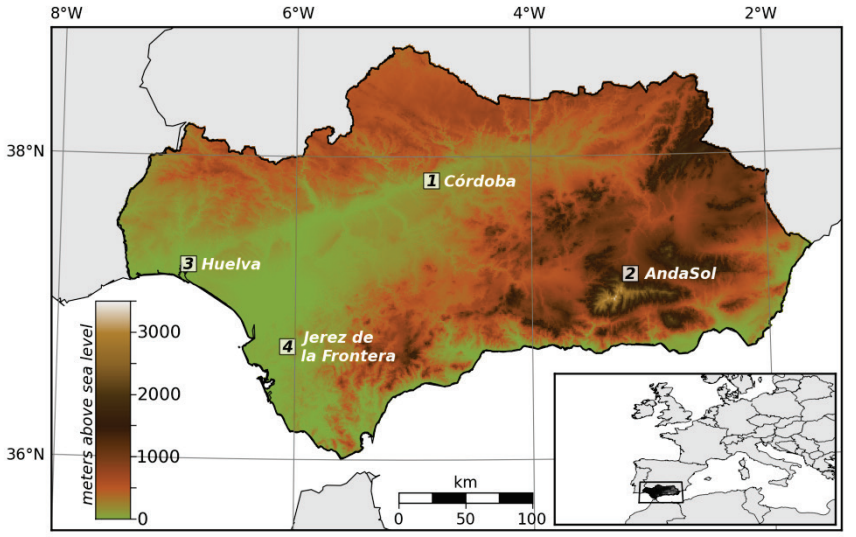

Figure 1. Study region and radiometric stations locations. Scale at the left represents the elevation.

GHI forecasts were evaluated based on hourly values measured, along the years 2007 and 2008, at the four ground station. Particularly, GHI forecasts evaluation was carried out for the month of August 2007 (as representative of the summer season), September 15th to October 15th 2007 (autumn), 15th February to 15th March 2008 (winter) and April 2008 (spring). DNI forecasts evaluation was carried out based on hourly irradiation data collected at the Córdoba and Andasol stations for the same above mentioned periods.

The stations at Córdoba, Huelva and Jerez de la Frontera are part of the AEMet (Spanish Weather Service) net of meteorological stations. GHI data were collected by Kipp & Zonen (K&Z) CM11 pyranometers in the stations of Huelva and Jerez de la Frontera, and by K&Z CM21 at Córdoba. DNI was measured by a K&Z CH1 pyrheliometer at this last station. Each sensor was calibrated biannually. Standard quality control tests, as proposed by the World Radiation Data Center (http://wrdc.mgo.rssi.ru), are routinely applied by AEMet. Data corresponding to the Andasol STPP were provided by Solar Millenium AG. GHI and DNI data collected at this location was measured based

on a RSR-2 rotating shadow-band radiometer (Gueymard and Myers, 2009). For the present work, 10-minutes measurements were averaged to produce hourly values.

An additional quality control check was applied both to AEMet and Andasol radiometric data. This test, described in Long *et al*. (2008), filtered out data above physical limits established and recommended by the Baseline Surface Radiation Network (BSRN) (http://www.gewex.org/bsrn.html). About 0.5% of the data failed to pass the test and was removed from the original data set. Finally, measurements corresponding to solar zenith angles greater than 85 degrees were also rejected, in order to avoid instrumental errors.

3.2.2 WRF Setup

Simulations with the non-hydrostatic WRF-ARW mesoscale model (version 3) were conducted for the above mentioned periods, representative of the different year seasons. The simulations were driven based on the National Centre for Environmental Prediction (NCEP) global analyses. The temporal resolution of the analyses was 6 hours, while the spatial resolution was 1ºx1º (aprox. 100 km^2 at the experiment latitudes). The model configuration included three successive nested domains, with grid-spacing of 27 km (the coarser grid, hereinafter domain 1), 9 km (domain 2) and 3 km (domain 3, the finest grid), and 27 unevenly spaced vertical levels. Estimates corresponding to the grid points of the innermost domain (domain 3) that enclose the experimental radiometric stations were used in the evaluation procedure.

The WRF model has a wide range of physical parameterizations, which allow setting the model to better describe the physical processes based on model domain, resolution, location and application. In this work, the different parameterizations were selected following Ruiz-Arias *et al*. (2008), who conducted an evaluation study of the performance of the different parameterizations of the model in the same study region. Hence, the model was operated with the Thompson microphysic scheme (Thompson *et al*., 2004), the YSU planetary

boundary layer scheme (Hong *et al*., 2006), the Kain-Fritsch cumulus scheme (Kain and Fritsch, 1993), the unified Noah land-surface model, the RRTM scheme for long-wave radiation (Mlawer *et al*., 1997) and the scheme of Dudhia for shortwave radiation (Dudhia, 1989).

For each day of the different evaluation periods, a three-days-ahead (72 hours) hourly resolution forecasting simulation was carried out, starting at 00 h UTC of the following day. Between domains 1 and 2, the two-way nesting option was selected to allow the interaction of grids in both directions. For the domain 3, convective parameterization was turned off due to its high spatial resolution. Hence, the interaction option with its parent domain (domain 2) was set to one-way.

3.2.3 Post-processing to derive DNI

Current radiation schemes of the WRF model, and the vast majority of the NWP models, provide GHI forecasts, but not DNI forecasts. In this case, and as first approach, a statistical model (e.g. Boland and Ridley, 2008; Ruiz-Arias *et al*., 2010c) can be used to derive DNI from GHI. Nevertheless, WRF offers a detailed description of the current and future state of the atmosphere and allows a more precise disaggregation of the global solar irradiance into its components. In this work, a physical post-processing, based on the WRF model outputs and satellite retrievals (Ruiz-Arias *et al*., 2011), has been used to obtain DNI forecasts. Since this post-process aims to be fully operational, only information readily available elsewhere at the time of running the WRF model has been considered. Figure 2 presents an outline of the procedure. Firstly, the total broadband transmittance for cloudless skies was computed based on the parameterization proposed by Gueymard (1998). The water vapor content was derived from the WRF outputs. The aerosol and ozone loads were obtained from the Giovanni online data system, developed and maintained by the NASA Goddard Earth Sciences (GES) Data and Information Services Centre (DISC) (http://disc.sci.gsfc.nasa.gov/giovanni/). The aerosol optical depth (AOD) at 550 nm and the Angström's exponent for the study region belong to the MODIS/Terra Aerosol Daily L3 Global 1Deg CMG

dataset (spatial resolution of 1º×1º). From this dataset, the Angström's turbidity coefficient was calculated using the Angström's law. The column amount of ozone belongs to the EOS Aura OMI version 3 daily level 3 global 0.25°×0.25° gridded data, based on the TOMS algorithm. Next, the water and ice clouds transmittance was derived from the WRF model following the parameterization proposed by Hu and Stamnes (1993).

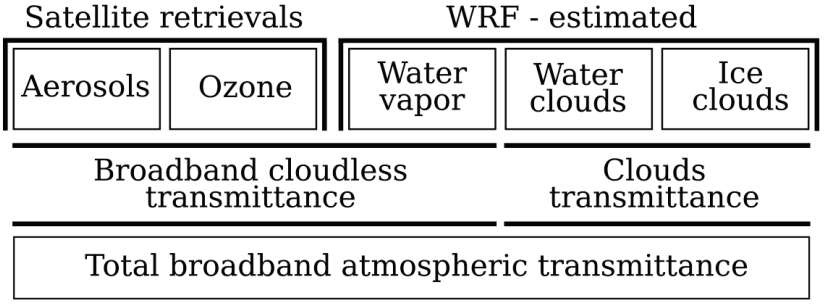

Figure 2. Outline of the post-processing methodology used to derive DNI based on the WRF model outputs and satellite retrievals.

To make the post-processing fully operational, both the aerosols and ozone content correspond to the previous day to the forecasting time and was kept fix along this period. Aerosol loads may experience important temporal variability at the inter-daily scale. Therefore, the use of the same aerosol information for the whole 72 forecast run may introduce error in the DNI forecasts (Breitkreuz *et al*., 2009). To address this issue, the autocorrelation of the AOD at 550 nm at the Córdoba and Andasol stations (Fig. 3) for the whole period of study were computed. As expected, autocorrelation values decreases as time lag increases, but still they are considerable high for both stations up to lag 3 days. Particularly, the decrease is considerable lower for the Andasol station, where aerosol loads tends to be more persistent and where, at lag 3 days, autocorrelation value is still close to 0.9. For the Córdoba station, autocorrelation values decreases up to 0.75 at lag 3 days. Therefore, the error in DNI forecasts associated with keeping the same aerosol information along the three days of forecast seems to be low for the Andasol station. However, for the Córdoba station, the error

should be slightly higher, particularly for the third day of the forecasting period. Nevertheless, it should be highlighted that these errors can be considerable higher when Saharan air intrusions (dust storms) take place over the study region. Mainly because the arrival of these air masses may introduce steep changes in the background AOD and its properties (Lyamani *et al*., 2005).

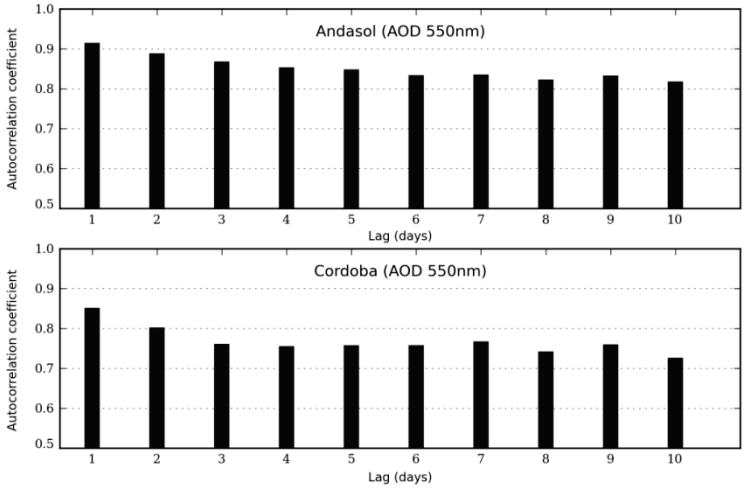

Figure 3. Autocorrelation values of the aerosol optical depth at 550 nm, derived from the MODIS data set, for Córdoba and Andasol validation stations. Autocorrelation values are displayed as a function of the time lag, in days, and were computed for the whole period of study (annual values).

3.2.4 Evaluation procedure

Two different evaluation analyses were carried out. In the first analysis, the performance of the WRF model is analysed based on the season of the year and for the annual period, with no distinction on the sky conditions. The aim is to evaluate the influence of the different meteorological conditions along the year in the forecasting skill of the model. The analysis is conducted independently for GHI (section 3.3.1) and DNI (section 3.3.2). Secondly (sections 3.3.3 and 3.3.4), the performance of the WRF model is analyzed based on the sky

conditions. Particularly, three different daily-scale scenarios were considered: clear sky, cloudy and overcast. Results shall be useful from an operational perspective since, overall, sky conditions for the following day (daily mean) are easier to forecast that individual (hourly mean) sky conditions. The three different sky conditions considered in this work were established based on the clearness index, defined as the ratio of the hemispherical horizontal total global solar irradiance to the total horizontal extraterrestrial irradiance (Perez *et al*., 1990b). Particularly, in this work, a daily averaged clearness index was used:

$$k_t = \frac{\int_{1h}^{24h} I dt}{\int_{1h}^{24h} I_0 \cdot e_0 \cdot \cos\theta_z dt} \approx \frac{\sum_{i=1h}^{24h} I_i}{I_0 \cdot \sum_{i=1h}^{24h} e_0 \cdot \cos\theta_{zi}} \tag{3.1}$$

where I is the hourly global horizontal irradiance at the earth' surface, I_0 = 1367 Wm^{-2} is the solar constant, $e_0 = (r / r_0)^2$ is the eccentricity correction factor at each hour and θ_z is the hourly averaged solar zenith angle. The k_t values were computed for each forecasted day based on the ground station measurements.

Values of daily k_t greater than 0.65 were considered indicative of clear sky conditions, values between 0.65 and 0.4 were considered as cloudy conditions and values below 0.4 as overcast conditions. These two analyses were carried out independently for each ground station and as a function of the forecasting lead time. Particularly, evaluation was carried out independently for the 24 hours (D1 hereinafter), 48 hours (D2) and 72 hours (D3) forecast horizons.

Forecasts were evaluated in terms of the mean bias error (MBE) and the root mean square error (RMSE), defined in absolute terms as (Wilks 1995, Perez *et al*., 1997, Lorentz *et al*., 2009b):

$$MBE = \bar{\varepsilon} = \frac{1}{N} \cdot \sum_{i=1}^{N} \varepsilon_i \tag{3.2}$$

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \varepsilon_i^2} \tag{3.3}$$

where $\varepsilon_i = x_f - x_o$ are the residuals (forecast errors), calculated as the difference between the forecasted values (x_f) and the observed values (x_o), and N is the total number of values. All error estimates were computed using hourly values for the considered period, night values with no irradiance were excluded from the evaluation. MBE quantifies the overall bias and detects if the model is producing overestimation (MBE > 0) or underestimation (MBE < 0). On the other hand, RMSE accounts for the spread of the error distribution. Also relative error measures (rMBE and rRMSE) were computed. Normalization is done with respect to the mean ground measurement irradiance in the considered period.

Finally, GHI and DNI forecasts were tested against persistence, a trivial reference model, in terms of the RMSE. Persistence model forecast errors were computed using the previous 24 hours observed values at the same hour for the D1 forecast. Similarly, for the D2 and D3 forecasts, persistence was computed using, respectively, the previous 48 and 72 hours observed values.

3.3 Results and discussion

3.3.1 GHI forecasts evaluation results: dependence on time horizon and seasonality

Tables 1 to 3 report the RMSE and MBE values for the locations of Andasol, Córdoba and Jerez, as a function of the forecast horizon and the season of the year, including the annual mean errors.

Positive values of the MBE in Tables 1 to 3 indicate that WRF model tends to overestimate the GHI for all the study locations and seasons of the year. As explained in section 3.3.3, this systematic overestimation can be related to the solar irradiance parameterization of the WRF model used in this work, namely the Dhudia scheme (Dudhia, 1989). MBE values tend to be lower in summer compared to the rest of the seasons, except for Andasol station (Table 1), where summer bias values are relatively high (about 16%). In addition, this station presents the highest MBE values for all the seasons. Particularly, for the annual data and the D1 forecasts, MBE values ranges from about 7% for Huelva station (not shown) to about 14% for Andasol station (Table 1).

Table 1. GHI forecast evaluation results as a function of the season of the year and time horizon for Andasol station. The WRF model MBE and RMSE forecasting values are show in absolute (W/m^2) and relative magnitude (in brackets at the right, in percentage). The last column shows the persistence trivial model forecasting RMSE values both in absolute and relative values.

Forecast horizon	WRF MBE	WRF RMSE	Persistence RMSE
Summer			
Day 1	81 (16)	168 (33)	192 (38)
Day 2	86 (17)	173 (34)	189 (37)
Day 3	59 (12)	174 (34)	167 (33)
Autumn			
Day 1	34 (8)	151 (37)	203 (50)
Day 2	46 (11)	172 (43)	212 (53)

Day 3	44 (11)	201 (50)	216 (54)
Winter			
Day 1	50 (13)	167 (43)	160 (41)
Day 2	40 (10)	177 (44)	173 (43)
Day 3	37 (9)	178 (44)	175 (43)
Spring			
Day 1	87 (17)	188 (36)	198 (38)
Day 2	72 (14)	179 (34)	227 (43)
Day 3	88 (17)	201 (38)	230 (44)
Anual Period			
Day 1	64 (14)	170 (37)	190 (41)
Day 2	62 (13)	175 (38)	202 (44)
Day 3	59 (13)	189 (41)	199 (43)

Overall, the higher MBE from autumn to spring, the rainy (cloudy) seasons in the study area, indicates the limited ability of WRF model to forecast cloudy conditions. This issue is further analyzed in section 3.3.3. Note that there is not a clear tendency of the MBE to increase with the forecast horizon.

The RMSE shows a clear seasonal dependence, with lower values in summer (Tables 1 to 3). For the rest of the seasons, RMSE values are alike and about one third higher. Again, these differences are related with the cloudiness, more difficult to forecast than the clear-sky conditions (see section 3.3.3). As in the case of the MBE, Andasol station

(Table 1) presents the highest RMSE values for almost all the seasons and forecast horizons. Differences are more accused in summer, when RMSE values are about one third higher for Andasol than for the rest of the validation stations. As a consequence, for the annual period and the D1 forecasts, RMSE ranges from 27% in the case of Huelva station (not shown) to 37% for Andasol station (Table 1). As expected, the

forecasts accuracy tends to decrease with the forecast lead time, except in some cases; for instance, during autumn for all the validation stations and summer for Jerez (Table 3) and Huelva (not shown) stations, the two stations with more similarities in the climate conditions.

The comparison between the WRF and the trivial persistence model performances shows some interesting features. Firstly, persistence model RMSE values tend to be, as expected, higher than the WRF model RMSE values. Nevertheless, performance differences vary with the validation station, forecast horizon and, mainly, season of the year. Particularly, WRF performs considerable better than the persistence during the summer for all the validation stations. For instance, for this season and except for Andasol station, the trivial model D1 forecasts RMSE values are about one third higher (Tables 2 and 3). For the rest of the seasons, differences are considerable lower. Even in some cases, the trivial model performs slightly better, as during winter for Andasol station (Table 1) or during the spring and D1 forecasts at Jerez (Table 3) and Huelva (not shown) stations. The persistence model accuracy decreases with the forecast lead time, particularly during spring (see Table 2 and 3). This is probably related to the unstable and highly variable weather condition that the study region undergoes during this part of the year. On the other side, the WRF model performance shows a lower dependence on the time horizon during this part of the year, indicative of a higher ability of the model to properly forecast cloudy conditions. As a consequence, except for Andasol station, differences between the WRF and the trivial model RMSE values reach the maximum during the spring for D3 forecasts (Tables 2 and 3).

Table 2. As in Table 1 but for the Córdoba station.

Forecast horizon	WRF MBE	WRF RMSE	Persistence RMSE
Summer			
Day 1	38 (7)	107 (19)	132 (23)
Day 2	41 (7)	101 (18)	144 (26)
Day 3	43 (8)	113 (20)	147 (26)
Autumn			

Day 1	51 (13)	134 (35)	158 (41)
Day 2	63 (16)	164 (43)	182 (48)
Day 3	53 (14)	160 (42)	207 (56)
Winter			
Day 1	59 (14)	140 (33)	157 (36)
Day 2	43 (10)	120 (28)	184 (41)
Day 3	59 (13)	140 (32)	172 (37)
Spring			
Day 1	57 (12)	160 (34)	196 (41)
Day 2	38 (8)	157 (33)	233 (48)
Day 3	43 (9)	160 (35)	265 (55)
Anual Period			
Day 1	50 (11)	136 (29)	163 (34)
Day 2	45 (9)	136 (29)	189 (39)
Day 3	48 (10)	141 (31)	204 (42)

Table 3. As in Table 1 but for the Jerez station.

Forecast horizon	WRF MBE	WRF RMSE	Persistence RMSE
Summer			
Day 1	32 (6)	106 (19)	152 (28)
Day 2	34 (6)	128 (23)	166 (30)
Day 3	35 (6)	139 (26)	166 (31)
Autumn			
Day 1	35 (8)	129 (31)	136 (33)
Day 2	28 (6)	136 (33)	151 (36)
Day 3	32 (8)	143 (34)	147 (35)
Winter			
Day 1	30 (7)	142 (35)	163 (40)
Day 2	46 (11)	148 (36)	180 (44)
Day 3	39 (9)	140 (34)	178 (43)

Spring			
Day 1	95 (19)	179 (37)	165 (34)
Day 2	94 (19)	171 (35)	215 (44)
Day 3	95 (19)	180 (37)	251 (51)
Anual Period			
Day 1	49 (10)	142 (30)	154 (33)
Day 2	48 (10)	147 (31)	180 (38)
Day 3	48 (10)	152 (32)	191 (41)

As a way to summarize the former results, the spatial mean of the relative RMSE values, obtained based on the results of the four validation ground locations, are shown in Figure 4. Particularly, the averages of the relative RMSE are displayed as a function of the forecasting lead time and season of the year. For the sake of comparison, the corresponding RMSE values of the trivial persistence model are also shown. This analysis, based on a spatial mean, aims to filter out local dependencies in the evaluation results, as microclimatic effects of the validation stations locations, providing general results for the study region. As expected from the results above, the WRF model forecast RMSE values shows a clear seasonal dependence over the study region. Particularly, the lowest values are found in summer (about 22% for the D1 forecasts); for the rest of the seasons similar values, but higher than during summer, are found (about 35% for the D1 forecasts). This leads to that, for the annual period, the relative RMSE value is about 30% for the D1 forecasts. Dependence on the forecast lead time is relatively low. For instance, for the annual period, RMSE values are just 32% and 34% for, respectively, the D2 and D3 forecasts. Unlike the WRF model, the persistence trivial model shows a considerable dependence on the lead time. In addition, this trivial model performs, in general, considerable worse than the WRF model, except for the D1 forecasts during winter and spring, when performance is similar. When analyzing the annual period scores, WRF performs about 5% better for the D1 forecasts, about 8% for D2 forecasts and about 10% for D3 forecasts.

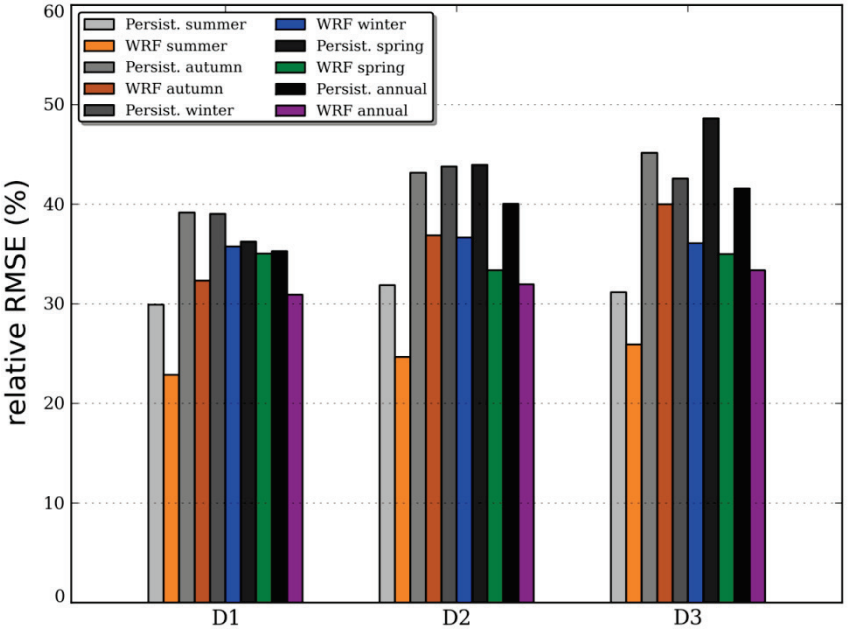

Figure 4. Relative RMSE values of the WRF model GHI forecasts in comparison to ground measured values. Values are the averaged of the four validation ground stations relative RMSE and are displayed for the different time horizon and season of the year. The corresponding relative RMSE values of the trivial persistence forecasting model are also shown.

3.3.2 DNI forecasts evaluation results: dependence on time horizon and seasonality

Tables 4 and 5 summarizes for, respectively, the Andasol and Córdoba ground validation stations, the performance of the WRF DNI forecasts against ground measurements, quantified by their absolute and relative RMSE and MBE values.

Table 4. As in Table 1 but for the DNI.

Forecast horizon	WRF MBE	WRF RMSE	Persistence RMSE
Summer			
Day 1	68 (13)	234 (46)	338 (66)
Day 2	82 (16)	247 (48)	349 (68)
Day 3	84 (17)	237 (45)	301 (57)
Autumn			
Day 1	118 (25)	311 (66)	399 (85)
Day 2	133 (28)	331 (71)	410 (87)
Day 3	145 (31)	348 (73)	415 (87)
Winter			
Day 1	201 (47)	364 (86)	359 (84)
Day 2	180 (41)	360 (81)	440 (99)
Day 3	163 (37)	353 (80)	425 (96)
Spring			
Day 1	75 (13)	308 (53)	360 (62)
Day 2	76 (13)	305 (53)	429 (74)
Day 3	78 (14)	333 (59)	460 (81)
Anual Period			
Day 1	109 (22)	304 (61)	364 (73)
Day 2	117 (24)	311 (62)	406 (81)
Day 3	119 (25)	319 (63)	403 (79)

As for the GHI, positive values of the MBE are found for the two validation stations, indicating that the WRF model tends to overestimate the DNI. The overestimation is present in all the seasons, but is of a considerable lower magnitude in summer compared to the rest of the seasons. Note that the MBE values are considerable higher for Córdoba station, particularly during autumn and spring, with values about twice than those at Andasol station. The MBE values are considerable higher for DNI than for GHI. For instance, for the annual

period analysis, at Andasol station and for the D1 forecasts, DNI MBE values are about twice higher than the corresponding GHI values (14% versus 22%; Tables 1 and 4). In the case of Córdoba station, differences are even higher (11% versus 35%; Tables 2 and 5). These differences between validation stations for the annual period are mainly associated with the differences during spring and autumn.

Table 5. As in Table 2 but for the DNI.

Forecast horizon	WRF MBE	WRF RMSE	Persistence RMSE
Summer			
Day 1	99 (16)	197 (33)	250 (42)
Day 2	103 (17)	204 (34)	299 (50)
Day 3	107 (18)	216 (36)	318 (54)
Autumn			
Day 1	255 (69)	362 (98)	290 (79)
Day 2	248 (68)	372 (102)	354 (102)
Day 3	251 (69)	368 (102)	387 (119)
Winter			
Day 1	232 (47)	370 (75)	368 (74)
Day 2	233 (47)	350 (69)	445 (84)
Day 3	273 (54)	393 (78)	419 (76)
Spring			
Day 1	148 (32)	264 (57)	276 (62)
Day 2	149 (32)	253 (54)	357 (82)
Day 3	150 (33)	258 (56)	439 (102)
Anual Period			
Day 1	172 (35)	294 (60)	292 (59)
Day 2	178 (36)	290 (58)	358 (73)
Day 3	182 (38)	305 (62)	386 (78)

The former results can be explained based, firstly, on the higher sensitivity of DNI to the cloudiness conditions, which enhances

differences between DNI and GHI forecasting MBE errors. Secondly, based on the location of the two validation stations. Córdoba station is located about 250 km to the west of Andasol station, in the middle of the Guadalquivir valley, open to the Atlantic Ocean and without important topographic features in between (Fig. 1). Therefore, the frontal systems coming from Atlantic Ocean leads to a relatively important cloudiness in this area. On the other side, these frontal systems, on their way to the location of Andasol station, encounter important topographic features that reduce the precipitation, and then the cloudiness, by orographic forcing.

As for the MBE, the RMSE shows a clear seasonal dependence, with lower values during summer for the two validation stations (Tables 4 and 5). Again, these seasonal differences can be explained based on the climate of the region, since cloudy conditions (which are much more common from autumn to spring than during the summer) are more difficult to be properly forecasted than clear-sky conditions. Note that, for the annual period and unlike the MBE, RMSE values are similar for both evaluation stations (about 60%). This result can be explained because the higher RMSE values for Córdoba station during autumn, associated with a greater cloudiness, is compensated by the higher RMSE values in Andasol during summer, associated with the presence of more cloudiness in this mountainous area during summer due to local convective activity. Similar results were reported by Wittman *et al*. (2008) in a forecast evaluation experiment in July at Andasol.

Persistence model DNI forecast RMSE values tends to be, as expected, higher than the corresponding WRF values. But considerable differences are found depending on the validation station, forecast horizon and season of the year. During summer, the WRF model performs considerable better, with RMSE values about one third lower at the two validation stations (Tables 4 and 5). For Andasol station, the WRF model performs considerable better during autumn and spring, although differences are lower than during the summer (Table 4). This leads to that, for the annual period, the WRF model performs about

15% better (in terms of the RMSE) than the trivial persistence model (relative RMSE 61% versus 73% for the D1 forecasts). On the other hand, for Córdoba ground station (Table 5), the trivial model performs slightly better for the D1 forecasts during autumn. For the D1 and D2, the WRF model tends to perform considerable better for all the season and validation stations. This is because the persistence model accuracy declines with the forecast lead time and it is considerable worse than that one of the WRF model.

As in Figure 4 for the GHI forecasts, Figure 5 shows the relative RMSE values, both for the WRF and the trivial model, averaged over the two validation ground stations. For the sake of comparison, the corresponding GHI values are also shown. As can be derived from the previous station-based analysis, the WRF model forecast RMSE values shows a clear seasonal dependence over the study region and a low dependence on the forecasting lead time. Particularly, the lowest RMSE values are found for summer (about 40% for the three lead times), followed by the spring (about 50% for the D1 and D2 forecasts and almost 60% for the D3). On the other hand, during autumn and winter, RMSE values are considerable higher (close to 80% for the two seasons and all lead times). Note that the WRF model only performs significantly better than the trivial model during summer and spring. During winter and autumn, the trivial model performs better. The rationale behind these results is the kind of sky conditions found for these seasons. Particularly, more than a half of the days presents cloudy conditions. As showed in section 3.3.4, these sky conditions are very stringent for the WRF model, leading to considerable high forecasting errors. As a consequence, for the annual period, the WRF model only performs slightly better than the trivial model. Particularly, for D1 forecasts, performance is very similar, while for D2 and D3 forecasts, the WRF forecasts RMSE values are about 5% better than the trivial model values (about 60% versus 65%).

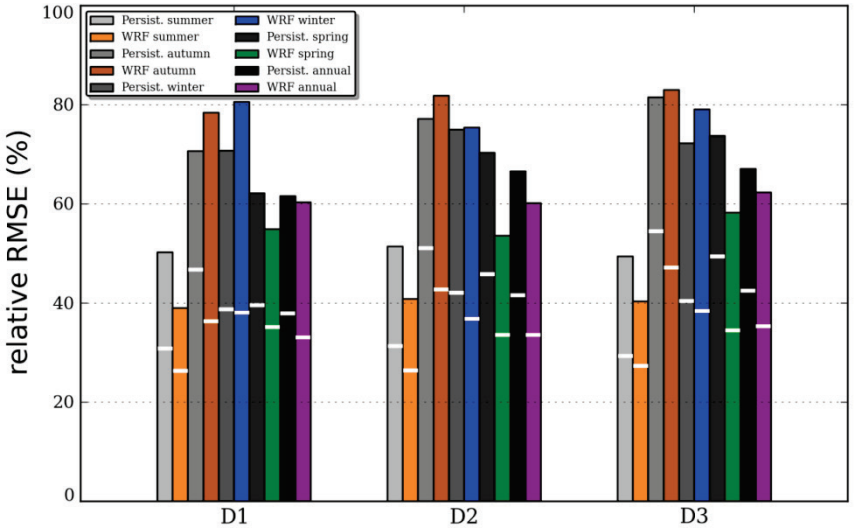

Figure 5. Relative RMSE values of the WRF model DNI forecasts in comparison to ground measured values. Values are the average of the two validation ground stations RMSE and are displayed for the different time horizon and season of the year. The corresponding relative RMSE values of the trivial persistence model and of the WRF GHI forecasts (horizontal lines), for the same two station, are also shown.

Compared to the GHI RMSE values, the corresponding DNI values are about twice higher except for the summer, when DNI errors are just one third higher. Note that the main difference between the GHI and DNI forecasting errors are found during autumn. For this period, the WRF model performs considerable better than the trivial persistence model for the GHI forecast while for DNI the performance of the trivial model is better.

3.3.3 GHI forecasts evaluation results: dependence on the sky conditions

Tables 6 to 8 summarize, for the validation locations of Andasol, Córdoba and Jerez, the performance of the WRF model GHI forecasts as a function of the sky conditions. Particularly, the different error values are presented as a function of the forecast horizon, for days D1

and D2, and the sky conditions for the different seasons of the year, including the annual mean errors.

Table 6. GHI forecast evaluation results as a function of the sky conditions, season of the year and forecast horizon, for days D1 and d2, for Andasol station. The WRF model MBE and RMSE forecasting values are showed in absolute (W/m2) and relative magnitude (in brackets at the right, in percentage).

	Forecast Horizon	Sky conditions	WRF MBE	WRF RMSE
Summer	Day 1	clear sky	40 (7)	113 (20)
		cloudy	110 (23)	149 (31)
		overcast	138 (48)	270 (94)
	Day 2	clear sky	21 (4)	147 (25)
		cloudy	142 (30)	188 (40)
		overcast	121 (42)	219 (76)
Autumun	Day 1	clear sky	4 (1)	49 (10)
		cloudy	36 (9)	166 (42)
		overcast	108 (60)	247 (138)
	Day 2	clear sky	-1 (0)	69 (14)
		cloudy	57 (14)	190 (48)
		overcast	142 (79)	277 (155)
Winter	Day 1	clear sky	24 (5)	56 (11)
		cloudy	67 (19)	192 (55)
		overcast	66 (31)	253 (121)
	Day 2	clear sky	12 (2)	72 (14)
		cloudy	66 (19)	216 (61)
		overcast	48 (23)	248 (119)
Spring	Day 1	clear sky	18 (3)	68 (11)
		cloudy	153 (34)	241 (53)
		overcast	267 (99)	358 (133)
	Day 2	clear sky	19 (3)	75 (12)
		cloudy	125 (28)	230 (51)

		overcast	202 (75)	329 (122)
Annual	Day 1	clear sky	23 (4)	79 (14)
		cloudy	83 (20)	191 (45)
		overcast	148 (62)	290 (121)
	Day 2	clear sky	15 (3)	98 (17)
		cloudy	96 (23)	202 (48)
		overcast	127 (54)	278 (118)

Table 7. As in Table 6 but for the Córdoba station.

	Forecast Horizon	Sky conditions	WRF MBE	WRF RMSE
Summer	Day 1	clear sky	13 (2)	33 (5)
		cloudy	109 (23)	205 (44)
		overcast	- -	- -
	Day 2	clear sky	11 (2)	34 (6)
		cloudy	111 (23)	193 (41)
		overcast	- -	- -
Autumun	Day 1	clear sky	5 (1)	43 (9)
		cloudy	44 (11)	112 (27)
		overcast	112 (49)	230 (100)
	Day 2	clear sky	5 (1)	43 (9)
		cloudy	44 (11)	121 (30)
		overcast	178 (77)	304 (132)
Winter	Day 1	clear sky	3 (1)	65 (13)
		cloudy	74 (19)	137 (34)
		overcast	210 (91)	270 (117)
	Day 2	clear sky	5 (1)	49 (10)
		cloudy	58 (14)	130 (32)
		overcast	148 (64)	230 (99)
Spring	Day 1	clear sky	10 (2)	61 (10)
		cloudy	110 (24)	235 (52)

		overcast	114 (73)	205 (132)
	Day 2	clear sky	9 (2)	84 (14)
		cloudy	84 (19)	231 (51)
		overcast	44 (28)	168 (108)
Annual	Day 1	clear sky	10 (2)	50 (9)
		cloudy	79 (18)	173 (40)
		overcast	139 (70)	231 (116)
	Day 2	clear sky	10 (2)	56 (10)
		cloudy	72 (17)	169 (39)
		overcast	112 (56)	233 (117)

As expected, a marked dependence of the forecasting error on the sky conditions was found for all the validation stations and seasons of the year. Particularly, at Córdoba validation station (Table 7), for the annual period and for the D1 forecasts, the MBE values are 2% for clear-sky-conditions, 18% for cloudy conditions and 70% for completely overcast conditions. Similar values are found for Andasol station (Table 6), Huelva station (not shown) and Jerez station (Table 8). This indicates that the WRF model tends to overestimate the GHI and this overestimation is strongly dependent on the sky conditions. Since the MBE increases as the cloud fraction increases, the overestimation can be related to limited ability of the WRF model to properly forecast cloudy conditions, forecasting more clear-sky-conditions than actually occurred. This is in agreement with the results of the previous sections 3.3.1, where higher MBE values where found during the rainy (cloudy) seasons in the study area. This systematic bias can partially be avoided by the introduction, not attempted in this work, of a cloud-related specific bias correction as in Lorenz *et al*. (2009a).

Table 8. As in Table 6 but for the Jerez station.

Forecast Horizon	Sky conditions	WRF MBE	WRF RMSE

Summer	Day 1	clear sky	27 (5)	55 (9)
		cloudy	65 (13)	145 (28)
		overcast	135 (65)	244 (88)
	Day 2	clear sky	19 (3)	74 (13)
		cloudy	93 (18)	157 (31)
		overcast	137 (67)	313 (113)
Autumun	Day 1	clear sky	7 (1)	68 (15)
		cloudy	29 (7)	125 (29)
		overcast	131 (54)	227 (93)
	Day 2	clear sky	3 (1)	78 (17)
		cloudy	2 (0)	136 (32)
		overcast	132 (54)	221 (91)
Winter	Day 1	clear sky	0 (0)	67 (13)
		cloudy	28 (7)	135 (35)
		overcast	120 (86)	266 (191)
	Day 2	clear sky	-4 (-1)	90 (17)
		cloudy	54 (14)	125 (32)
		overcast	166 (119)	291 (209)
Spring	Day 1	clear sky	29 (5)	55 (9)
		cloudy	152 (35)	213 (49)
		overcast	274 (171)	371 (232)
	Day 2	clear sky	0 (0)	142 (24)
		cloudy	120 (28)	195 (45)
		overcast	111 (69)	227 (142)
Annual	Day 1	clear sky	20 (4)	59 (11)
		cloudy	57 (13)	150 (35)
		overcast	146 (75)	291 (149)
	Day 2	clear sky	7 (1)	103 (18)
		cloudy	54 (12)	150 (35)
		overcast	107 (55)	259 (133)

Differences between sky conditions vary by a factor of more than 10 in terms of the relative RMSE. For instance, at Córdoba validation station (Table 7), for the annual period and for the D1 forecasts, relative RMSE values are 9% for clear sky conditions, 40% for cloudy conditions and 116% for completely overcast conditions. Similar values are found for the rest of the validation stations. Regarding the season of the year, summer presents the lowest RMSE values for clear-sky conditions, except for Andasol station, which maximum values are found during this season. Figure 6 shows the three-days-ahead GHI and DNI forecast, along the period February 28th to March 1st, for Andasol station. Note that during the first day of forecast (D1), the model was able to properly forecast the steep change in GHI, caused by clouds, that ocurred before solar noon. Nevertheles, along the second day of forecast (D2), the model failed the forecast of clouds during the morning and evening. In this second day of forecast, the RMSE values for the trivial model and for the WRF model are similar.

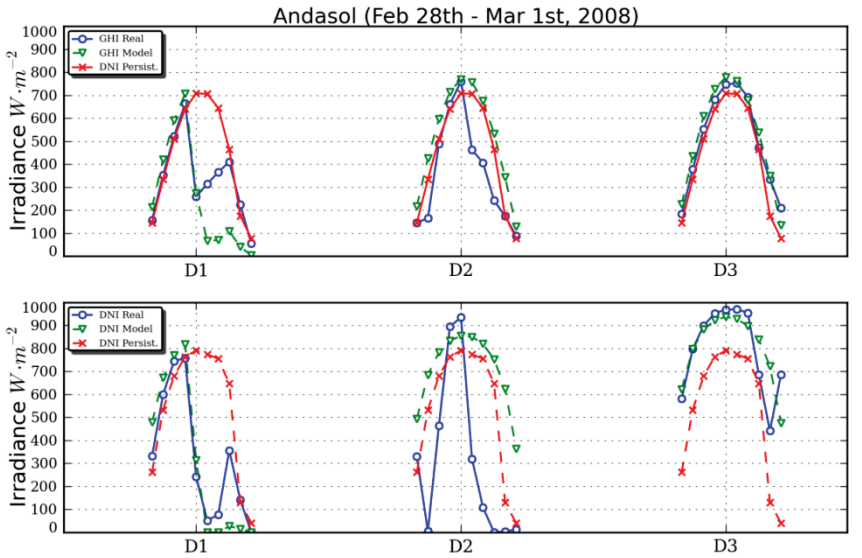

Figure 6. WRF model three-days-ahead GHI (above) and DNI (below) forecasts for the Andasol station along the period February 28th to March 1st. The trivial model forecasts and the observed values are also displayed.

As in the case of WRF, the persistence model accuracy strongly depends on the sky conditions. Particularly, the performance substantially decreases as the cloudiness increases for all the seasons and validation stations. For instance, at Córdoba station (Table 7), for the annual period and D1 forecasts, persistence model forecasts RMSE values are 20% for clear sky conditions, 43% for cloudy conditions and 146% for complete overcast conditions. Similar values are found for the rest of the validation stations. Compared to the trivial model, and for D1 forecasts, performance of WRF is substantially better for clear sky conditions, similar for cloudy conditions and slightly better for overcast conditions. For instance, for Córdoba station (Table 7), for the annual period and D1 forecasts, RMSE values for clear sky conditions are 9% for the WRF model and 20% for the persistence model. For cloudy conditions, RMSE values are similar, around 40%. However, under overcast skies, RMSE values are 116% and 146% for WRF and the persistence model, respectively. For the D2 and D3 (not shown) forecast, performance of the trivial model substantially decreases compared to the WRF model for clear sky and overcast conditions. For cloudy conditions, nevertheless, performance keeps similar.

As a way to summarize the former results, Figure 7 shows the spatial mean of the relative RMSE values for the D1 forecast as a function of the clear-sky index kt. Values were obtained based on the results of the four validation ground locations. The WRF forecast accuracy for the whole region shows a marked dependence on the sky conditions. For the annual period, values range from below 10% for clear sky (kt~0.8), about 50% for cloudy conditions (kt ~0.5) and up to more than 100% for overcast conditions (kt<0.3). For cloudy and overcast conditions there is a clear seasonal dependence, with higher values during spring and lower values during autumn, while for clear sky (kt>0.7) performance is very similar for the different seasons of the year. Compared to the trivial model, the WRF model performs considerable better for clear skies (kt>0.7): for all the seasons, persistence model RMSE values are about twice higher than those of the WRF model. For overcast conditions (kt<0.4) the WRF model also performs better, except during spring and winter, when both models

perform similarly. Finally, for cloudy conditions (0.3<kt<0.6), performance is similar except in autumn, when WRF performs better.

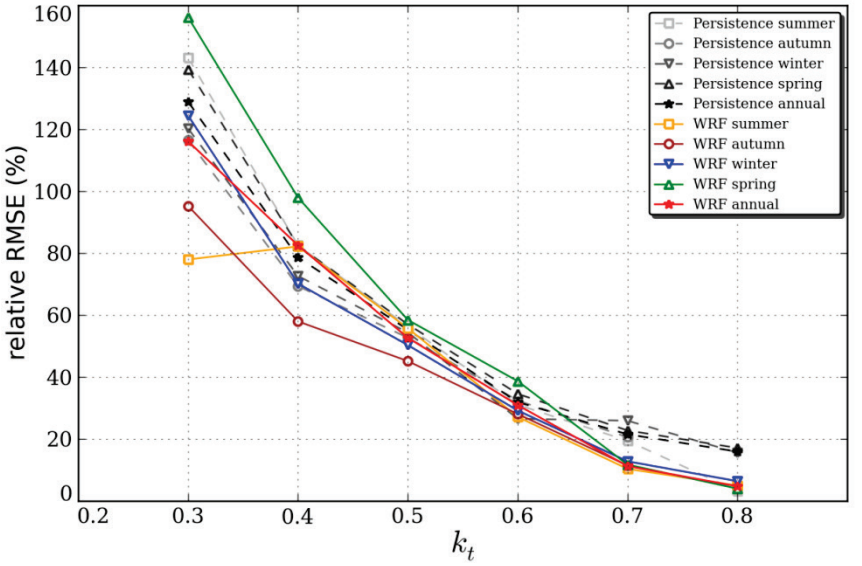

Figure 7. Relative RMSE values of the WRF model GHI one-day-ahead forecast in comparison to ground measured values. Values are the average of the four validation ground station results for all the seasons and are displayed as a function of the clear sky index. The corresponding relative RMSE values of the trivial persistence model are also shown.

The systematic overestimation of the GHI during clear sky conditions for all the validation stations can be explained on the basis of the solar irradiance parameterization of the WRF model used in this work, the Dhudia scheme (Dudhia, 1989). This parameterization allows solar radiation varying with cloud amount and composition, humidity and sun's zenith angle. Additionally, to account for aerosol and other scattering effects, a climatic clear-sky scattering factor, equivalent to a 0.1 total aerosol optical depth (AOD), is assumed by the model (Zamora *et al*., 2003). In addition, this solar radiation scheme neglects stratospheric ozone absorption. During summer, AOD (monthly mean) measured by the MODIS sensor is 0.24 and 0.36 at, respectively, Córdoba and Andasol stations. Therefore, the scattering factor used in the Dudhia model seems to be too low to account for the aerosol effect

at the study region, leading to an overestimation of the GHI, particularly significant at Andasol station in summer. Similar conclusions were provided by (Zamora *et al*., 2005) in an evaluation study of the WRF GHI forecasts for central and southern U.S. Particularly, they showed that when AOD doubles the value considered in the Dudhia scheme, this radiation scheme overestimate the GHI by ~100 W/m^2 at solar noon. Figure 8 shows, for Córdoba and Andasol stations, the GHI and DNI one-day-ahead forecasts (D1) during two clear-sky days: 20th and 21th of August 2007. Note that, for GHI and for Andasol station, at solar noon, a similar overestimation as those reported by (Zamora *et al*., 2005) is observed. For the rest of the seasons of the year, similar conclusion, regarding the AOD effect on the results under clear sky conditions, can be derived.

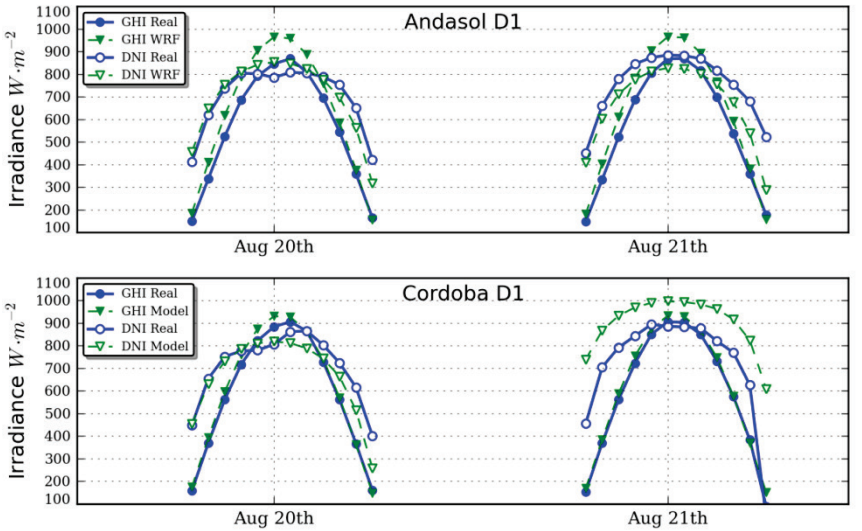

Figure 8. WRF model GHI and DNI one-day-ahead forecasts corresponding to the day 20th and 21th of August 2007. Above, for the Andasol station, below, for the Córdoba station. Observed values are also displayed.

3.3.4 DNI forecasts evaluation results: dependence on the sky conditions

Tables 9 and 10 summarizes for, respectively, the Andasol and Córdoba ground stations, the performance of the WRF DNI forecasts as a function of the sky conditions. As for GHI (Tables 6 to 8), a marked dependence on the sky conditions is found for all the validation stations and seasons. Particularly, at the Córdoba validation station (Table 10), for the annual period and D1 forecasts, MBE values are about 10% for clear sky, 75% for cloudy conditions and 507% for overcast conditions. Similar values are found for other seasons and lead times. On the other hand, Andasol validation station (Table 9) shows negative MBE during clear sky conditions along the whole year except in winter, leading to a MBE value of -4% for the annual analysis. During cloudy and overcast conditions, MBE values are similar to those found for Córdoba station. Note that positive MBE values during cloudy and overcast conditions indicates, as for GHI, the limited ability of the WRF model to forecast cloud conditions in the study area. On the other hand, the positive and negative bias found during clear sky conditions can be explained on the base of the proposed model to obtain the DNI (see section 3.2.3). Particularly, the negative bias found for Andasol station seems to be mainly caused by an overestimation of the aerosol load. Note that, under clear sky conditions and during the summer, the effect of the aerosol on DNI variability tends to be the dominating one. On the other hand, for Córdoba station, the positive bias observed for all the stations seems to be associated with an underestimation of the aerosol load at this location. Based on the results of the previous section regarding the GHI, it seems that, during the summer, the actual AOD value at Andasol station is higher than 0.1 but lower than those provided by the MODIS platform (0.36). On the other hand, for Córdoba station, AOD values seem to be slightly higher than those provided by the MODIS (0.24 for the summer). This can be observed in Figure 8, where, at solar noon, GHI is clearly overestimated at Andasol and slightly overestimated at Córdoba. On the other hand, DNI is underestimated at Andasol and overestimated at Córdoba.

Table 9. As in Table 6 but for the DNI.

	Forecast Horizon	Sky conditions	WRF MBE	WRF RMSE
Summer	Day 1	clear sky	-37 (-5)	150 (22)
		cloudy	150 (37)	281 (69)
		overcast	164 (98)	299 (179)
	Day 2	clear sky	-60 (-9)	185 (27)
		cloudy	197 (50)	284 (71)
		overcast	203 (161)	318 (252)
Autumun	Day 1	clear sky	-21 (-3)	86 (12)
		cloudy	154 (39)	351 (90)
		overcast	368 (493)	489 (654)
	Day 2	clear sky	-48 (-7)	145 (20)
		cloudy	188 (48)	370 (94)
		overcast	432 (577)	512 (686)
Winter	Day 1	clear sky	21 (3)	152 (20)
		cloudy	330 (130)	454 (179)
		overcast	297 (525)	453 (801)
	Day 2	clear sky	-1 (0)	141 (18)
		cloudy	324 (130)	461 (185)
		overcast	307 (543)	465 (822)
Spring	Day 1	clear sky	-53 (-7)	185 (24)
		cloudy	251 (66)	434 (114)
		overcast	300 (230)	423 (324)
	Day 2	clear sky	-49 (-6)	198 (26)
		cloudy	197 (52)	433 (114)
		overcast	226 (173)	385 (294)
Annual	Day 1	clear sky	-28 (-4)	156 (21)
		cloudy	207 (57)	373 (103)
		overcast	280 (262)	421 (394)
	Day 2	clear sky	-42 (-6)	175 (24)
		cloudy	220 (60)	378 (104)

		overcast	289 (305)	430 (454)

Concerning the RMSE, differences between sky conditions are higher than for the GHI forecasts, indicating the higher sensitivity of the DNI to the cloudiness. Particularly, at Andasol station, for the annual period and D1 forecasts (Table 9), RMSE values range from 21% for clear sky conditions, to 103% for cloudy conditions and up to 394% for overcast conditions. The respective values for GHI at Andasol station are (Table 6), 14%, 45% and 121%. Very similar values are found for Córdoba station (Table 10). Note in Figure 6 (D2 forecasts) the higher sensitivity of DNI to the presence of clouds, compared to GHI. This yields considerable higher RMSE values when cloudy conditions are not properly forecasted. Regarding the season of the year, RMSE values tend to be considerable higher during winter for Andasol and during spring for Córdoba. Finally, differences in the RMSE values between forecasting lead times tend to be lower.

Table 10. As in Table 7 but for the DNI.

	Forecast Horizon	Sky conditions	WRF MBE	WRF RMSE
Summer	Day 1	clear sky	49 (7)	117 (17)
		cloudy	243 (73)	336 (101)
		overcast	--	--
	Day 2	clear sky	47 (7)	117 (17)
		cloudy	243 (70)	332 (96)
		overcast	--	--
Autumun	Day 1	clear sky	75 (12)	139 (22)
		cloudy	255 (67)	354 (92)
		overcast	399 (482)	495 (599)
	Day 2	clear sky	66 (10)	138 (21)
		cloudy	242 (64)	359 (95)
		overcast	418 (505)	520 (629)
Winter	Day 1	clear sky	75 (10)	178 (25)
		cloudy	361 (110)	444 (136)
		overcast	467 (446)	597 (571)

	Day 2	clear sky	97 (13)	175 (24)
		cloudy	347 (106)	426 (130)
		overcast	422 (403)	569 (544)
Spring	Day 1	clear sky	111 (16)	181 (26)
		cloudy	212 (66)	362 (112)
		overcast	156 (765)	288 (1415)
	Day 2	clear sky	115 (16)	185 (26)
		cloudy	183 (57)	358 (111)
		overcast	91 (448)	227 (1118)
Annual	Day 1	clear sky	73 (10)	151 (22)
		cloudy	264 (75)	370 (106)
		overcast	311 (507)	451 (736)
	Day 2	clear sky	78 (11)	153 (22)
		cloudy	250 (72)	366 (105)
		overcast	277 (451)	435 (710)

Performance of the persistence model (not shown) substantially decreased as the cloudiness increased. As for GHI, the WRF model performs, overall, considerable better for clear skies and similarly for cloudy conditions. Particularly, for the former, the RMSE of the WRF model are about half of the RMSE values of the persistence model. For overcast conditions, considerable differences in the performance are found depending on the season and validation station. For instance, for Andasol station, WRF performs considerable better for all the seasons except for the winter. Particularly, for the annual period and D1 forecasts, WRF model RMSE value is 394% while the persistence RMSE value is 439%. At Córdoba station, WRF performs considerable better except for the autumn.

As a way to summarize the former results, Figure 9 shows the spatial mean of the relative RMSE values for the D1 forecasts as a function of the clear-sky index kt. Values were obtained based on the results of the two validation ground locations. The WRF forecast accuracy for the whole region shows a marked dependence on the sky

conditions. This dependence is even higher than for the case of GHI (Figure 7). For the annual period, relative RMSE values range from below 20% for clear skies (kt~0.8), to above 100% for cloudy conditions (kt ~0.5) and up to more than 400% for overcast conditions (kt<3). This is about twice the corresponding RMSE values for GHI during clear and cloudy conditions and more than 3 times the error for overcast conditions. Note that, unlike for GHI, when the maximum errors during cloudy and overcast conditions are found in spring, maximum errors for the DNI are found during winter.

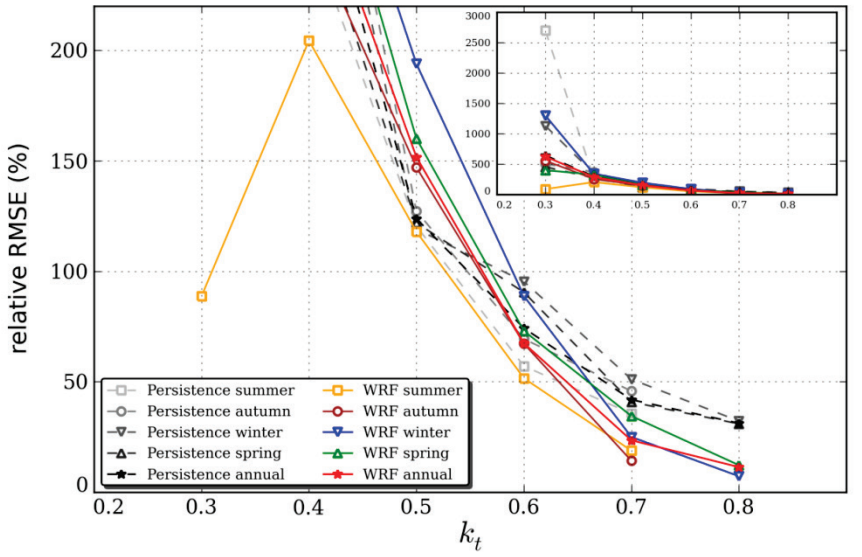

Figure 9. As in figure 7 but for the DNI forecasts. Only two ground stations (Córdoba and Andasol) were considered, compared to the four stations (Córdoba, Andasol, Jerez and Huelva) evaluated for the GHI. The small figure (top right hand corner) represents the same values but using a different vertical scale.

Similarly, minimum errors are found during summer for the DNI, while for GHI they are minimal in autumn. Finally, WRF performs considerable better than the persistence model for clear (kt>0.7) and overcast conditions (kt<0.4). For cloudy conditions, things are more complicated. Particularly, for kt~0.6, performance is similar while for kt~0.5 the persistence model performs better.

3.4 Summary and conclusions

In this work, an evaluation of the reliability of three-days-ahead GHI and DNI forecasts provided by the WRF mesoscale atmospheric model for Andalusia (southern Spain) was conducted. GHI forecasts were provided directly by the model, while DNI forecasts were obtained based on a physical post processing based on the WRF outputs and satellite retrievals for aerosols and ozone. Hourly time resolution and 3 km spatial resolution estimates of the WRF model were tested, in terms of the MBE and RMSE, against ground measurements collected at four radiometric stations along the years 2007 and 2008. Two different analyses were carried out for GHI and DNI forecasts. In the first one, these forecasts were evaluated independently for the different seasons of the year, without distinction of the sky conditions. In the second analysis, the evaluation was carried out on the light of three different sky conditions: clear, cloudy and overcast. In both cases, three different leading times were considered: 24, 48 and 72 hours.

The first analysis (seasonal analysis) showed that WRF tends to overestimate GHI for all the seasons of the year (MBE about 10%) but with lower MBE in summer (about 5%). The relative RMSE also showed a clear seasonal dependence with values ranging from about 20% during summer to about 35% for the rest of the seasons for 24 hours ahead forecast. As expected, the forecasts accuracy decreased with the forecast lead time. Nevertheless, this decrease was modest except in autumn. The trivial persistence model performed considerable worse than WRF, except for 24 hours ahead forecast during winter and spring, when performance was similar. Regarding DNI, MBE and RMSE values were about twice higher than the corresponding GHI values, except for the RMSE in summer, when they were just one third higher. Particularly, the lowest relative RMSE values were found in summer (about 40%), followed by the spring (about 55%), both for the 24 hours lead time. On the other hand, during the rest of the seasons, RMSE values were considerable higher. The WRF model only performed significantly better than the trivial model during summer and spring. During winter and autumn, the trivial model performed better.

The most significant differences between the GHI and DNI forecasting errors were found in autumn. For this period, WRF performed considerable better than the trivial persistence model for the GHI forecast while for the DNI performance of the trivial model was better.

The second analysis (sky-conditions-dependent analysis) showed, as expected, a marked dependence of the forecasting error on the sky conditions for all the seasons and both for the MBE and the RMSE. Particularly, for 24 hours lead time, the MBE values were 2% for clear-sky conditions, 18% for cloudy conditions and 54% for overcast conditions. This indicates that the WRF model tends to overestimate the GHI for all sky conditions. Since the MBE increases as the cloudiness increases, this overestimation can be related to the limited ability of WRF to properly forecast cloudy conditions, forecasting more clear-sky-conditions than actually occurred. The systematic overestimation of the GHI during clear sky conditions was explained on the basis of the WRF's solar irradiance parameterization used in this work, which accounts for a lower AOD than the real in the study area. Regarding the RMSE values, differences between sky conditions varied by a factor of more than 10. For instance, for the annual analysis and for 24 hours lead times, relative RMSE values ranged from below 10% for clear skies, to about 50% for cloudy conditions and up to more than 100% for overcast conditions. This is because the model showed to be able to accurately forecast steep changes in the sky (cloudiness) conditions. Nevertheless, in many cases, the existence of clouds was not forecasted, leading to increasing RMSE forecasting errors with increasing cloudiness. As far as DNI is concerned, the MBE values showed a marked dependence on the sky conditions ranging, for the Córdoba validation station and for 24 hours lead time, from about 10% for clear sky, to 75% and 507% for cloudy and overcast conditions, respectively. Concerning the RMSE, differences between sky conditions were higher than for the GHI forecasts, indicating the higher sensitivity of the DNI to the cloud conditions. Particularly, for the annual period and 24 hours lead times, relative RMSE values ranged from below 20% during clear sky (kt~0.8), to above 100% for cloudy conditions (kt~0.5) and to more than 400% for overcast conditions

(kt<3). This is about twice the corresponding RMSE values for the GHI during clear and cloudy conditions and more than 3 times the error for overcast conditions. Differences in the RMSE values between forecasting lead times were of low magnitude. Finally, the WRF model performed considerable better than the persistence for clear sky and overcast conditions. For cloudy conditions, the performance showed to be similar.

In many cases, particularly under cloud and overcast conditions, the two days and three days ahead forecast errors are similar (an even in some cases better) than the one day ahead errors. From a meteorological perspective, quality of the forecasts descends when increasing the lead time. Therefore, these results are probably related with the relative short (four months along the year) evaluation period. This can also explain why the persistence model performance is close to the WRF model performance at some test periods (especially in autumn).

The former results are indicative of the region as a whole and were obtained by averaging the results of the four analysed evaluation station. Nevertheless, both the seasonal and the sky-conditions-dependent analyses showed considerable differences depending on the validation station, especially for the DNI forecast. Mainly, because the region presents a wide range of weather and climate conditions. Particularly, forecasting errors for the stations located at the west of the region were, on the whole, lower than the corresponding located at the east.

The usefulness of the WRF model DNI forecast regarding the Spanish Solar Electricity premium feed-in tariff model is difficult to assess based solely on the results here presented. On the one hand, it is clear that during clear sky conditions, the WRF model is a valuable tool to participate in the electricity market, performing better than the persistence models. Nevertheless, cloud forecasts is still a big issue for the WRF model and considerable improvements should be obtained before reliable DNI and GHI forecasts are obtained during cloudy and overcast conditions. Therefore, from an economic perspective, the

balancing between the negative effects of the unreliable forecasts under cloudy conditions and the positive effect of reliable forecasts should be analyzed in a case study, as those of Wittmann *et al*. (2008). This will be attempted in a future work. Finally, it should be highlighted that none statistical post processing to improve the forecast accuracy was carried out upon the WRF output. Future research will be focussed, also, on this issue. Particularly, following the methodology proposed in Ruiz-Arias *et al*. (2010a), the eventual improvement in the forecasting RMSE and MBE values obtained by spatially averaging the model output will be analysed. In addition, a correction of the systematic errors (bias correction) will be also attempted.

Acknowledgements

The Spanish Ministry of Science and Technology (Project ENE2007-67849-C02-01) and the Andalusian Ministry of Science and Technology (Project P07-RNM-02872) financed this study. Authors would like to thank AEMET (Spanish Meteorological Service) and Milenio Solar (Solar Millenium AG) for providing the ground observational measurements. Authors also would like to thank the Informatic and Scientific Center of Andalusia (CICA), mainly to Marceliano Marrón, for helping with computational aspects. We are in debt with Dr. Elke Lorenz for the interesting and fruitful discussions regarding this work during Vicente Lara stay at the Meterology and Energy group of the University of Oldenburg (Germany). Finally, the authors would like to acknowledge the very helpful comments made by three anonymous reviewers, which helped to substantially improve the manuscript.

Chapter 4

Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe

---

Perez R, Lorenz E, Pelland S, Beauharnois M, Van Knowe G, Karl Hemker Jr., Detlev Heinemann, Jan Remund, Stefan C. Müller, Wolfgang Traunmüller, Gerald Steinmauer, David Pozo, Jose A. Ruiz-Arias, ***Vicente Lara-Fanego****, Lourdes Ramirez-Santigosa, Martin Gaston-Romero, Luis M. Pomares (2013). Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. Sol Energy 94:305–326.*

Citations in WEB of Science: 59

Impact factor: 3.685

4.1 Introduction

Solar power generation is highly variable due its dependence on meteorological conditions. The integration of this fluctuating resource into the energy supply system requires reliable forecasts of the expected power production as a basis for management and operation strategies. During the last years the contribution of solar power to the electricity supply has been increasing fast leading to a strong need for accurate solar power predictions (in Germany, for instance, the PV production already exceeds 40% of electrical demand on sunny summer days).

Following this new and rapidly evolving situation on the energy market, substantial research effort is currently being spent on the development of irradiance and solar power prediction models, and

several models have been proposed recently by research organizations as well as by private companies. Common operational approaches to short-term solar radiation forecasting include (1) numerical weather prediction (NWP) models that infer local cloud information – hence, indirectly, transmitted radiation – through the dynamic modeling of the atmosphere up to several days ahead (e.g., see Remund *et al*., 2008); (2) models using satellite remote sensing or ground based sky measurements to infer the motion of clouds and project their impact in the future. Earlier contributions by some of the authors have shown that satellite-derived cloud motion tends to outperform NWP models for forecast horizons up to 4–5 h ahead depending on location (e.g., Perez *et al*., 2010; Heinemann *et al*., 2006b). Short-term forecasting using ground-based sky imagery with very high spatial and temporal resolution is suitable for intra-hour forecasting (Chow *et al*., 2011); (3) statistical time series models based on measured irradiance data are applied for very short term forecasting in the range of minutes to hours (e.g., see Pedro and Coimbra, 2012). In this paper we focus our attention on solar radiation forecasts based on NWP models which are most appropriate for day-ahead and multi-day forecast horizons. Day-ahead predictions are of particular importance for application in the energy market, where day-ahead power trading plays a major role in many countries.

This article combines and discusses three independent validations of global horizontal irradiance (GHI) multi-day forecast models that were performed in the US, Canada and Europe in the framework of the IEA SHC Task 36 "Solar resource knowledge management" (http://archive.iea-shc.org/task36/). Comparing the performance of different models gives valuable information both to researchers, to rank their approaches and inform further model development, and to forecast users, to assist them in choosing between different forecasting products. It is important that a standardized methodology for evaluation is used for the comparison in order to achieve meaningful results when comparing different approaches. Therefore, a common benchmarking procedure has been set up in the framework of the IEA SHC Task 36. As a basis for the benchmarking we have prepared several ground

measurement data sets covering different climatic regions and a common set of accuracy measures has been identified.

The paper first gives an overview of the different forecasting approaches. Then we present the ground measurement datasets used for the validation. Next, the concept of evaluation is introduced, and finally, we provide the results of the forecast comparison along with a short discussion and conclusions.

4.2. Forecast models

The evaluation includes forecasts based on global, multiscale and mesoscale NWP models. Hourly site-specific forecasts are derived from direct NWP model output with different methods ranging from simple averaging and interpolation techniques to advanced statistical postprocessing tools and meteorologists' interpretation to combine the output of various NWP models. The models considered for this evaluation are listed below, along with the acronyms that will be used to present results:

1. The Global Environmental Multiscale (GEM) model from Environment Canada in its regional deterministic configuration (Mailhot *et al*., 2006).
2. An application of the European Centre for Medium-Range Weather Forecasts (ECMWF) model (Lorenz *et al*., 2009a, b).
3. Several versions of the Weather Research and Forecasting (WRF) model (Skamarock *et al*., 2005, 2008) initialized with Global Forecast System (GFS) forecasts (GFS, 2010) from the US National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Prediction (NCEP).
    - WRF-ASRC, a version used as part of an operational air quality forecasting program at the Atmospheric Sciences Research Center of the University of Albany (AQFMS, 2010).

• WRF-AWS, a version of WRF operated at AWS Truepower in the US.
• WRF-Meteotest, a version of WRF operated at Meteotest in Europe.
• WRF-UJAEN, a version operated at the University of Jaén, Spain (Lara-Fanego *et al*., 2012).

4. The MASS model (Manobianco *et al*., 1996).
5. The Advanced Multiscale Regional Prediction System (ARPS) model (Xue *et al*., 2001).
6. The regional weather forecasting system Skiron (Kallos, 1997) operated and combined with statistical postprocessing based on learning machines at Spain's National Renewable Energy Center (CENER). (Skiron-CENER, Gastón *et al*., 2009).
7. The High Resolution Limited Area Model (HIR- LAM, 2010) operational model from the Spanish weather service (AEMet) combined with a statistical postprocessing at CIEMAT (HIRLAM-Ciemat).
8. A model based on cloud cover predictions from the US National Digital Forecast Database, (NDFD) proposed by Perez *et al*. (2010).
9. BLUE FORECAST: statistical forecast tool of Bluesky based on the GFS predictions from NCEP.
10. Forecasts based on meteorologists' cloud cover fore- casts by Bluesky (BLUESKY-meteorologists).

The first two models are directly based on global (planetary) NWP systems, respectively GEM, and ECMWF.

The native time step of the regional configuration of the GEM model and its ground resolution are 7.5 min and rv15 km, respectively. GEM forecasts of downward short- wave radiation flux at the surface (DSWRF) originating at 00:00Z and 12:00Z were de-archived by the Canadian Meteorological Centre at an hourly time step for this analysis. The de-archived forecasts cover North America and adjacent waters. As described by Pelland *et al*. (2011), the GEM solar forecasts were postprocessed by taking an average of the irradiance forecasts over a

square region centered on the location of each site used in the validation. The size of this square grid was optimized for each station by selecting a size that minimized forecast root mean square error during a 1 year training period prior to the evaluation period used here.

ECMWF irradiance forecasts used here had a temporal resolution of 3 h and a spatial resolution of 25 km. The ECMWF site-specific, hourly data prepared for the present analysis according to Lorenz *et al*. (2009b) are obtained via time interpolation of the 3-hourly global clear sky indices. In addition, a bias correction that is dependent upon the cloud situation was performed for the European sites. This postprocessing was based on historic ground measured irradiance values for Germany and Switzerland, and on satellite derived irradiance values for Spain and Austria. For the US and Canadian sites no additional training to ground measured data was applied.

Models 3–7 are mesoscale models that use global weather models as an input to define regional boundary conditions, but add high resolution terrain and other features to produce higher resolution forecasts. In all cases analyzed here, the global weather model input is NOAA's GFS model. The GFS model dataset used for this project has a time resolution of 3 h and a nominal ground resolution of one by one degree (i.e., rv80 x 100 km in the considered latitude range). All the mesoscale models produce hourly output.

The WRF version of the model run by AWS Truepower as well as the MASS and ARPS models have a final ground resolution of 5 km. They are tested in two operational modes: with and without Model Output Statistics (MOS) postprocessing. The MOS process consists of integrating ongoing local irradiance measurements, when available, to correct localized errors from the numerical weather prediction process. This is a common operational forecasting practice: taking advantage of ongoing local surface and upper air measurements to deliver better forecasts.

The Advanced Research WRF model currently used in operational forecasting at the Atmospheric Sciences Research Center (WRF-

ASRC) is a next-generation mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. It features multiple dynamical cores, a 3-dimensional variational (3DVAR) data assimilation system, and a software architecture allowing for computational parallelism and system extensibility. The operational version of this WRF model is version 3.2.1 and is run at a horizontal grid resolution of 12 km for domain encompassing the eastern section of the United States and Canada.

The two applications of WRF for Europe (Meteotest and U. Jaén) do not integrate postprocessing with ground measured values. WRF forecasts processed by the University of Jaén for a study region in Andalusia show a final spatial resolution of 3 km. The choice of the different parameterizations was based on a calibrating experiment for MM5, a former version of the WRF model, carried out for an optimum adjustment for the study region by Ruiz-Arias *et al*. (2008). WRF forecasts at Meteotest for Central Europe are processed with a grid size of 5 km x 5 km for the innermost domain. The forecasts are averaged using 10 x 10 model pixels around the point of interest corresponding to an area of 50 x 50 km.

Models 6 and 7 apply a postprocessing procedure to predictions of a mesoscale NWP model. CENER's solar global irradiance prediction scheme (model 6) is based on the regional weather forecasting system Skiron (Kallos, 1997), developed at the Hellenic National Meteorological Service, and operated with a final spatial resolution of 0.1o x 0.1o. The applied statistical postprocess is based on learning machines (Gastón *et al*., 2009). CIEMAT applies a bias correction to forecasts of the HIRLAM operational model of the Spanish weather service (AEMet) with a spatial resolution of 20 km x 20 km.

The statistical forecast tool BLUE FORECAST (model 9) is also based on the global GFS model. The original GFS forecasts with temporal resolutions of 3 and 6 h and spatial resolutions of 1o x 1o and 0.5o x 0.5o are integrated into a statistical postprocessing procedure using different methods of data mining such as ridge regression,

automatic quadratic models or neural networks, based on meteorological inputs (see Natschläger *et al*., 2008).

The NDFD forecast does not provide irradiance per se, but cloud amount that extends up to 7 days ahead with a ground resolution of rv5 km and a time resolution of 3 h up to 3 days ahead and 6 h beyond that. The NDFD is also based on the GFS global model. GFS forecasts are individually processed by regional NOAA offices using mesoscale models and local observations and gridded nationally into the NDFD. The forecast cloud amounts are modeled into irradiance using an approach developed by Perez *et al*. (2010).

A similar approach is also operated by Bluesky for model 10. The meteorologists in the operational weather service use the results of several meteorological forecast models and combine these with their meteorological knowledge and forecasting experience. The result is cloud cover forecasts with hourly resolution in a first step. These are converted to solar irradiance forecasts using an equation including the cloud cover coefficient and clear sky irradiances.

All forecasts are set to nominally originate at 00:00Z. In addition, some of the models are also tested with an origination time of 12:00Z. This 00:00Z common reference results in a slight performance handicap for the European validations compared to the North American validations; however as can be gauged from the results, e.g., by comparing the 00:00Z and 12:00Z performance, this is a very small effect.

4.3 Validation

The evaluation was performed for sites in the US, Canada and Europe covering different climatic conditions. These include Mediterranean climate in Southern Spain, humid continental climate in Canada, mostly continental climate in Central Europe and some high alpine stations in Switzerland, and finally arid, sub-tropical, semi-arid, and continental conditions in the US.

Because of operational contingencies not all the models could be validated at all the sites. Models 1–5 and 8 were validated in the US. Models 1, 2 and 3 (without MOS application) were validated against Canadian sites. Models 2, 3, 6, 7, 9 and10 were validated in Europe. The common denominators to all the validations are (1) the ECMWF model and (2) the GFS-driven WRF model applied by various operators under slightly different conditions.

4.3.1 Validation measurements

All benchmark measurement stations are part of net- works operated by each considered country's weather services and include well maintained and quality controlled Class I instruments and data.

United States

Validation measurements consist of hourly averaged global horizontal irradiance (GHI) recorded for a 1 year period (May 1st, 2009, through April 30th, 2010) at the seven stations of the SURFRAD network (SURFRAD, 2010). The stations are listed in Table 1.

Table 1. Location and climate type for the US sites.

Station	Latitude (°)	Longitude (°)	Elevation (m)	Climate
Goodwin Creek	34.25	89.87	98	Humid continental
Desert Rock	36.63	116.02	1107	Arid
Bondville	40.05	88.37	213	Continental
Boulder	40.13	105.24	1689	Semi-arid
Penn State	40.72	77.93	376	Humid continental
Sioux Falls	48.73	96.62	473	Continental
Fort Peck	48.31	105.1	643	Continental

Some of the models were only processed for a subset of the SURFRAD sites. The ARPS, MASS and WRF model processed by AWS Truepower could only be run at Desert Rock, Goodwin Creek and Penn State, while the WRF-ASRC model, run as part of the air

quality forecast model, was only available for Goodwin Creek and Penn State.

All models were processed to deliver data up to 48 h ahead (next day and 2 day forecasts). The ECMWF fore- casts were processed up to 3 days ahead, and the NDFD up to 7 days ahead.

Canada

The three sites used for evaluating irradiance forecasts in Canada are listed in Table 2. The validation period runs from June 1, 2009 to May 31, 2010. The GEM, ECMWF and WRF-ASRC forecasts originating at 00:00Z were processed for forecast horizons of 0–48 h ahead, and compared to hourly average irradiance measured at the three ground stations. The mean of the individual forecast models was also evaluated against the ground station data to investigate whether this yields any benefits, as reported in the case of wind forecasts (Ernst *et al*., 2007).

Table 2. Location and climate type for the Canadian sites.

Station	Latitude (°)	Longitude (°)	Elevation (m)	Climate
Egbert	44.23	79.78	250	Humid continental
Bratt's Lake	50.20	104.71	580	Humid continental
Varennes	45.63	73.38	36	Humid continental

In the case of WRF, forecasts were only available for two stations (Egbert and Varennes) for the last 2 months of the evaluation period (i.e. April 1, 2010 to May 31, 2010).

Europe

The selected data sets with hourly average values of measured irradiance for Europe cover four countries: Southern Germany, Switzerland including mountain stations, Austria, and Southern Spain. The period of evaluation for all sites and forecasting approaches is July 2007 to June 2008.

1 German sites

For the German sites (see Table 3) ground measurement data for three locations were provided by the German weather service (DWD). Forecast data of ECMWF, WRF-Meteotest, and BLUE FORECAST were considered for horizons up to 3 days ahead. Skiron-CEN-ER forecasts were processed for 48 h.

Table 3. Location and climate type for the German sites.

Station	Latitude (°)	Longitude (°)	Elevation (m)	Climate
Fürstenzell	48.55	-13.35	476	Continental
Stuttgart	48.83	-9.20	318	Continental
Würzburg	49.77	-9.97	275	Continental

2 Austrian sites

In addition to irradiance forecasts of ECMWF, WRF-Meteotest, Skiron-CENER and BLUE FORECAST, irradiance forecasts based on cloud coverforecasts by the meteorologists' of Bluesky up to 48 h ahead were evaluated. The Austrian ground measurements (see Table 4) were recorded by BLUESKY in two locations.

Table 4. Location and climate type for the Austrian sites.

Station	Latitude (°)	Longitude (°)	Elevation (m)	Climate
Linz	48.30	-14.30	266	Continental
Vienna	48.20	-16.37	171	Continental

3. Swiss sites

The models considered for the Swiss validation include ECMWF, WRF-Meteotest, and BLUE FORECAST. Ground measurements for sixteen sites are from the MeteoSwiss network. The sites considered for Switzerland cover a considerable variety in climatic conditions (see Table 5).

Table 5. Location and climate type for the Swiss sites.

Station	Latitud	Longitud	Elevation	Climate

	e (°)	e (°)	(m)	
Basel-Binningen	47.54	-7.58	316	Temperate Atlantic
Payerne	46.81	-6.94	490	Moderate maritime/continental
La Chaux-de-Fonds	47.09	-6.80	1018	Temperate Atlantic
Bern-Liebefeld	46.93	-7.42	565	Moderate maritime/continental
Buchs-Suhr	47.38	-8.08	387	Moderate maritime/continental
Napf	47.00	-7.94	1406	Moderate maritime/continental
Zürich SMA	47.38	-8.57	556	Moderate maritime/continental
Säntis	47.25	-9.34	2490	Alpine
St. Gallen	47.43	-9.40	779	Moderate maritime/continental
Genève-Cointrin	46.24	-6.12	420	Moderate maritime/continental
Sion	46.22	-7.34	482	Dry alpine
Montana	46.31	-7.49	1508	Alpine
Jungfraujoch	46.55	-7.99	3580	High alpine
Locarno-Magadino	46.16	-8.88	197	Warm temperate, humid
Weissfluhjoch	46.83	-9.81	2690	Alpine
Davos	46.81	-9.84	1590	Continental/alpine

4. Spanish sites

Forecasts for Spain were processed based on the global ECMWF model and three different mesoscale models (WRF-Jaén, Skiron-

CENER and HIRLAM-CIEMAT). The three ground measurement stations (see Table 6) operated by the Spanish Weather Service AEMet are located in the South of Spain.

Table 6. Location and climate type for the Spanish sites.

Station	Latitude (°)	Longitude (°)	Elevation (m)	Climate
Huelva	37.28	-6.91	19	Mediterranean
Córdoba	37.84	-4.85	91	Mediterranean
Granada	37.14	-3.63	687	Mediterranean

4.3.2 Overview of forecast model benchmarking tests

A summary of the models tested as part of this article is presented in Table 7. The ECMWF model and the GFS- driven WRF model are the only common denominators to all tests, noting that the WRF model was run by different entities in different countries, with slightly differing operational settings and was not available at some of the US and Canadian sites.

Table 7. Overview of forecast validations.
[a] Models run both with and without MOS.

	Forecast models – the number in () corresponds to the descriptive number in the text	Time horizon (days)
Europe		
Germany	ECMWF (2)	3
	WRF-Meteotest (3)	3
	SKIRON-CENER (6)	3
	BLUE FORECAST (9)	2
Switzerland	ECMWF (2)	3
	WRF-Meteotest (3)	3
	BLUE FORECAST (9)	3
Austria	ECMWF (2)	3

	WRF-Meteotest (3)	3
	CENER (6)	3
	BLUE FORECAST (9)	2
	BLUESKY-Meteorologists (10)	2
Spain	ECMWF (2)	3
	WRF-UJAEN (3)	3
	CENER (6)	3
	HIRLAM (7)	2
	BLUE FORECAST (9)	3
USA		
USA	GEM (1)	2
	ECMWF (2)	3
	WRF-ASRC(3)	2
	WRF-AWS[a] (3)	2
	MASS[a] (4)	2
	ARPS[a] (5)	2
	NDFD (8)	7
Canada		
Canada	GEM (1)	2
	ECMWF (2)	2
	WRF-ASRC (3)	2

4.3.3 Concept of evaluation

To compare the performance of the different methods, hourly forecasts for the evaluation sites as provided by the different research groups and private companies were evaluated against hourly mean values of measured irradiance, regardless of the original spatial and temporal resolution of the underlying NWP models. The analysis presented focuses on the "end-use" accuracy of these site-specific, hourly irradiance predictions derived by the different forecast providers from gridded NWP data rather than on the evaluation of the direct NWP model output. To assess the performance of forecast algorithms, in general, a lot of different aspects have to be taken into account. In this paper, which aims at the inter-comparison of different models we

focus on a few, basic measures of accuracy that are considered to be most relevant for the intended application of solar power prediction.

The validation metrics include the root mean square error, RMSE, to compare predicted irradiance $I_{pred,i}$, to measured irradiance $I_{meas,i}$.

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \left(I_{pred,i} - I_{meas,i}\right)^2} \tag{4.1}$$

Here, N is the number of evaluated data pairs. The RMSE is often considered as the most important model validation metric in the context of renewable power forecasting. Because it is based on the square of differences between modeled and measured values, large forecast errors and outliers are weighted more strongly than small errors. It is suitable for applications where small errors are more tolerable and large forecast errors have a disproportionately high impact, which is the case for many aspects of grid management issues.

Table 8. Relative RMSE US.

% RMSE		Day	Bondville	Boulder	Desert Rock	Fort Peck	Goodwin Creek	Penn State	Sioux Falls	Composite
Mean GHI (W/m2)			335	374	466	326	363	298	328	356
Ref. sat. mod.			21%	25%	15%	23%	20%	28%	22%	22%
Persistence	0:00Z	1	59%	51%	29%	46%	51%	65%	51%	50%
GEM	0:00Z	1	35%	38%	21%	30%	33%	38%	38%	33%
GEM	12:00Z	1	33%	36%	20%	29%	33%	38%	36%	32%
ECMWF	0:00Z	1	34%	38%	21%	32%	31%	39%	38%	33%
NDFD	0:00Z	1	40%	44%	25%	38%	38%	48%	44%	40%
NDFD	12:00Z	1	40%	43%	23%	37%	37%	45%	43%	38%
WRF-ASRC		1	46%				43%	51%		44%
MASS	0:00Z	1			31%		53%	67%		55%
MASS	12:00Z	1			32%		55%	64%		54%
MAS-MOS	0:00Z	1			24%		38%	44%		38%
MAS-MOS	12:00Z	1			24%		38%	44%		38%
WRF-AWS	0:00Z	1			25%		45%	54%		45%
WRF-AWS	12:00Z	1			26%		47%	58%		47%

WRF-AWS-MOS	0:00Z	1			23%		41%	47%		40%
WRF-AWS-MOS	12:00Z	1			23%		41%	46%		40%
ARPS	0:00Z	1			33%		54%	69%		56%
ARPS-MOS	0:00Z	1			24%		43%	48%		42%
GEMS/ECMWF	0:00Z	1	32%	37%	20%	30%	31%	37%	37%	32%
Persistence	0:00Z	2	64%	57%	32%	49%	60%	72%	57%	56%
GEM	0:00Z	2	37%	37%	21%	32%	35%	40%	37%	34%
GEM	12:00Z	2	34%	36%	22%	30%	33%	39%	36%	33%
ECMWF	0:00Z	2	38%	39%	22%	34%	34%	41%	39%	35%
NDFD	0:00Z		43%	45%	27%	39%	40%	49%	45%	41%
NDFD	12:00Z	2	42%	45%	25%	38%	39%	48%	45%	40%
WRF-ASRC		2	50%				45%	55%		46%
MASS	0:00Z	2			31%		57%	68%		56%
MASS	12:00Z	2			31%		58%	66%		55%
MAS-MOS	0:00Z	2			24%		40%	46%		39%
MAS-MOS	12:00Z	2			24%		40%	46%		39%
WRF-AWS	0:00Z	2			27%		47%	59%		47%
WRF-AWS	12:00Z	2			26%		46%	58%		46%
WRF-AWS-MOS	0:00Z	2			24%		42%	49%		41%
WRF-AWS-MOS	12:00Z	2			23%		41%	48%		40%
ARPS	0:00Z	2			33%		55%	70%		57%
ARPS-MOS	0:00Z	2			25%		44%	50%		42%
GEMS/ECMWF	0:00Z	2	35%	37%	20%	31%	33%	38%	37%	33%
Persistence	0:00Z	3	67%	58%	32%	54%	63%	77%	58%	58%
ECMWF	0:00Z	3	40%	41%	23%	35%	37%	45%	41%	37%
NDFD	0:00Z	3	47%	46%	29%	39%	44%	54%	46%	44%
NDFD	12:00Z	3	45%	46%	28%	38%	42%	51%	46%	42%
Persistence	0:00Z	4	69%	59%	33%	54%	62%	79%	59%	59%
NDFD	0:00Z	4	49%	46%	29%	39%	46%	55%	46%	44%
NDFD	12:00Z	4	47%	46%	29%	38%	45%	55%	46%	44%
Persistence	0:00Z	5	71%	59%	33%	52%	63%	78%	59%	59%
NDFD	0:00Z	5	52%	47%	29%	41%	47%	58%	47%	46%
NDFD	12:00Z	5	51%	47%	29%	40%	48%	58%	47%	45%
Persistence	0:00Z	6	68%	59%	33%	54%	60%	78%	59%	59%
NDFD	0:00Z	6	56%	49%	29%	43%	50%	61%	49%	48%
NDFD	12:00Z	6	56%	50%	30%	42%	48%	59%	50%	48%
Persistence	0:00Z	7	67%	60%	34%	54%	60%	75%	60%	59%
NDFD	0:00Z	7	57%	51%	31%	45%	54%	61%	51%	50%
NDFD	12:00Z	7	56%	51%	30%	44%	52%	59%	51%	49%

The MAE – the mean absolute error:

$$MAE = \frac{1}{N} \cdot \sum_{i=1}^{N} \left| I_{pred,i} - I_{meas,i} \right| \tag{4.2}$$

is a useful complement to the RMSE that is effective at quantifying the tightness of the measured-modeled scatter plot near the 1-to-1 line. In particular it is appropriate for applications with linear cost functions, that is, where the costs that are caused by a wrong forecast are proportional to the forecast error.

Table 9. Relative RMSE Central Europe.

% RMSE		Day	Fürstenzell	Stuttgart	Würzburg	Composite Germany	Linz	Wien	Composite Austria	Composite Swizerland
Mean GHI (W/m2)			227	233	224	228	206	241	224	270
Reference satellite model										
Persistence	0:00Z	1	66%	63%	61%	64%	71%	57%	64%	58%
ECMWF-OL	0:00Z	1	40%	40%	42%	40%	50%	42%	46%	40%
BLUE FORECAST	0:00Z	1	41%	42%	42%	42%	46%	43%	45%	41%
WRF-Meteotest	0:00Z	1	48%	51%	57%	52%	64%	47%	55%	44%
CENER	0:00Z	1	46%	51%	53%	50%	63%	53%	58%	
Meteorologists	0:00Z	1					55%	46%	50%	
Persistence	0:00Z	2	74%	69%	68%	70%	78%	63%	70%	64%
ECMWF-OL	0:00Z	2	41%	42%	42%	42%	52%	43%	47%	42%
BLUE FORECAST	0:00Z	2	43%	45%	44%	44%	49%	41%	45%	42%
WRF-Meteotest	0:00Z	2	51%	55%	59%	55%	64%	53%	59%	46%
CENER	0:00Z	2	48%	54%	56%	53%	65%	54%	60%	
Meteorologists	0:00Z	2					55%	44%	49%	
Persistence	0:00Z	3	75%	74%	71%	73%	78%	65%	72%	67%
ECMWF	0:00Z	3	44%	46%	45%	45%	54%	47%	51%	43%
BLUE FORECAST	0:00Z	3	45%	47%	45%	46%	51%	45%	48%	44%
WRF-Meteotest	0:00Z	3	57%	62%	63%	61%	67%	58%	63%	51%

The MBE-mean bias error:

$$MBE = \frac{1}{N} \cdot \sum_{i=1}^{N} \left(I_{pred,i} - I_{meas,i} \right) \tag{4.3}$$

describes systematic deviations of a forecast. The agreement between the distribution functions of measured and predicted time series can be evaluated using the Kolmogorov–Smirnov test integral (KSI) (e.g., see Perez *et al*., 2010). We decided to use a robust interpretation of the KSI metric that simply describes the integrated absolute difference between the predicted and measured normalized cumulative distributions CDF_{pred} and CDF_{meas} integrated over all irradiance levels I and normalized to 1,

$$KSI = \frac{1}{I_{max}} \cdot \int_0^{I_{max}} \left| CDF_{pred} - CDF_{mean} \right| dI \tag{4.4}$$

The evaluation of distribution functions is helpful e.g. for applications where decisions are related to threshold values. However, the KSI metric is less important for fore- cast evaluation than the other metrics introduced and is given here only for the Canadian and US sites, where a discretized version of Eq. (4.4) was used to evaluate the KSI metric.

The accuracy measures are calculated using only day- time hours (I > 0) (i.e., night values with zero irradiance are excluded from the evaluation.) The evaluation results are grouped according to forecast days. For a model run at 00:00Z, the results for the first forecast day (intraday) integrate forecast horizons up to 24 h, the second forecast day (day-ahead) integrates forecast horizons from 25 to 48 h, and so on. The reason for grouping results according to forecast days rather than forecast hours is the strong dependency of forecast accuracy on the daytime caused by the daily course of irradiance.

Relative values of the error measures are obtained by normalization to the mean ground measured irradiance of the considered period.

As an additional quality check, forecasts often are compared to trivial reference models, which are the result of simple considerations and not of modeling efforts. It is worthwhile to implement and run a complex forecasting tool if it is able to clearly outperform trivial (i.e.,

self-evident) reference models. The most common such reference model for short term forecasts is persistence. Persistence consists of projecting currently and recently measured conditions into the future while accounting for solar geometry changes. Here, where we are inter-comparing NWP models originating nominally at 00:00Z, designed for next and subsequent day forecasts, the benchmark persistence is obtained by determining the daily global clear sky index kt* (ratio of measured irradiance to irradiance for clear sky conditions) from the last available day and projecting this index for all subsequent forecast days/hours.

Forecast skill can be gauged by comparing the forecast and reference (i.e., persistence) errors as follows:

$$MSE\ skill\ score = \frac{\left(MSE_{ref} - MSE_{forecast}\right)}{MSE_{ref}} \tag{4.5}$$

where MSE is the mean square error (square of the RMSE as defined in Eq. (4.1)). A MSE skill score of one corresponds to a perfect model. Negative MSE skill scores indicate performance worse than persistence.

For the US sites, the satellite irradiance model developed by Perez *et al*. (2002) and used in the NSRDB (2005) and SolarAnywhere (2010) is used as a complementary reference to gauge the performance of the forecast models – note that this reference model is an "existing conditions" and not a forecast model.

Results of the forecast evaluation are provided at different levels of detail. Tables 8–21 give the different validation metrics for the single sites. (As an exception, for Switzerland with more than 15 stations and the same forecast models available for all stations, the average of the errors of the individual sites is given instead.) These detailed results allow for directly assessing and comparing the performance of different forecasting methods for a given location with its particular climatic conditions, which is of interest not only from the scientific point of

view. Forecast users, e.g. a utility company or a plant operator, are also often interested in applying the forecasts and hence in the relevant information about forecast accuracy for a certain location or region.

Table 10. Relative RMSE SPAIN.

% RMSE		Day	Cordoba	Granada	Huelva	Composite Spain
Mean GHI (W/m2) Reference satellite model			443	409	407	420
Persistence	0:00Z	1	34%	36%	34%	35%
ECMWF-OL	0:00Z	1	23%	23%	20%	22%
CENER	0:00Z	1	26%	25%	26%	25%
WRF-UJAEN	0:00Z	1	28%	27%	25%	26%
HIRLAM	0:00Z	1	26%	32%	26%	29%
Persistence	0:00Z	2	37%	39%	38%	38%
ECMWF-OL	0:00Z	2	25%	22%	21%	23%
CENER	0:00Z	2	30%	26%	27%	27%
WRF-UJAEN	0:00Z	2	29%	29%	27%	28%
HIRLAM	0:00Z	2	29%	36%	32%	33%
Persistence	0:00Z	3	29%	41%	39%	40%
ECMWF	0:00Z	3	29%	24%	22%	23%
WRF-UJAEN	0:00Z	3	29%	30%	30%	30%
HIRLAM	0:00Z	3	29%	39%	36%	35%

Table 11. Relative RMSE Canada.

[a] The WRF model was only run on a 2 month data subset and results were prorated using the other models as a template.

% RMSE		Day	Egbert	Bratt's Lake	Varennes	Composite
Mean GHI (W m-2) Reference satellite model			320	306	306	311
Persistence	0:00Z	1	52%	52%	58%	54%
GEM	0:00Z	1	32%	31%	37%	33%
ECMWF	0:00Z	1	32%	31%	35%	32%
WRF-ASRC[a]	0:00Z	1	40%		44%	42%
GEM/ECMWF/WRF-ASRC[a]	0:00Z	1	31%		33%	30%
GEM/ECMWF	0:00Z	1	31%	29%	34%	31%
Persistence	0:00Z	2	56%	57%	63%	59%
GEM	0:00Z	2	33%	35%	38%	35%
ECMWF	0:00Z	2	34%	35%	38%	36%
WRF-ASRC[a]	0:00Z	2	43%		45%	44%
GEM/ECMWF/WRF[a]	0:00Z	2	32%		36%	32%
GEM/ECMWF	0:00Z	2	32%	33%	36%	34%

In addition to the evaluation and model comparison for the single sites, all-site composite errors for the different evaluation regions (US, Canada, and Europe) are calculated by averaging the errors of the individual sites, in order to give a quick overview of model performances. For some of the models forecasts are available only for a

subset of sites in a given region. For these models i an estimate of the all-site composite value, e.g. the $RMSE^*_{all\text{-}sites,i}$, is prorated with the following equation:

$$RMSE^*_{all-sites,i} = RMSE_{subset,i} \cdot \frac{\sum_{j=1}^{M} RMSE^*_{all-sites,j}}{\sum_{j=1}^{M} RMSE_{subset,j}} \quad (4.6)$$

i.e. by multiplying the composite $RMSE_{subset,i}$ for the subset of sites at which the forecast is available with the ratio of the average all-site composite RMSE to the average subset composite RMSE of all models j that are available for all sites. This estimate of the average performance is of course provided with some uncertainty. In particular, averaging over sites with different climatic conditions may result in biased overall estimates – note that this is also the reason why composite values for Northern and Southern Europe are given separately. However, given the availability of the detailed site-specific results in Tables 8–21, we consider it to be a reasonable simplification.

4.4 Results and discussion

An overview of the all-site composite relative RMSE values for the different study regions US, Canada, Central Europe and Spain is given in Figs. 1–4. Corresponding RMSE values for the single sites are given Tables 8–11 respectively. Figs. 5–14 accordingly provide composite summaries for the MAE, MBE and KSI metrics, also completed by the detailed site specific results in Tables 12–15 for the MAE, 16–19 for the MBE and 20–21 for the KSI.

We first give a description and discussion of the US results, which include the largest number of different forecast models and also cover different climate zones. Next, the discussion is extended to the evaluation for Canada and Europe and some additional findings are highlighted.

Table 12. Relative MAE US.

% MAE		Day	Bondville	Boulder	Desert Rock	Fort Peck	Goodwin Creek	Penn State	Sioux Falls	Composite
Mean GHI (W m-2)			335	374	466	326	363	298	328	356
Reference satellite model			14%	17%	9%	16%	13%	18%	15%	14%
Persistence	0:00Z	1	39%	34%	18%	29%	34%	44%	34%	33%
GEM	0:00Z	1	24%	24%	11%	19%	21%	26%	24%	21%
GEM	12:00Z	1	23%	23%	11%	18%	22%	26%	23%	21%
ECMWF	0:00Z	1	21%	23%	11%	19%	20%	25%	23%	21%
NDFD	0:00Z	1	26%	28%	14%	23%	23%	30%	28%	25%
NDFD	12:00Z	1	26%	27%	14%	23%	23%	29%	27%	24%
WRF-ASRC		1	30%				28%	34%		28%
MASS	0:00Z	1			21%		39%	49%		40%
MASS	12:00Z	1			22%		40%	47%		39%
MAS-MOS	0:00Z	1			15%		27%	31%		27%
MAS-MOS	12:00Z	1			15%		27%	32%		27%
WRF-AWS	0:00Z	1			16%		29%	37%		29%
WRF-AWS	12:00Z	1			16%		29%	39%		31%
WRF-AWS-MOS	0:00Z	1			14%		28%	34%		28%
WRF-AWS-MOS	12:00Z	1			14%		28%	33%		27%
ARPS	0:00Z	1			23%		39%	49%		40%
ARPS-MOS	0:00Z	1			15%		30%	34%		29%
GEMS/ECMWF	0:00Z	1	21%	23%	11%	18%	19%	25%	23%	20%
Persistence	0:00Z	2	44%	39%	19%	32%	41%	50%	39%	38%
GEM	0:00Z	2	25%	24%	12%	20%	22%	27%	24%	22%
GEM	12:00Z	2	23%	23%	12%	19%	21%	26%	23%	21%
ECMWF	0:00Z	2	24%	24%	12%	21%	21%	27%	24%	22%
NDFD	0:00Z	2	28%	29%	16%	24%	25%	32%	29%	26%
NDFD	12:00Z	2	27%	29%	15%	24%	24%	30%	29%	25%
WRF-ASRC		2	32%				29%	37%		30%
MASS	0:00Z	2			21%		42%	49%		40%
MASS	12:00Z	2			21%		43%	47%		40%
MAS-MOS	0:00Z	2			15%		28%	33%		27%
MAS-MOS	12:00Z	2			15%		28%	33%		27%
WRF-AWS	0:00Z	2			17%		30%	39%		31%
WRF-AWS	12:00Z	2			16%		29%	39%		30%
WRF-AWS-MOS	0:00Z	2			15%		29%	34%		28%
WRF-AWS-MOS	12:00Z	2			14%		29%	34%		28%
ARPS	0:00Z	2			23%		40%	50%		41%
ARPS-MOS	0:00Z	2			15%		31%	34%		29%
GEMS/ECMWF	0:00Z	2	23%	23%	11%	19%	21%	26%	23%	21%
Persistence	0:00Z	3	46%	40%	20%	36%	44%	54%	40%	40%
ECMWF	0:00Z	3	25%	25%	12%	21%	23%	30%	25%	23%
NDFD	0:00Z	3	31%	31%	17%	25%	28%	36%	31%	28%
NDFD	12:00Z	3	30%	30%	16%	24%	26%	34%	30%	27%
Persistence	0:00Z	4	49%	41%	20%	35%	43%	56%	41%	41%
NDFD	0:00Z	4	34%	32%	17%	25%	30%	38%	32%	30%
NDFD	12:00Z	4	32%	32%	17%	24%	29%	37%	32%	29%
Persistence	0:00Z	5	50%	41%	20%	34%	44%	55%	41%	41%
NDFD	0:00Z	5	36%	32%	17%	26%	31%	41%	32%	31%
NDFD	12:00Z	5	35%	32%	17%	26%	31%	40%	32%	30%
Persistence	0:00Z	6	48%	41%	20%	35%	42%	55%	41%	40%
NDFD	0:00Z	6	39%	34%	17%	28%	34%	43%	34%	33%
NDFD	12:00Z	6	39%	34%	18%	27%	32%	42%	34%	32%

Persistence	0:00Z	7	47%	42%	21%	36%	41%	53%	42%	40%
NDFD	0:00Z	7	40%	36%	18%	29%	37%	43%	36%	34%
NDFD	12:00Z	7	40%	35%	18%	29%	35%	42%	35%	33%

RMSE all-site composite values for the US given in Fig. 1 show a considerable spread for the different models. They range between 32% and 47% for Day 1 forecasts and – showing only a slight increase – between 34% and 48% for Day 2 forecasts. The corresponding values of MAE(Fig. 5) lie between 20% and 29% for Day 1 and between 22% and 31% for Day 2 forecasts.

Table 13. Relative MAE Central Europe.

% MAE		Day	Fürstenzell	Stuttgart	Würzburg	Composite	Linz	Wien	Composite Austria	Composite Switzerland
Mean GHI (W m-2) Reference satellite model			227	233	224	228	206	241	224	270
Persistence	0:00Z	1	42%	40%	39%	41%	46%	36%	41%	39%
ECMWF-OL	0:00Z	1	26%	26%	27%	26%	32%	26%	29%	26%
BLUE FORECAST	0:00Z	1	26%	28%	28%	27%	28%	27%	28%	27%
WRF-Meteotest	0:00Z	1	30%	32%	37%	33%	40%	29%	35%	28%
CENER	0:00Z	1	29%	32%	33%	32%	43%	35%	39%	
Meteorologists	0:00Z	1					35%	28%	32%	
Persistence	0:00Z	2	48%	45%	44%	46%	50%	41%	46%	43%
ECMWF-OL	0:00Z	2	27%	28%	28%	28%	34%	27%	31%	27%
BLUE FORECAST	0:00Z	2	28%	30%	29%	29%	30%	27%	28%	28%
WRF-Meteotest	0:00Z	2	32%	34%	38%	35%	40%	34%	37%	29%
CENER	0:00Z	2	31%	34%	36%	34%	44%	35%	40%	
Meteorologists	0:00Z	2					35%	27%	31%	
Persistence	0:00Z	3	49%	48%	47%	48%	51%	42%	47%	45%
ECMWF	0:00Z	3	29%	30%	29%	30%	35%	30%	32%	28%
BLUE FORECAST	0:00Z	3	29%	31%	31%	30%	32%	30%	31%	30%
WRF-Meteotest	0:00Z	3	36%	38%	40%	38%	42%	37%	40%	32%

Table 14. Relative MAE Spain.

% MAE		Day	Cordoba	Granada	Huelva	COMPOSITE SPAIN
Mean GHI (W m-2) Reference satellite model			443	409	407	420
Persistence	0:00Z	1	20%	19%	19%	19%
ECMWF-OL	0:00Z	1	15%	13%	12%	13%

CENER	0:00Z	1	16%	16%	17%	16%
WRF-UJAEN	0:00Z	1	15%	14%	13%	14%
HIRLAM	0:00Z	1	19%	25%	19%	21%
Persistence	0:00Z	2	22%	21%	21%	22%
ECMWF-OL	0:00Z	2	16%	13%	12%	14%
CENER	0:00Z	2	18%	17%	17%	17%
WRF-UJAEN	0:00Z	2	16%	15%	14%	15%
HIRLAM	0:00Z	2	21%	27%	23%	24%
Persistence	0:00Z	3	24%	23%	23%	23%
ECMWF	0:00Z	3	16%	14%	13%	14%
WRF-UJAEN	0:00Z	3	16%	15%	16%	16%
HIRLAM	0:00Z	3	22%	30%	25%	24%

Table 15. Relative MAE Canada.

[a] The WRF model was only run on a 2 month data subset and results were prorated using the other models as a template.

% MAE		Day	Egbert	Bratt's Lake	Varennes	Composite
Mean GHI (W m-2)			320	306	306	311
Reference satellite model						
Persistence	0:00Z	1	37%	37%	41%	38%
GEM	0:00Z	1	23%	20%	25%	23%
ECMWF	0:00Z	1	20%	19%	22%	21%
WRF-ASRC[a]	0:00Z	1	27%		30%	28%
GEM/ECMWF/WRF-ASRC[a]	0:00Z	1	21%		22%	20%
GEM/ECMWF	0:00Z	1	21%	19%	23%	21%
Persistence	0:00Z	2	41%	39%	46%	42%
GEM	0:00Z	2	23%	22%	25%	23%
ECMWF	0:00Z	2	22%	21%	25%	23%
WRF-ASRC[a]	0:00Z	2	30%		32%	31%
GEM/ECMWF/WRF[a]	0:00Z	2	23%		24%	22%
GEM/ECMWF	0:00Z	2	22%	21%	24%	22%

Table 16. Relative MBE US.

% MBE		Day	Bondville	Boulder	Desert Rock	Fort Peck	Goodwin Creek	Penn State	Sioux Falls	Composite
Mean GHI (W m-2)			335	374	466	326	363	298	328	356
Reference satellite model			0%	0%	2%	0%	1%	2%	2%	1%
Persistence	0:00Z	1	-2%	-1%	-2%	-2%	-1%	-1%	-1%	-1%
GEM	0:00Z	1	8%	10%	2%	6%	8%	11%	10%	8%
GEM	12:00Z	1	7%	7%	3%	5%	7%	12%	7%	7%
ECMWF	0:00Z	1	6%	14%	5%	9%	6%	12%	14%	10 %
NDFD	0:00Z	1	-7%	-9%	-1%	4%	-6%	-8%	-9%	-5%
NDFD	12:00Z	1	-9%	-10%	-1%	2%	-8%	-9%	-10%	-6%
WRF-ASRC		1	9%				13%	13%		15 %
MASS	0:00Z	1			19%		34%	41%		37 %
MASS	12:00Z	1			18%		34%	40%		36 %
MAS-MOS	0:00Z	1			-1%		1%	0%		0%
MAS-MOS	12:00Z	1			0%		-1%	0%		-1%
WRF-AWS	0:00Z	1			1%		19%	23%		17

										%
WRF-AWS	12:00Z	1			1%		18%	22%		16 %
WRF-AWS-MOS	0:00Z	1			1%		-1%	0%		0%
WRF-AWS-MOS	12:00Z	1			0%		0%	0%		0%
ARPS	0:00Z	1			20%		33%	41%		37 %
ARPS-MOS	0:00Z	1			0%		-2%	0%		-1%
GEMS/ECMWF	0:00Z	1	7%	12%	4%	8%	7%	12%	12%	9%
Persistence	0:00Z	2	-2%	-2%	-2%	-2%	-2%	-1%	-2%	-2%
GEM	0:00Z	2	6%	7%	2%	5%	8%	11%	7%	7%
GEM	12:00Z	2	6%	7%	3%	4%	6%	10%	7%	6%
ECMWF	0:00Z	2	5%	14%	5%	9%	6%	10%	14%	9%
NDFD	0:00Z	2	-8%	-9%	-3%	3%	-6%	-7%	-9%	-6%
NDFD	12:00Z	2	-8%	-10%	-2%	2%	-8%	-9%	-10%	-6%
WRF-ASRC		2	8%				14%	12%		14 %
MASS	0:00Z	2			18%		36%	40%		35 %
MASS	12:00Z	2			18%		37%	36%		34 %
MAS-MOS	0:00Z	2			-1%		-1%	-2%		-1%
MAS-MOS	12:00Z	2			-1%		1%	-2%		-1%
WRF-AWS	0:00Z	2			1%		18%	22%		15 %
WRF-AWS	12:00Z	2			1%		17%	21%		15 %
WRF-AWS-MOS	0:00Z	2			0%		-1%	-1%		-1%
WRF-AWS-MOS	12:00Z	2			0%		-1%	-1%		-1%
ARPS	0:00Z	2			20%		31%	41%		34 %
ARPS-MOS	0:00Z	2			0%		-3%	0%		-1%
GEMS/ECMWF	0:00Z	2	6%	10%	4%	7%	7%	11%	10%	8%
Persistence	0:00Z	3	-2%	-2%	-2%	-1%	-2%	-1%	-2%	-2%
ECMWF	0:00Z	3	5%	13%	5%	9%	6%	10%	13%	9%
NDFD	0:00Z	3	-6%	-9%	-4%	3%	-8%	-10%	-9%	-6%
NDFD	12:00Z	3	-8%	-10%	-3%	1%	-8%	-10%	-10%	-7%
Persistence	0:00Z	4	-2%	-2%	-2%	-1%	-3%	-2%	-2%	-2%
NDFD	0:00Z	4	-5%	-7%	-3%	4%	-7%	-7%	-7%	-5%
NDFD	12:00Z	4	-7%	-8%	-4%	4%	-9%	-9%	-8%	-6%
Persistence	0:00Z	5	-2%	-2%	-2%	-1%	-2%	-2%	-2%	-2%
NDFD	0:00Z	5	-4%	-5%	-3%	6%	-7%	-7%	-5%	-4%
NDFD	12:00Z	5	-6%	-6%	-3%	4%	-8%	-7%	-6%	-5%
Persistence	0:00Z	6	-2%	-2%	-2%	-1%	-2%	-2%	-2%	-2%
NDFD	0:00Z	6	-3%	-4%	-1%	6%	-6%	-8%	-4%	-3%
NDFD	12:00Z	6	-4%	-5%	-2%	6%	-8%	-6%	-5%	-4%
Persistence	0:00Z	7	-2%	-2%	-2%	-1%	-2%	-2%	-2%	-2%
NDFD	0:00Z	7	-2%	-5%	-1%	5%	-7%	-7%	-5%	-3%
NDFD	12:00Z	7	-2%	-5%	-2%	5%	-8%	-6%	-5%	-3%

Lowest MAE and RMSE values are found for the global model ECMWF and GEM irradiance forecasts. All considered mesoscale-model forecasts (WRF-AFS, WRF-ASRC, ARPS, MAS) as well as the NDFD based forecasts show larger forecast errors. This indicates some

shortcomings in the selected mesoscale models' radiation and/or cloud schemes. Another reason might be the use of lateral boundary conditions from GFS, used to initialize all mesoscale models evaluated here. In recent work by Mathiesen and Kleissl (2011), the GFS model irradiance forecasts were found to have a similar performance to those of the ECMWF model when applying a simple post- processing. This suggests that the performance difference noted here between the ECMWF and GEM model on the one hand and the different mesoscale models initialized with GFS on the other hand has more to do with the mesoscale models themselves than with the GFS boundary conditions. Additional detailed studies comparing, e.g., the performance of mesoscale models as a function of the boundary conditions from different global models, are required to confirm this assertion.

Table 17. Relative MBE Central Europe.

% MBE		Day	Fürstenzell	Stuttgart	Würzburg	Composite	Linz	Wien	Composite Austria	Composite Switzerland
Mean GHI (W m-2) Reference satellite model			227	233	224	228	206	241	224	270
Persistence	0:00Z	1	-3%	-2%	-1%	-2%	-11%	-2%	-6%	-6%
ECMWF-OL	0:00Z	1	-1%	-4%	-4%	-3%	12%	2%	7%	0%
BLUE FORECAST	0:00Z	1	0%	-4%	-1%	-1%	0%	1%	1%	-3%
WRF-Meteotest	0:00Z	1	1%	0%	-1%	0%	28%	14%	21%	1%
CENER	0:00Z	1	7%	3%	8%	6%	21%	6%	14%	
Meteorologists	0:00Z	1					9%	-1%	0%	
Persistence	0:00Z	2	-3%	-3%	-2%	-3%	-11%	-2%	-7%	-7%
ECMWF-OL	0:00Z	2	-1%	-4%	-5%	-3%	12%	0%	6%	0%
BLUE FORECAST	0:00Z	2	1%	-3%	-1%	-1%	1%	0%	1%	-1%
WRF-Meteotest	0:00Z	2	4%	-1%	-11%	-5%	25%	2%	13%	1%
CENER	0:00Z	2	5%	0%	6%	4%	18%	1%	9%	
Meteorologists	0:00Z	2					8%	-2%	3%	
Persistence	0:00Z	3	-4%	-3%	-2%	-3%	-11%	-2%	-7%	-7%
ECMWF	0:00Z	3	-2%	-4%	-5%	-3%	13%	2%	8%	-1%
BLUE FORECAST	0:00Z	3	2%	−3%	−1%	−1%	1%	3%	2%	−1%
WRF-Meteotest	0:00Z	3	2%	-7%	-12%	-6%	24%	2%	14%	-1%

Table 18. Relative MBE Spain.

% MBE	Day	Cordoba	Granada	Huelva	COMPOSITE SPAIN

Mean GHI (W m-2)			443	409	407	420
Reference satellite model						
Persistence	0:00Z	1	0%	1%	0%	0%
ECMWF-OL	0:00Z	1	-2%	2%	0%	0%
CENER	0:00Z	1	2%	2%	-4%	-1%
WRF-UJAEN	0:00Z	1	9%	7%	4%	6%
HIRLAM	0:00Z	1	-6%	-16%	-7%	-10%
Persistence	0:00Z	2	0%	1%	0%	0%
ECMWF-OL	0:00Z	2	-3%	2%	-1%	0%
CENER	0:00Z	2	-1%	1%	-3%	-1%
WRF-UJAEN	0:00Z	2	9%	6%	5%	7%
HIRLAM	0:00Z	2	-5%	-17%	-10%	-12%
Persistence	0:00Z	3	0%	1%	0%	0%
ECMWF	0:00Z	3	-2%	1%	0%	0%
WRF-UJAEN	0:00Z	3	9%	7%	5%	7%
HIRLAM	0:00Z	3	-7%	-18%	-9%	-9%

Table 19. Relative MBE Canada.

[a] The WRF model was only run on a 2 month data subset and results were prorated using the other models as a template.

% MBE		Day	Egbert	Bratt's Lake	Varennes	Composite
Mean GHI (W m-2)			320	306	306	311
Reference satellite model						
Persistence	0:00Z	1	-4%	-8%	-6%	-6%
GEM	0:00Z	1	2%	2%	-2%	1%
ECMWF	0:00Z	1	4%	4%	0%	3%
WRF-ASRC[a]	0:00Z	1	2%		-1%	0%
GEM/ECMWF/WRF-ASRC[a]	0:00Z	1	2%		0%	1%
GEM/ECMWF	0:00Z	1	3%	3%	-1%	2%
Persistence	0:00Z	2	-5%	-9%	-6%	-6%
GEM	0:00Z	2	1%	1%	-1%	1%
ECMWF	0:00Z	2	1%	5%	-1%	2%
WRF-ASRC[a]	0:00Z	2	0%		6%	2%
GEM/ECMWF/WRF[a]	0:00Z	2	-1%		3%	1%
GEM/ECMWF	0:00Z	2	1%	3%	-1%	1%

Table 20. KSI * 100 US.

KSI*100		Day	Bondville	Boulder	Desert Rock	Fort Peck	Goodwin Creek	Penn State	Sioux Falls	Composite
Mean GHI (W m-2)			335	374	466	326	363	298	328	356
Reference satellite model			1.2	0.5	1.0	0.8	0.9	1.2	0.5	0.9
Persistence	0:00Z	1	1.7	3.1	2.0	2.0	1.6	2.0	3.1	2.2
GEM	0:00Z	1	3.4	3.9	1.6	2.5	3.2	4.3	3.9	3.3
GEM	12:00Z	1	3.4	3.2	1.6	2.1	3.7	4.0	3.2	3.0
ECMWF	0:00Z	1	2.6	5.6	2.6	3.0	3.3	4.0	5.6	3.8
NDFD	0:00Z	1	2.6	3.6	1.7	1.5	2.8	2.4	3.6	2.6
NDFD	12:00Z	1	3.1	3.6	1.6	0.8	3.5	2.8	3.6	2.7
WRF-ASRC		1	3.7				4.9	4.3		4.1
MASS	0:00Z	1			7.9		11	12		11
MASS	12:00Z	1			7.5		11	11		11
MAS-MOS	0:00Z	1			0.6		3.1	2.2		2.1
MAS-MOS	12:00Z	1			1.3		2.5	1.9		2.1

WRF-AWS	0:00Z	1			0.9		7.3	7.6		5.6
WRF-AWS	12:00Z	1			0.9		6.7	6.8		5.1
WRF-AWS-MOS	0:00Z	1			1.4		3.1	2.9		2.6
WRF-AWS-MOS	12:00Z	1			1.3		3.1	3.0		2.7
ARPS	0:00Z	1			8.2		11	11		11
ARPS-MOS	0:00Z	1			1.2		3.5	2.8		2.7
GEMS/ECMWF	0:00Z	1	3.3	4.8	2.2	2.8	3.4	4.2	4.8	3.6
Persistence	0:00Z	2	1.7	3.1	2.1	2.0	1.6	2.0	3.1	2.3
GEM	0:00Z	2	3.4	3.6	1.6	2.5	3.3	4.0	3.6	3.1
GEM	12:00Z	2	3.1	3.3	1.7	2.0	3.4	4.0	3.3	3.0
ECMWF	0:00Z	2	2.3	5.4	2.5	3.0	3.2	3.7	5.4	3.6
NDFD	0:00Z	2	3.1	3.8	2.0	1.2	3.0	2.3	3.8	2.7
NDFD	12:00Z	2	3.1	3.9	1.8	0.9	3.4	2.7	3.9	2.8
WRF-ASRC		2	3.0				5.0	3.9		4.0
MASS	0:00Z	2			7.6		12	11		11
MASS	12:00Z	2			7.4		12	10		11
MAS-MOS	0:00Z	2			1.0		3.1	3.3		2.7
MAS-MOS	12:00Z	2			1.1		2.5	2.9		2.3
WRF-AWS	0:00Z	2			0.9		6.5	6.7		5.1
WRF-AWS	12:00Z	2			1.0		6.1	6.6		4.9
WRF-AWS-MOS	0:00Z	2			1.5		2.9	3.1		2.7
WRF-AWS-MOS	12:00Z	2			1.3		2.9	3.2		2.6
ARPS	0:00Z	2			8.1		10	11		10
ARPS-MOS	0:00Z	2			1.1		3.0	2.9		2.6
GEMS/ECMWF	0:00Z	2	3.1	4.4	2.2	2.7	3.5	4.0	4.4	3.5
Persistence	0:00Z	3	1.8	3.1	2.1	2.0	1.6	2.1	3.1	2.2
ECMWF	0:00Z	3	2.2	5.3	2.3	3.0	3.2	3.4	5.3	3.5
NDFD	0:00Z	3	3.3	4.2	2.6	1.6	4.0	3.2	4.2	3.3
NDFD	12:00Z	3	3.6	4.2	2.1	1.2	3.8	3.0	4.2	3.2
Persistence	0:00Z	4	1.8	3.1	2.1	1.9	1.7	2.1	3.1	2.2
NDFD	0:00Z	4	3.7	4.5	2.4	2.4	4.2	3.1	4.5	3.5
NDFD	12:00Z	4	3.8	4.5	2.5	2.0	4.3	3.2	4.5	3.5
Persistence	0:00Z	5	1.8	3.1	2.1	2.0	1.7	2.1	3.1	2.3
NDFD	0:00Z	5	4.1	4.7	2.0	2.9	4.7	3.6	4.7	3.8
NDFD	12:00Z	5	4.2	4.7	2.3	2.3	4.5	3.2	4.7	3.7
Persistence	0:00Z	6	1.8	3.1	2.1	2.0	1.7	2.1	3.1	2.2
NDFD	0:00Z	6	4.6	4.9	1.7	3.0	4.8	4.4	4.9	4.0
NDFD	12:00Z	6	4.8	4.9	2.0	2.8	5.0	3.8	4.9	4.0
Persistence	0:00Z	7	1.8	3.1	2.1	2.0	1.6	2.1	3.1	2.3
NDFD	0:00Z	7	5.1	4.9	1.8	2.8	5.3	4.5	4.9	4.2
NDFD	12:00Z	7	5.1	4.9	2.0	2.9	5.2	3.9	4.9	4.1

Table 21. KSI*100 Canada.

[a] The WRF model was only run on a 2 month data subset and results were prorated using the other models as a template.

KSI*100			Egbert	Bratt's Lake	Varennes	Composite
Mean GHI (W m-2)			320	306	306	311
Reference satellite model						
Persistence	0:00Z	Day 1	3.2	3.4	3.6	3.3
GEM	0:00Z	Day 1	2.6	1.7	3.1	2.4
ECMWF	0:00Z	Day 1	2.1	1.7	2.2	1.9
WRF-ASRC[a]	0:00Z	Day 1	1.5		0.7	1.0
GEM/ECMWF/WRF-ASRC[a]	0:00Z	Day 1	1.9		2.2	1.7
GEM/ECMWF	0:00Z	Day 1	2.4	1.8	2.8	2.2
Persistence	0:00Z	Day 2	3.2	3.4	3.5	3.3
GEM	0:00Z	Day 2	2.5	1.8	2.8	2.3
ECMWF	0:00Z	Day 2	1.6	2.1	2.1	1.7
WRF-ASRC[a]	0:00Z	Day 2	0.5		1.7	1.1
GEM/ECMWF/WRF[a]	0:00Z	Day 2	1.9		2.3	1.8
GEM/ECMWF	0:00Z	Day 2	2.3	2.0	2.7	2.2

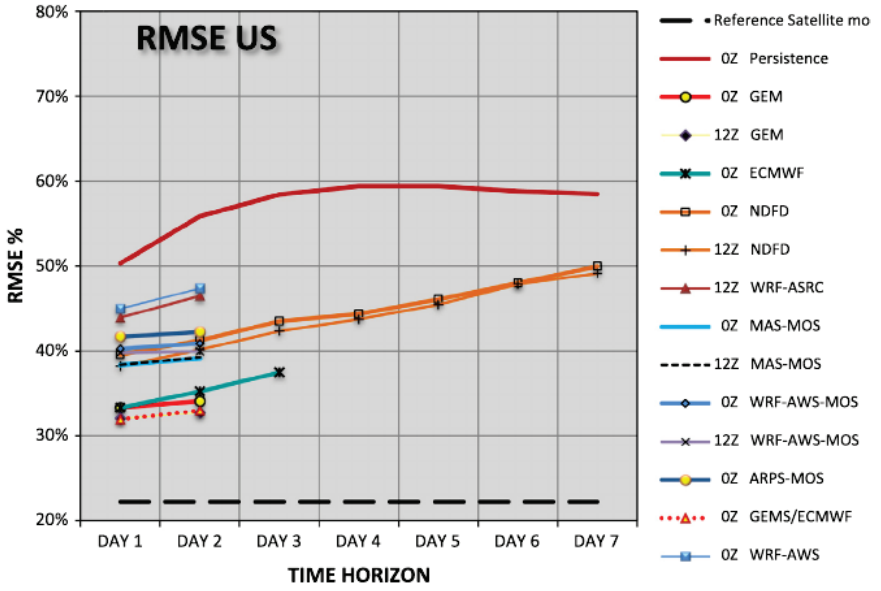

Figure 1. Composite RMSE as a function of prediction time horizon, United States.

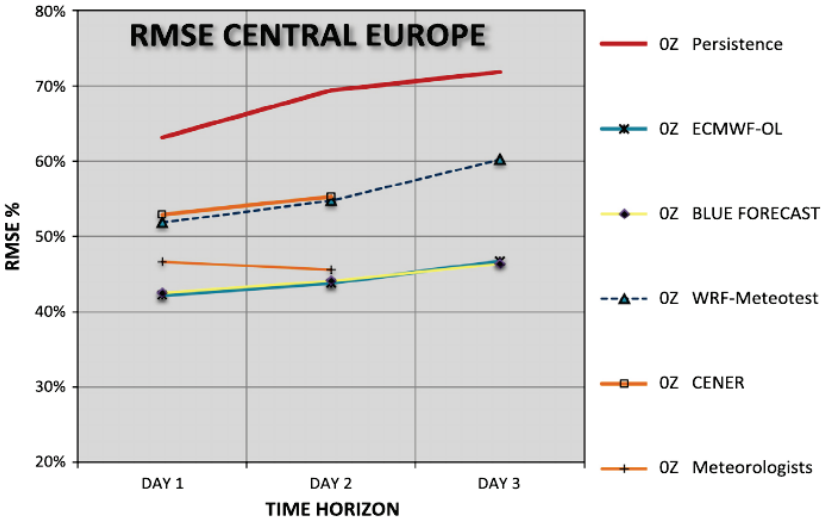

Figure 2. Composite RMSE as a function of prediction time horizon, Central Europe.

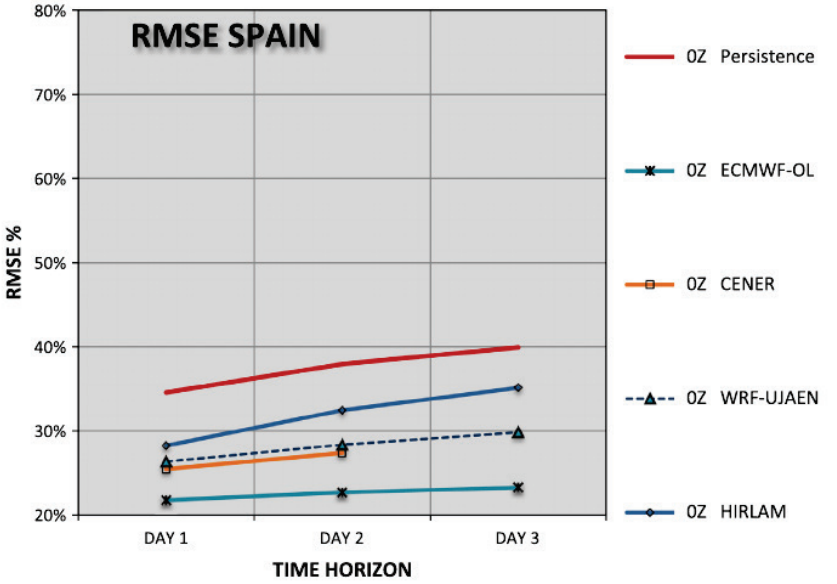

Figure 3. Composite RMSE as a function of prediction time horizon, Spain.

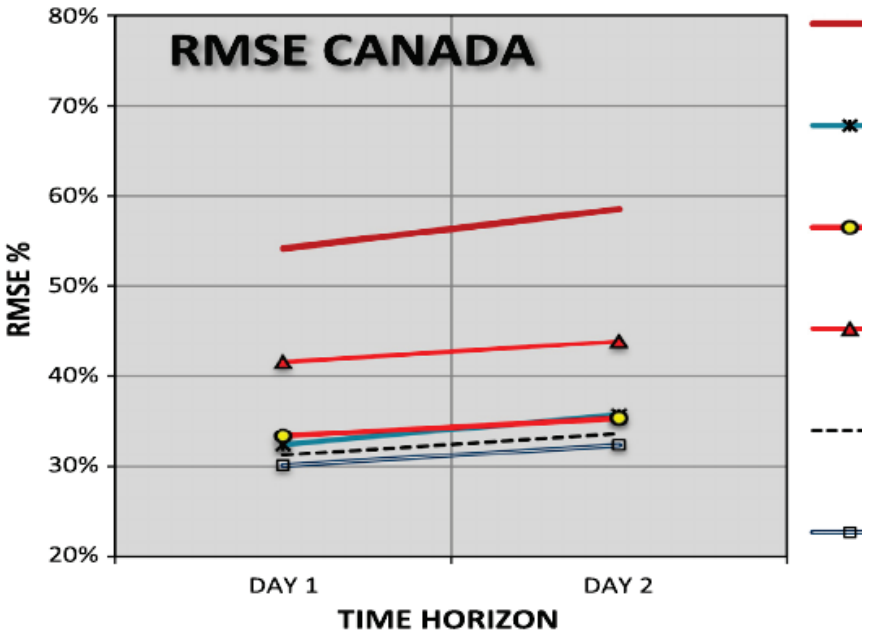

Figure 4. Composite RMSE as a function of prediction time horizon, Canada.

A comparison of the 00:00Z and 12:00Z runs (Figs. 1 and 5) shows slight advantages for the later 12:00Z runs for both RMSE and MAE (Figs. 1 and 5) as expected.

Almost all forecast models considered here outperform the persistence forecasts in terms of RMSE (Fig. 1) and MAE (Fig. 5), thus passing the basic test that confirms the skill of these forecasts with respect to trivial models. Exceptions are some pre-MOS models in the US evaluation (MASS and AEPS, see Tables 8, 12 and 16, not included in Figs. 1, 5 and 9). RMSEs and MAEs for persistence fore- casts are significantly larger for Day 2 than for Day 1, while for the other forecast models the increase in these error metrics is fairly modest.

There is a considerable variation of accuracy in terms of RMSE and MAE for the different sites and climates in the US (Tables 8 and 12),

where in the following only models available for all sites are considered in the discussion. For an arid climate (Desert Rock, US) with many sunny days, relative RMSEs in the range of 20–25% for Day 1 forecasts, are considerably smaller than for the other sites for all investigated models, where the RSME values exceed 30%. Largest Day 1 RMSE values between 38% and 48% are found for Penn state with the lowest mean irradiance. Persistence shows a similar trend ranging from 29% for Desert Rock to 65% for Penn State. Forecasts' skill with respect persistence measured by the MSE skill score is lower for Desert Rock (e.g. for Day 1 forecasts: MSE skill score of 0.52 for GEM0Z and 0.37 for NDFD0Z) than for Penn State (for Day 1 forecasts: MSE skill score of 0.65 for GEM0Z and 0.52 for NDFD0Z).

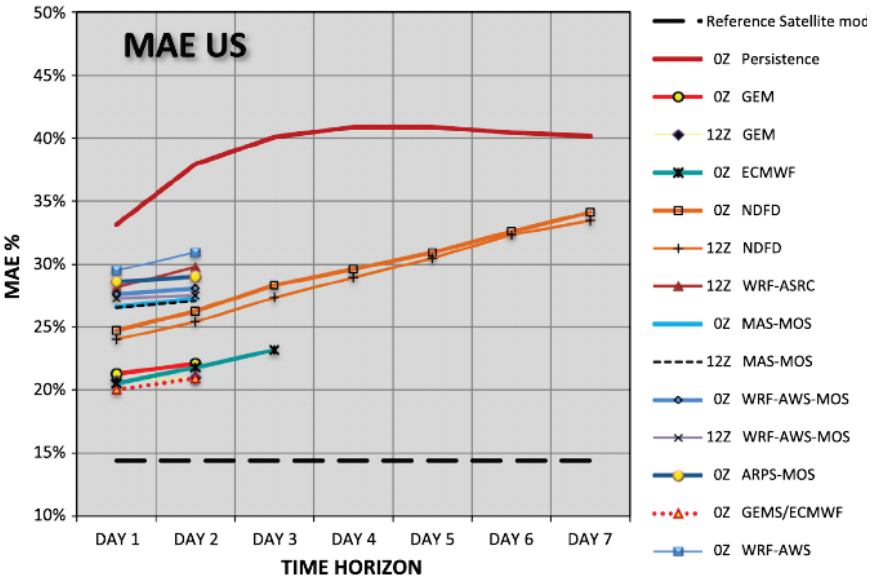

Figure 5. Composite MAE as a function of prediction time horizon, Unated States.

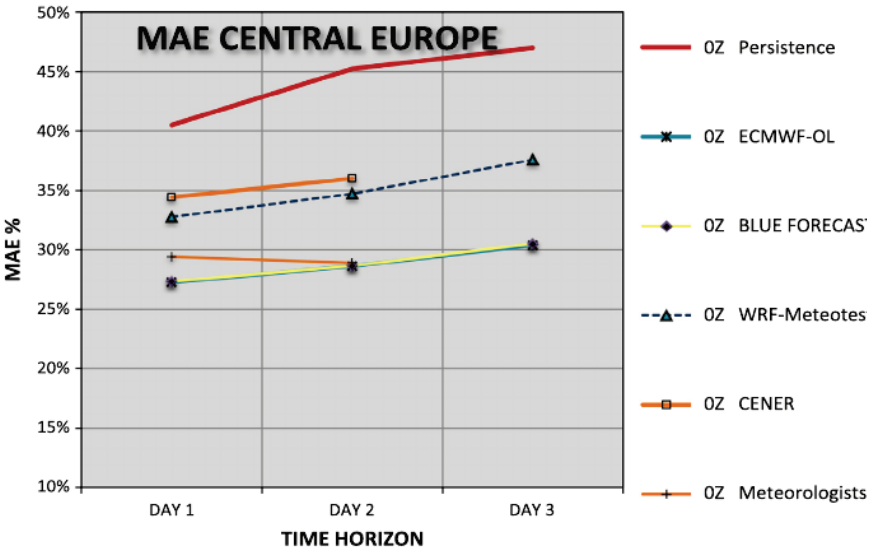

Figure 6. Composite MAE as a function of prediction time horizon, Central Europe.

Extending the model comparison from US to Canada (Figs. 4 and 8) and Europe (Figs. 2, 3, 6, and 7), the finding that ECMWF based irradiance forecasts show a higher accuracy than irradiance forecasts with WRF and the other investigated mesoscale models is confirmed. For Canada, like for the US, the performance of the Canadian GEM model is similar to the performance of the ECMWF model. For the Central European evaluation (Figs. 2 and 6) the GFS-based statistical method BLUE FORECAST performs similarly to the ECWMF based forecasts. Good results were also achieved with a method using cloud cover forecasts by meteorologists, as shown in the evaluations for Austria (Tables 9 and 13). Especially for local weather phenomena, such as fog or orographic effects, this approach may be advantageous (see also Traunmüller and Steinmaurer, 2010). However, this method is restricted to areas well-known by the experts interpreting and combining different forecast models.

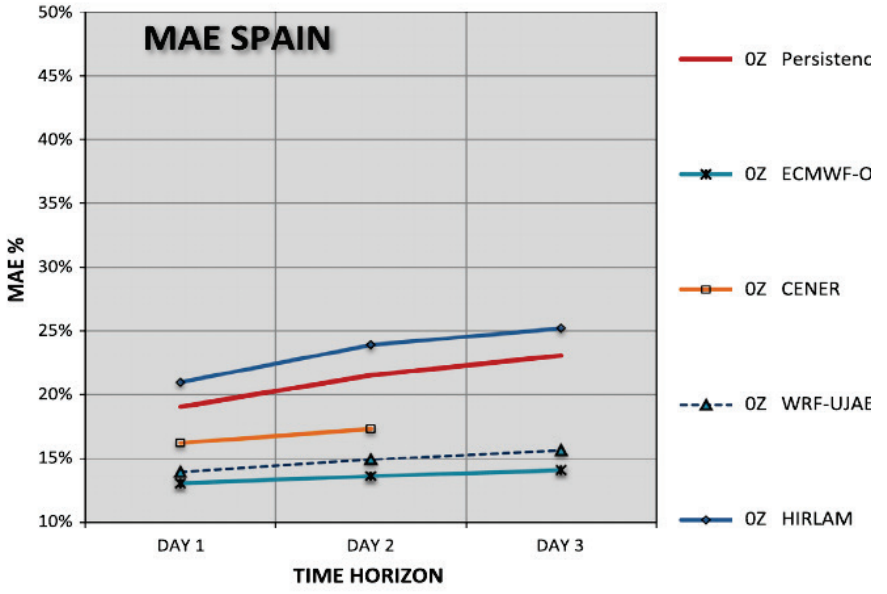

Figure 7. Composite MAE as a function of prediction time horizon, Spain.

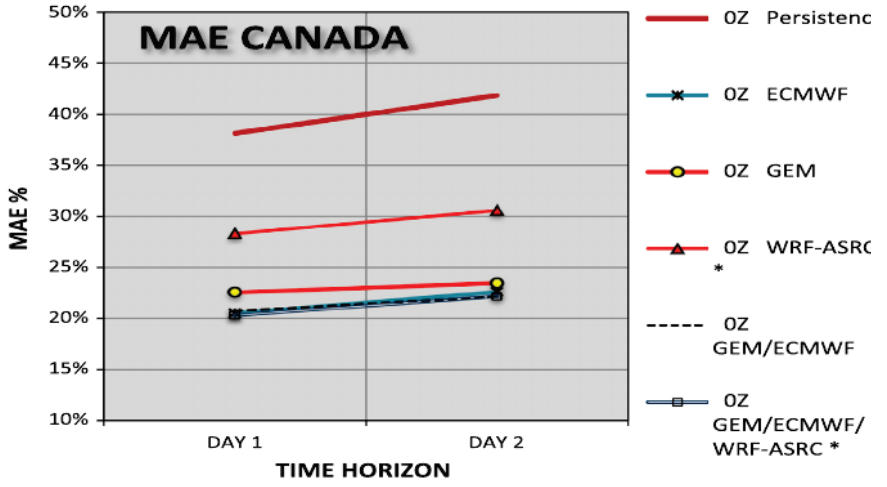

Figure 8. Composite MAE as a function of prediction time horizon, Canada.

When looking at the inter-comparison between WRF and the other two mesoscale models in Europe (Figs. 2 and 3), it has to be considered that both WRF-meteotest and WRF – UJAEN did not include any adaptation to measured data, while the SKIRON based forecasts provided by CENER, showing a similar performance to WRF in terms of RMSE, and HIRLAM based forecasts included a statistical postprocessing. This suggests that without post-precessing applied, forecasts with SKIRON and HIRLAM would show higher errors than the forecasts processed with WRF.

In addition to the evaluation of the single forecast models, a combination of some of the forecasts was investigated for the North American sites. The simple averaging of the two best performing models – ECMWF and GEM – does slightly better than individual models in both the US and Canadian evaluations (Figs. 1 and 4). Furthermore, as shown in Fig. 4 and Table 11 for the Canadian sites, the average of the WRF, ECMWF and GEM models also outperforms the individual models in terms of RMSE and MAE, and outperforms the ECMWF/GEM combination even though the WRF model has higher RMSEs and MAEs than the other two models. Forecast errors of the different models are not fully correlated and partly compensate each other. These observations indicate that combining independently run forecast models is a worthwhile option for improving forecast performance.

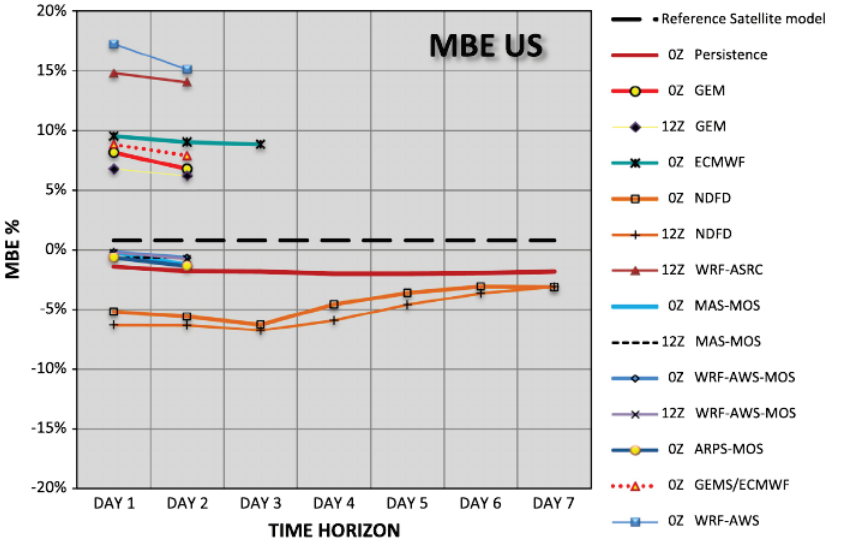

Figure 9. Composite MBE as a function of prediction time horizon, Unated States.

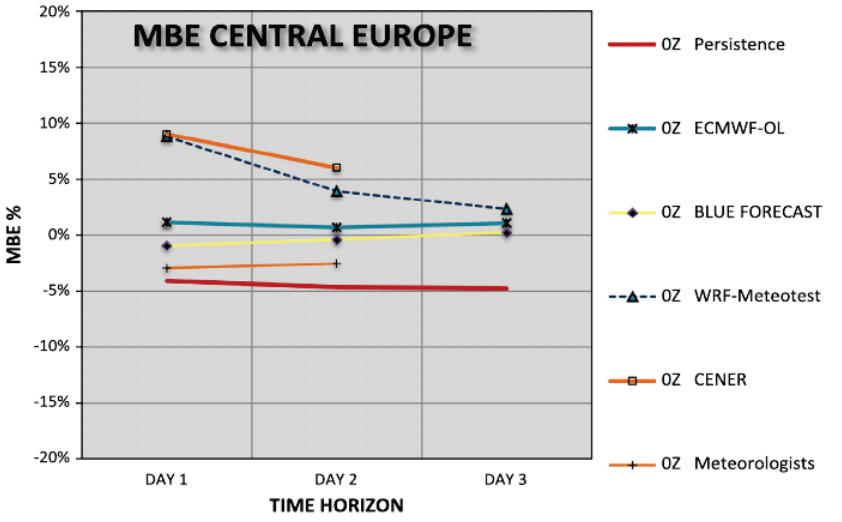

Figure 10. Composite MBE as a function of prediction time horizon, Central Europe.

With respect to the comparison of forecast performance for the different investigated regions we found lowest RMSE values in the range of 20% to 35% for the Mediterranean region Southern Spain (Fig. 3). For the Canadian stations with a humid continental climate, RMSE values between 30% and 45% are found (Fig. 4). For the US stations located in different climates (arid, sub-tropical, semi- arid, continental), RMSE values show a strong variation from station to station. All site-composite RMSE values for the US (Fig. 1) are similar to Canada. For the Central European stations with mostly continental climate and some alpine stations included average relative RMSE values range from 40% to 60% (Fig. 2).

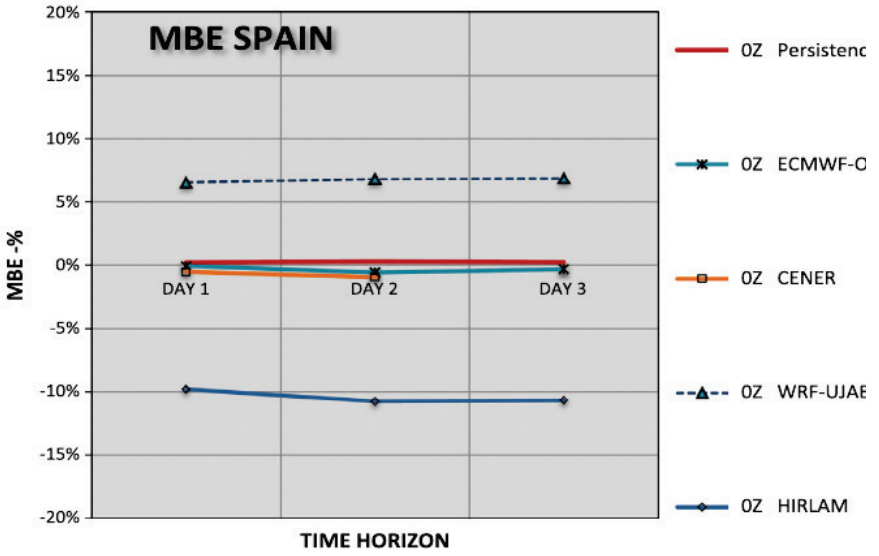

Figure 11. Composite MBE as a function of prediction time horizon, Spain.

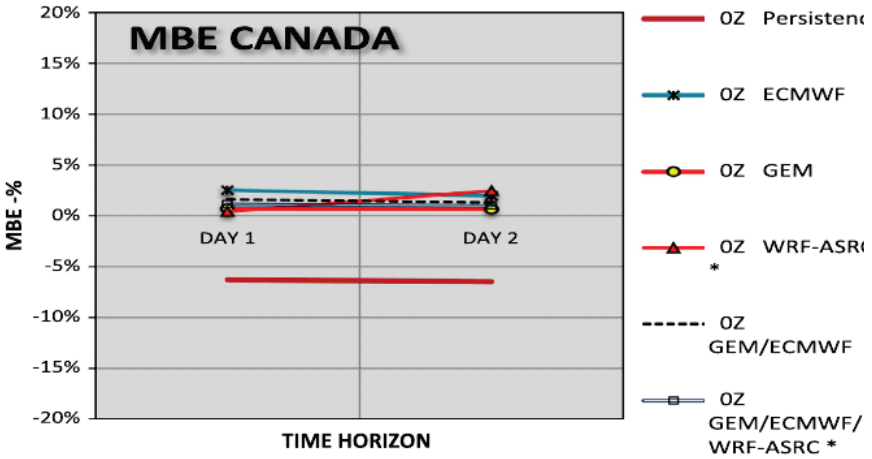

Figure 12. Composite MBE as a function of prediction time horizon, Canada.

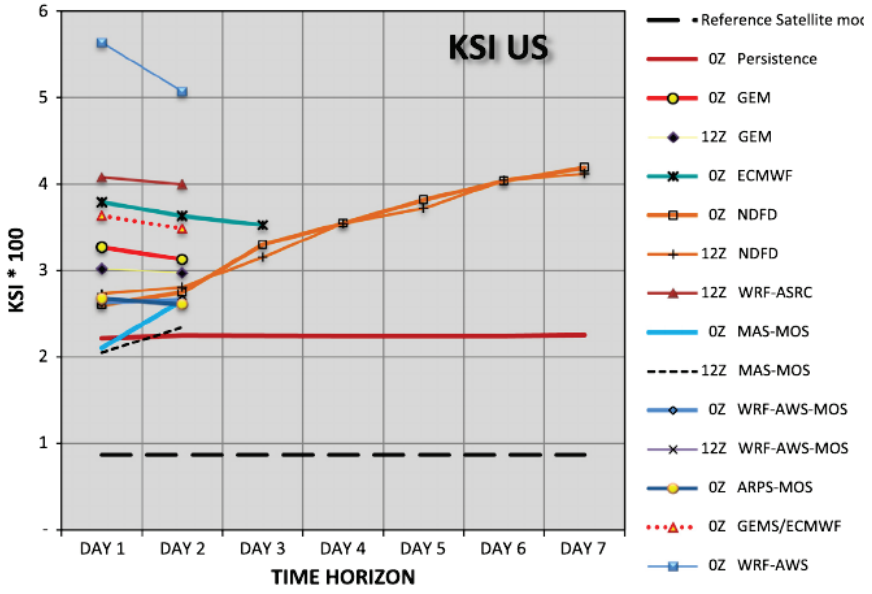

Figure 13. Composite KSI*100 as a function of prediction time horizon, Unated States.

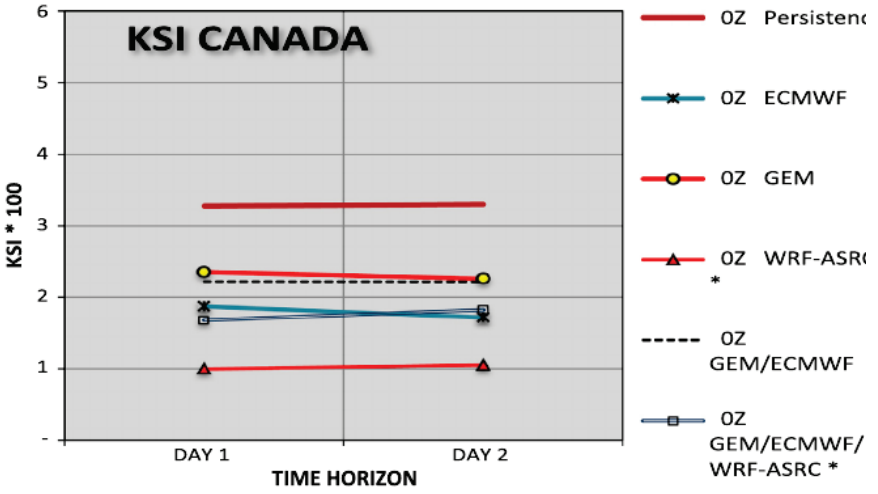

Figure 14. Composite KSI*100 as a function of prediction time horizon, Canada.

4.5 Conclusions

We have presented three validation studies comparing NWP based irradiance multi-day forecast for the US, Canada and Europe. The focus of the comparison was on the end-use accuracy of the different models including global, multiscale and mesoscale NWP models as a basis and different postprocessing techniques to derive hourly site-specific forecasts ranging from very simple interpolation to advanced statistical postprocessing.

Two models are common to the three validation efforts – the ECMWF global model and the GFS-driven WRF mesoscale model that was run in different configurations by various forecast providers – and allow the general observation that the global-model ECMWF forecasts perform significantly better than the GFS-based WRF-model forecasts. This trend is observed for all sites and different climatic conditions.

All other investigated mesoscale models available either for the US or for Europe showed even higher forecast errors than WRF. The potential of MOS to improve fore- casts with large systematic deviations was shown for some of the mesoscale models in the US. A forecast performance similar to the ECMWF forecasts in North America was achieved with the Canadian GEM model and in Central Europe with a statistical tool based on GFS forecasts. Furthermore, it was found that simple averaging of models' output tends to perform better than individual models.

Currently, major research efforts are spent on irradiance forecasting, driven by the strong need for reliable solar power forecasts which is arising from the continuously increasing amount of solar power installed in many countries. Weather services and research groups are working on improving cloud parameterizations and radiation schemes in NWP models and investigating the use of ensemble prediction systems and rapid update models. Another focus of current research is the application of intelligent statistical methods like machine learning to improve or combine the output of NWP systems. Accordingly, evaluation and comparison of different approaches for irradiance forecasting will be continued and new comparison studies will reflect the new developments in this field.

Acknowledgements

This work was supported by NREL under Contract No. AGJ04029001 as a contribution to IEA SHC Task 36. Financial support for the comparisons at Canadian sites was provided by Natural Resources Canada through the ecoENERGY Technology Initiative, which is a component of ecoACTION, the Canadian government's actions to- wards clean air and greenhouse gas emission reductions.

Chapter 5

Evaluation of DNI forecast based on the WRF mesoscale atmospheric model for CPV applications

***Lara-Fanego V**, Ruiz-Arias J A, Pozo-Vázquez D, Gueymard C A, Tovar-Pescador J (2012). Evaluation of DNI forecast based on the WRF mesoscale atmospheric model for CPV applications. Am Ins Phys Conf Proc 1477(1):317–322. doi:10.1063/1.4753895*

5.1 Introduction

Technologies for harnessing the solar resource have experienced a significant development in re-cent years. Their future looks even more promising. The International Energy Agency expects that, ac-cording to a reference scenario, the world's installed solar power capacity will increase from 14 GW in 2008 to 119 GW in 2035, with a 8.3% average annual increase (IEA, 2011). Therefore, the challenge for the next few years is to achieve a high level of development and integration, to make this resource competitive compared to traditional sources of energy, or even to more established renewable sources, like wind. A major effort is being made in this regard (Szuromi, *et al*., 2007). To achieve this goal, a key aspect concerns the resource itself (technology aside). The safe and optimal integration of large-scale solar electric power production into the energy grid of any country depends on the knowledge of the solar production capacity, which in turn is directly related to the available resource.

An important intrinsic characteristic of solar radiation is its very high variability over space and time, itself directly dependent on

weather characteristics. This intermittency in the resource makes a solar plant's operation and management particularly difficult. It also makes solar production troublesome for grid system operators, since it is hardly controlled and may not be available when it would be of greatest value (IEA, 2011). This ultimately translates into incremental exploitation and integration costs. There-fore, prior knowledge of the available resource of the near future is essential. Previous experience with the wind energy sector has shown that accurate fore-casts play a key role toward the successful integration of variable energy sources.

CPV systems use the beam component of solar radiation—or direct normal irradiance (DNI)—as their energy source. DNI is primarily affected by clouds, aerosols, and water vapor. Clouds are normally the principal factor affecting the incident solar radiation at the earth's surface, since they are most often completely opaque to DNI. In contrast, aero-sols are most influential under cloudless conditions. The uncertainty in the determination of the physical parameters associated with these atmospheric constituents is the main source of error in DNI predictions. This study focuses on how the latter is affected by cloudiness forecasts.

Numerical weather prediction (NWP) models have been proved to be powerful tools for solar radiation forecasting (Lorenz *et al*., 2009a; Lara-Fanego *et al*., 2011). One particular tool that is widely used by the research community is the Weather Research and Forecasting (WRF) model (Skamarock *et al*., 2008). WRF, like other NWP models, has a wide range of physical parameterizations, providing the possibility to achieve high spatial and temporal resolutions. It is commonly assumed that the higher the resolution, the better the physical description and results will be. This should apply, for instance, to the representation of processes that lead to the formation of clouds. In turn, high resolutions are computationally very expensive. Therefore, an optimal spatial resolution may exist in solar forecasting.

This contribution evaluates the role of the WRF model's horizontal spatial resolution in the reliability of the DNI forecasts that it can (indirectly) generate. Additionally, the intentional use of spatial

averaging of the gridded WRF-derived solar field to improve the model's accuracy is evaluated. The methodology applied is described first. A description of the forecast results is presented in a second step.

5.2 Methodology

5.2.1 Observations and evaluation procedure

This study is conducted for the Andasol Solar Thermal Power Plant (37.228º N, 3.069º W; 1100 m.a.s.l.), Fig. 1. Ten-minute DNI measurements are collected with an RSR2 radiometer. This instrument is well maintained and calibrated. Data for the 12-month period 01/12/2009 to 30/11/2010 were first corrected for spectral effects, and finally filtered with a series of quality control tests. For this study, irradiance values corresponding to solar zenith angles above 85º were filtered out to avoid the high measurement uncertainties associated with low-sun conditions. The original 10-minute data were also averaged to obtain hourly values. From a climatological standpoint, 2010 was an exceptionally rainy—and therefore cloudy—year.

Two forecast horizons are studied separately here: hours 1–24 (day 1, or "day ahead"), and hours 25–48 (day 2). Sky conditions are characterized by the clearness index (kt) to separate clear-sky ($0.65 < kt$), cloudy ($0.4 \leq kt \leq 0.65$) and complete overcast ($kt < 0.4$) conditions. The forecast reliability is objectively evaluated in terms of mean bias error (MBE) and root mean squared error (RMSE), and their relative values (in %), obtained by normalization to the mean of the ground measurement irradiance for the considered period. The forecast errors (residuals) are calculated as the difference between forecasted values and observations. A positive MBE is thus indicative of an overestimation of the modeled DNI. Finally, the trivial persistence model is used as the skill reference model.

5.2.2 WRF configuration

The model's domain configuration is represented in Fig. 1. The dynamical downscaling is driven by the use of four nested domains with progressively decreasing horizontal resolutions of 27, 9, 3, and 1 km for the outermost to innermost domains. The atmospheric column is decomposed into 28 vertical levels. The ECMWF/IFS weather forecasts are used as initial and boundary conditions. For each day of the evaluation period a WRF (ARW, version 3) simulation of 60 hours is run. The first 12 forecasted hours are considered as model spin up, and discard-ed. The next 48 hours are evaluated independently for the first and second 24-hour periods. The WRF parameterizations are selected based on (Ruiz-Arias *et al*., 2009). In particular, Dudhia's scheme is used for the shortwave radiation parameterization.

5.2.3 DNI derivation

NWP models (e.g., WRF) do not usually provide DNI as an output variable. Therefore, DNI needs to be derived in a post-processing step based on WRF's comprehensive forecasted information and an external radiative model (Lara-Fanego *et al*., 2011; Ruiz-Arias *et al*., 2011). Recent studies have rigorously analyzed the performance of a large set of different radiative models (Gueymard, 2010, 2012a). Results showed that meteorological radiative models achieve a very high performance in DNI estimation under clear-sky conditions. In contrast, statistical/empirical models show lower performance but more simplicity. Since this work focuses on the WRF aspects related to cloud modeling, the simplest way to derive DNI is preferred; nevertheless better results can be expected by using the first kind of radiative models mentioned above (Ruiz-Arias *et al*., 2012). An empirical statistical model (Ruiz-Arias *et al*., 2010c) is simply applied here to obtain DNI from the WRF global horizontal irradiance (GHI) forecasts.

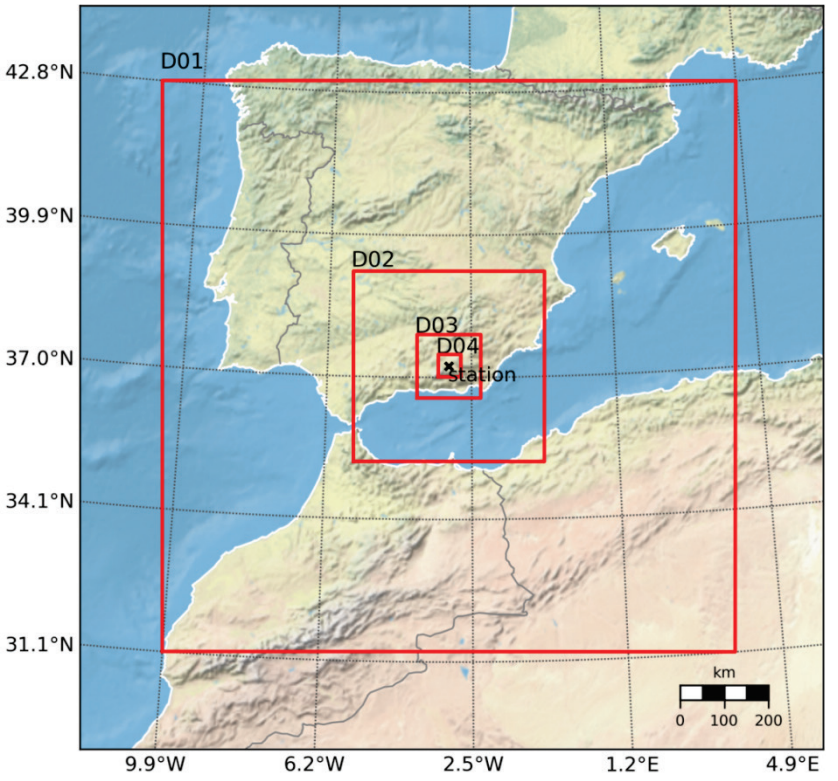

Figure 1. Configuration of the WRF domains. The spatial resolutions are 27 km, 9 km, 3 km and 1 km for do-mains D01, D02, D03 and D04, respectively. The radio-metric station is located at the center of all domains.

5.3 Results

5.3.1 Dependence on horizontal resolution

Figures 2 and 3 show the performance results for the day-ahead forecast horizon over the whole 12-month period. Extremely large errors are obvious for complete overcast conditions (Fig. 2). Since DNI is very sensitive to the presence of clouds, any misrepresentation of cloudiness—in either space or time—in the model's predictions may cause significant errors. The high variability of cloud type and cloud amount enhances this effect. The most important result is that the errors (RMSE and MBE) increase with spatial resolution, contrarily to what would have been expected. An interesting exception is the case of

clear-sky conditions, under which the RMSE decreases (Fig. 3). These results mean that clouds are not better resolved at higher spatial resolution by the model, at least in terms of their effect on solar radiation. In contrast, under cloudless conditions the topographic effects, which are better resolved at finer resolutions, become more relevant. This is of particular interest to CPV, since this technology can normally be installed on uneven terrain.

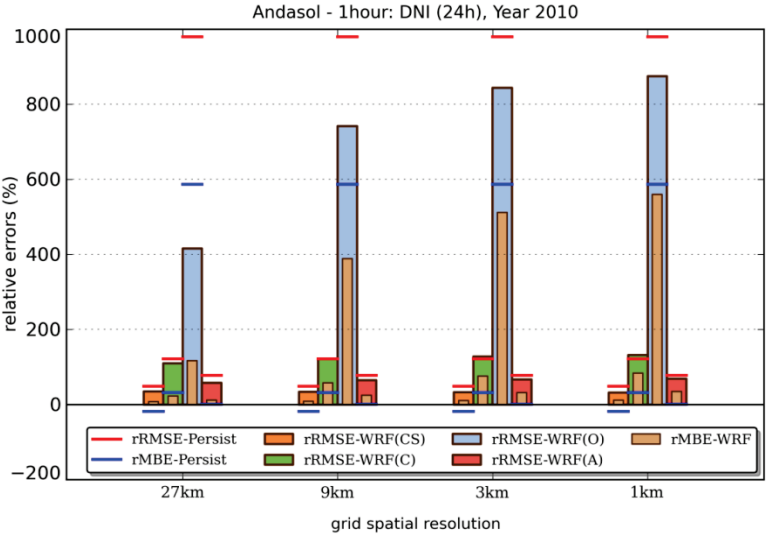

Figure 2. Relative errors of the DNI day-ahead forecasts over the complete data period. For RMSE, the orange color corresponds to clear skies (CS), the green color to cloudy conditions (C), the blue color to complete overcast (O) and the red color to all-sky conditions (A). The inner bars indicate the MBE. The persistence model errors are indicated by horizontal segments: red for RMSE and blue for MBE.

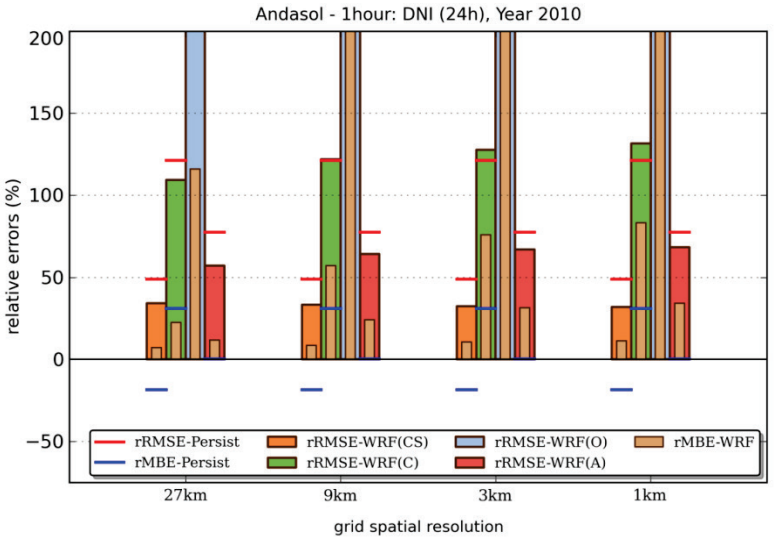

Figure 3. Same as Fig. 2 but with a different scale for better visualization.

MBE is found positive in all cases. Under clear-sky conditions, WRF tends to overestimate both GHI and DNI, most probably because of a too low aerosol optical depth (AOD) in the model. This overestimation also occurs under cloudy conditions, which means that the model predicts less cloudiness than will occur in reality. Nevertheless, it is important to note that the WRF-based DNI forecasts most generally outperform the persistence model.

The dependence of the forecast performance on the forecast horizon is an important topic, since energy sales in the daily electricity market must be made ≈24 hours early. From this standpoint, the second day of forecast may be more important than the first one.

Figure 4 compares the performance results of DNI forecasts according to time horizon (24h vs. 48h), for all possible spatial resolutions. These results show that the DNI forecasts are remarkably stable over time. Both forecast horizons exhibit similar dependence on spatial resolution. It should be pointed out that these results correspond to raw model outputs.

Better performance is achievable in DNI fore-casts with post-processing, as demonstrated in the next section. Moreover, the exceptionally rainy conditions that prevailed during the study period can explain a part of the large errors.

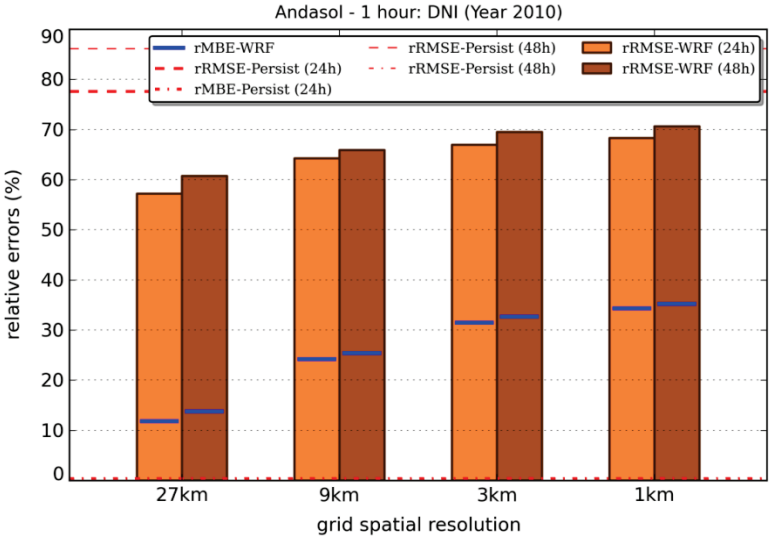

Figure 4. All-sky performance results for the 24h and 48h forecast horizons. Bars correspond to RMSE, and blue segments to MBE. The red dashed lines correspond to the persistence model's RMSE (its MBE is ≈0).

5.3.2 Spatial averaging (post-processing)

The inability of the WRF model to simulate high-frequency spatial-temporal cloud changes (thus, DNI changes) can be worked around by applying a spatial filtering algorithm. In particular, DNI values are gathered from the model's grid by averaging the DNI forecasts over windows of varying incremental size, with the target station always at the center of these windows. Spatial averaging of the predicted solar radiation is a commonly used filtering technique to remove the high-frequency variability in forecasts. Figure 5 shows the performance results for the 48h forecast horizon.

Both MBE and RMSE are reduced by the spatial averaging process, for all initial spatial resolutions. This improvement depends on spatial resolution. The best results are obtained again for the coarser initial resolution (27 km) using an averaging window of 100x100 km. These results are in agreement with previous findings (Lorenz *et al*., 2009a).

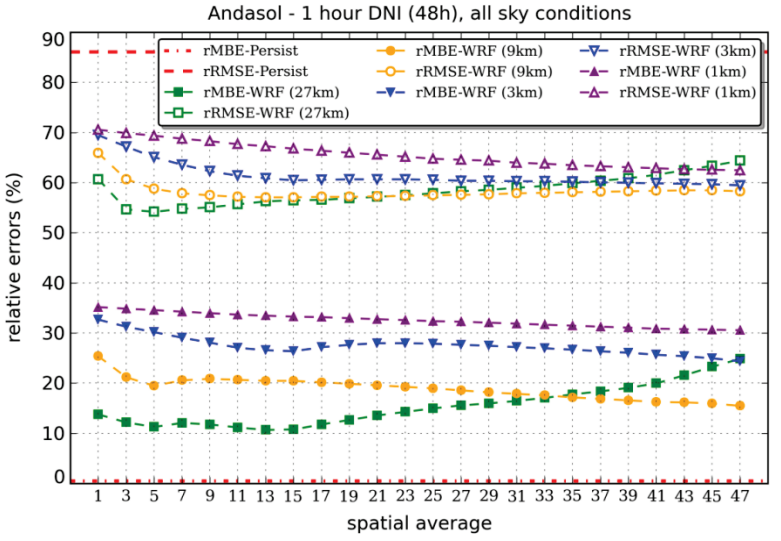

Figure 5. RMSE and MBE for the 48h forecast horizon and a spatial average of surrounding points from 1 (nearest grid point from the station location) to 47 (square side of farthest grid points).

Figure 6 shows how sky conditions (clear vs. cloudy) affect the performance of the spatial averaging. Under clear skies, the post-processing step is not as effective as it is under cloudy conditions.

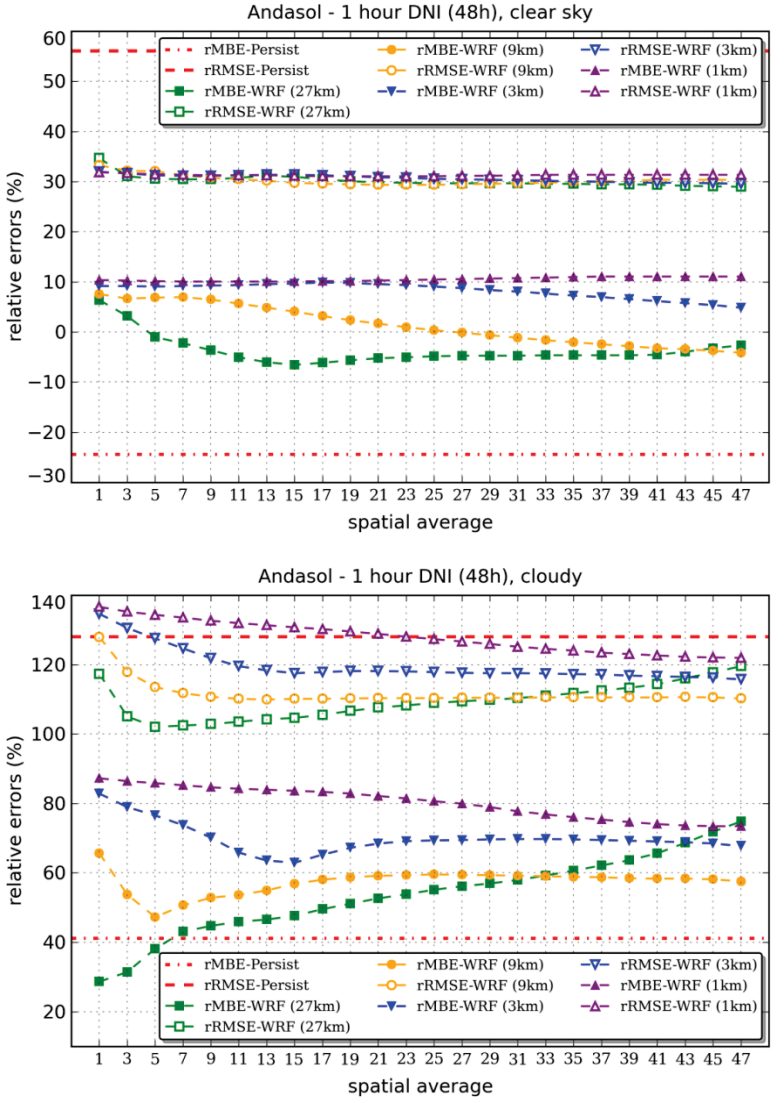

Figure 6. Same as Fig. 5 but for clear-sky (top) and cloudy (bottom) conditions.

5.4 Conclusions

DNI forecasts based on the WRF model are analyzed with respect to the performance impact of two important aspects of the model's optimization: horizontal spatial resolution, and spatial averaging of the model's outputs. A complete 12-month period of DNI forecasts are evaluated against measurements collected at a solar power plant in the southeast part of the Iberian Peninsula. Results show that an increase in spatial resolution does not enhance the reliability of the WRF-based DNI forecasts, except under clear-sky conditions. Therefore, clouds are not resolved better (in space or time) at higher resolution, from the stand-point of solar radiation forecasting. It is possible that the independent column approximation that is usually assumed in this kind of model limits the performance of solar radiation forecasts at high spatial resolutions. Another source of error results from the use of a simple empirical model to derive DNI from global irradiance.

Spatial averaging (in post-processing) notably reduces errors. The best results are obtained for the coarser domain (27 km) and spatial averaging of approximately 100x100 km. Cloud representation at high spatial resolution is a big issue for WRF (or for any NWP model). Further research is needed to improve cloud parameterizations in WRF, and how DNI can be directly derived from them. For CPV applications, additional research should investigate the forecasting of spectral irradiance based on advanced WRF radiation schemes possibly coupled with external radiative models providing spectral information (e.g., SMARTS).

Acknowledgments

This study was financed by the Consejería de Innovación, Ciencia y Empresa (CICE) of the Junta de Andalucía (Spain), under Project P07-RNM-02872, the Spanish Ministry of Science (Project CGL2011-30377-C02-01) and FEDER funds. The authors would like to thank

Milenio Solar (Solar Millenium AG) for providing the ground observations.

Chapter 6
Summary and conclusions

6.1 Introduction

The research conducted in this work aimed the enhancement of the current solar resource knowledge for the promotion and development of solar industry applications. In particular, two types of research analyses are carried out. Firstly, the focus is on improving the solar resource assessment for feasibility studies in solar power plant projects. To this end, in Chapter 2, a novel method for obtaining a representative year for the characterization of solar irradiance at multi-year scales is presented and evaluated. Secondly, the focus is on solar energy integration issues, once the solar power plant is operating. To this end, the development of accurate solar radiation forecasting systems is mandatory. In Chapter 3, the performance of the WRF NWP model for solar irradiance forecasting is assessed. Next, Chapter 4 presents a benchmark comparison of solar radiation forecasts derived with WRF and other NWP models. Finally, based on the results of the previous studies, an analysis of the influence of the horizontal spatial resolution in the skill of solar radiation forecasts is presented (Chapter 5).

6.2 TSY generation for improved solar resource assessment

Solar resource assessment can be regarded as the corpus of methods, techniques and their applications for the long-term characterization of solar radiation at a location of interest. One key ingredient of solar resource assessment for the solar energy industry is the so-called Typical Meteorological Year (TMY). Despite the foundations of TMYs are not fully accepted by the solar energy scientific community, they

are widely used and required by the solar energy industry, mainly for the design and bankability analysis of solar projects. In particular, the use of TMYs has become a standard for typifying the expected energy production of the solar plants. Nevertheless, the methods for obtaining TMYs are varied, due fundamentally to the lack of scientific consensus. This causes an unwanted inconsistency, since the application of different methodologies to the same dataset could result in different TMYs. Ultimately, this increases the uncertainty in the solar resource assessment, which is just the opposite of what is intended. By enhancing the reliability of expected solar resource, risk surcharges can be reduced and hence effective costs of production are lessened. In this sense, standardization of the methods for deriving TMYs is a demand from solar industry, because it establishes a reference for the quality of data and helps to reduce the uncertainty.

In this context, Chapter 2 presents a novel method aimed at contributing to the standardization of Typical Solar Year (TSY) datasets. TSYs, like the most common typical meteorological years (TMYs), are annual time series artificially formed to characterize the solar resource of a location of interest. Despite both –TMY and TSY- can provide the same amount of information (see discussion below) in Chapter 2 the TSY is considered according to its simpler definition, that is, as a time-series of a single solar irradiance variable. The main objective of the method, referred to as the EVA method –an acronym of the Spanish words for seasonality and variability-, is to determine the accumulated monthly solar energy values whose annual sum is equal to the annual target value for a determined probability of exceedance. Thus, it is possible to select the corresponding months of the long-term time series closest to the estimated months. The EVA method provides the selection of the months based on a composite of weights that accounts for: i) the variability of the seasonal adjusted monthly distributions, ii) the individual monthly energy contribution of each calendar month respect to the total annual energy. The novelty of this method relies on, firstly, its statistical basis -by means of an original approach- and, secondly, in its solid analytical formulation. This provides robustness and reliability to the method, while facilitating its

algorithmic implementation (*vide infra* Annex A). In addition, the EVA method has the property of being valid for the analysis of any solar radiation component indistinctly, GHI or DNI depending on the solar technology. However, its most outstanding feature is its capability to generate TSYs for any expected scenario of solar resource availability. This includes the worst-cases of extremely low energy, which is of key importance for the bankability of the solar projects. Currently, there are few methods that allow this type of analysis, and some of them are proprietary. Finally, the method seems to be consistent regarding the determination of the uncertainty. In particular, this uncertainty can be easily quantified and combined with the total uncertainty value of the solar resource assessment by means of the Gauss law of error propagation. Nevertheless, additional research is needed to analyze further this issue. The determination of the uncertainty is fundamental for solar projects trustworthiness and hence for their bankability.

In Chapter 2 the main results of the assessment of the EVA method are also presented. In general, it is shown that EVA method performs well. The differences between the annual target value for a certain probability of exceedance and the corresponding annual value of the irradiance generated by the EVA-derived TSY are low. The differences between the annual irradiation values of the TSYs and the target annual percentile values are generally below 1%, for both GHI and DNI. In addition, the method preserves the long-term statistics when the constructed TSY of a determined probability of exceedance is set to the original higher time resolution of its constituent months –typically 1 hour.

Nevertheless, additional research should be conducted to extend the set of locations used to evaluate the method. In this sense, it would be important to analyze the feasibility of applying the EVA methodology to satellite data. This is mainly because in practice, by far satellite information is the most common source of long-term time-series of solar irradiance data for solar industry applications. In this sense, it is timely to compare the results of the application of the method using long-term time series of measured data and those derived from satellite.

Additionally, it would be very important to extend this research by analyzing the solar energy yield of a plant based on the expected scenarios of solar resource, as described by the TSYs derived from the EVA method. This would allow estimating the actual typical response of the plant under the conditions derived from the TSYs. In any case, it can be concluded that the particular objective proposed in this thesis concerning the solar resource assessment was successfully achieved.

Finally, it should be clarified that the EVA method does not provide the customary TMY, but the so-called TSY. It is important to remark that TSY does not have less –but slightly different- information than TMY. In this sense, TSY can be directly complemented with the required meteorological variables by simply adding their values the corresponding time records. The difference is that the traditional generation of TMYs makes use of the ancillary meteorological information by means of pre-established non-consensual weights, while the TSY only uses the main variable, that is, it sets the 100% of the weight to the main irradiance variable –GHI or DNI-. In some sense, the use of the adjective "meteorological" in TMY seems to be a little bit ambitious. Mainly, because there is not a clear evidence about the convenience of combining -with certain degree of influence- the secondary variables with the main irradiance variables –GHI and/or DNI-. Consequently, there is no consensus either on the required external variables and the weight that each one should have. Hence, for solar energy applications it seems to be advisable to adopt a method that only consider the fundamental irradiance variables. This would reduce the degrees of freedom in the determination of the uncertainty such as, for instance, the ones associated to the determination of the meteorological variables. In this regard, sometimes this can lead to issues in the application of the TMY concept in the solar industry. Despite both TMYs and TSYs are not recommended for production analysis, they are actually used to infer the expected yield of the plant according to the scenarios defined by them. In this way, in practical situations the interest of a solar plant promoter is to present the most attractive project to win a bidding process. To this end, the quantities of expected energy yield are determinant and, hence, the estimated

amount of energy. Since the tables of weights to be used for the combination of the variables to generate the TMYs are part of the knowhow of the company, they can be slightly manipulated to favor a more convenient result for the interests of the promoter. For example, by adding a little increment to the percentage of the weight associated to the temperature, or to any other variables; logically with extremely careful to not exceed the expectations too much. In the end, a set of different TMYs can be obtained for the same long-term dataset, and the promoter can choose the most convenient. In this sense, the use of TSYs instead of TMYs allows eluding these situations. In addition, the standardization of the methods will also help to avoid these unnecessary imprecisions, providing a unique solution for common use.

6.3 Performance of WRF solar radiation forecasting for solar energy integration

Given the chaotic nature of the weather, solar radiation is highly variable in space and time, and reliable forecast of solar power production is of paramount importance for the large-scale integration of greater shares of solar energy in the power generation and distribution systems. Solar power forecasting contributes to minimize the supply system risks associated to the solar power production fluctuations and to maximize the economic revenues by scheduling the power delivery according to the expected production and the market situation. In this context, nowadays, forecast of solar power production are a basic tool of Transmission System Operators (TSOs) for grid management. It is also fundamental for solar power plants operators for energy trading and plant operation and management. Among the methodologies for predicting surface solar radiation, NWP models stand out as the most powerful tool for forecasting horizons beyond about 5 hours. Their physical foundation allows them to provide comprehensive weather forecasts, maintaining the spatial and temporal coherence over large extended regions and short to medium-range periods. In particular, these models can provide solar irradiance forecast, as well as forecast of ancillary variables of interest for solar energy applications. The

spatial and temporal scales of the forecasts provided by these models are in the range of 10-25 km and 1-3 hours, in the case of the global circulation models, and only few kilometers and sub-hourly, in the case of the regional models. Within the latter type of NWP models, the WRF model stands out as certainly one of the most advanced. It has an extensive set of parameterizations that provide the model a great flexibility to be adapted for a specific task and to the geophysical characteristics of a particular region. It is also one of the most used models around the world, provided it is a community model being under continuous research and development. Hence, it is appropriate choice for the purposes of this research.

In general, NWP models are not specifically devised for solar energy applications. As a consequence, there are still few research works, such as the one attempted here, assessing the performance of this type of models on the specific task of solar radiation forecasting.

The starting point of this research is addressed in Chapter 3. In this initial study, a comprehensive evaluation of GHI and DNI forecasts reliability provided by the WRF model is conducted. The analysis is carried out by comparing the model forecasts against ground measurements from several radiometric stations located in the region of Andalusia (southern Spain). Time period of analysis is one year. The time resolution is the usual hourly-base, while the spatial resolution is 3 km. This high spatial resolution is a markedly difference compared to the coarser spatial resolutions achieved by the GCM -usually greater than 10 km-. The study includes different aspects affecting solar radiation forecasting (clouds, aerosols, etc.). In addition, no post-processing was applied in order to focus the analysis only on the model performance. In particular, the skill of the prediction was evaluated for different sky conditions, namely: complete overcast, cloudy, clear sky and all-sky. Seasonal and year-around independent analyses were conducted. Additionally, model skill for different forecasting horizons (24h, 48h and 72h) was assessed. Furthermore, unlike GHI, which is directly provided by the model, DNI was derived from the model outputs as it was not already an output variable included in the WRF

version 3.2.1, used in this study. The current versions of WRF already provide DNI and diffuse components. However, this was the first study that analyzed DNI forecasts derived from WRF in the context of solar energy applications. Nonetheless, today it is still not properly studied the relative reliability of DNI forecasts of the NWP models vis-à-vis DNI forecasts derived by means of an external model, such as all-sky separation models or radiative transfer models as the one used here. From a practical stand point this could be important, both for using the most reliable forecasts, as well as for having a valuable reference for DNI forecast assessment. Notwithstanding, the goal should be to enhance the NWP model performance.

As a result of the evaluation, it was found that, in general, the WRF model tends to overestimate solar irradiance, under all sky conditions and for all analyzed periods. For cloudy conditions, it was found that this was mainly due to the fact that the WRF model underpredicts the cloud amount. In the case of clear sky conditions, this result suggests that the AOD values used in the model underestimate the actual values. As it was expected, errors for DNI were found to be markedly higher than those obtained for GHI. This is due to the greater sensibility of DNI, overall to the presence of clouds, but also to the uncertainty of the aerosol load. Results also showed that the model performed generally better than the trivial persistence model, except in periods where the presence of clouds was more significant, when performance was similar. Thus, summer presents better results, showing that the presence of clouds is the most important factor in the estimation of the solar irradiance by far. In summary, it was concluded that cloud forecast is still a big issue regarding solar radiation forecasting. Therefore, substantial improvements about cloud representation in NWP models should be obtained in the future. This will help to obtain the stringent requirement of the solar energy industry.

The next step in the solar radiation forecasting analysis was the intercomparison of the approach proposed here, based on the WRF NWP model, against other approaches based on different regional and global NWP models. This analysis was conducted in an extensive

benchmarking exercise described in Chapter 4. This work established a reference framework that allowed evaluating the WRF solar radiation forecasts respect to the forecasts of other NWP models. The study was carried out only for GHI mainly because of the lack of DNI observations. It should be noted that, unlike the approach followed here, some models participating in this benchmarking exercise used post-processing procedure to enhance their performances. In this sense, this benchmarking study is also interesting in order to know the possibilities of improving the WRF model estimates. Indeed, results show that the post-processing has a great potential to enhance forecasts with large systematic deviations, as it is the case of WRF according to previous results obtained in Chapter 3. In general, it was concluded that the performance of GCMs was better than those of the regional models. Overall, the WRF model errors were significantly larger than the errors of the ECMWF model, which obtained the best results. It was concluded that one of the reasons was the low bias errors provided by the GCMs; unlike WRF, which showed a significant overestimation. Another interesting result of the benchmarking study suggested that the performance differences between the regional models and the GCMs has more to do with the regional models themselves than with the initial and boundary conditions used -provided by GFS model in this study-. In this regard, additional detailed studies are required to confirm this assertion. However, contrary to expected, it was found that the horizontal spatial resolution did not play an important role in the GHI forecasting reliability. This issue motivated a further research, which was presented in Chapter 5.

Following results obtained in Chapter 4, the next step in the research plan was to analyze the role of the horizontal spatial resolution. This is an important aspect concerning solar resource forecasting from the stand point of model-output's optimization. Additionally, the effect of the spatial average of solar radiation in the reliability of the WRF solar irradiance forecast was also assessed. This simple post-processing method is extensively used, because it usually reduces the absolute error, which is directly related to the deviations of the expected power production. Both issues were analyzed in Chapter 5. This study was

conducted for both components: GHI and DNI, although only results for DNI were shown. However, the main conclusions are valid for both variables.

On the one hand, the role of horizontal spatial resolution was examined in order to clarify its effect on the reliability of the solar irradiance forecast. Contrary to what would be expected, the results showed that an increment in spatial resolution does not necessarily enhance the reliability of the WRF-based DNI forecasts. Only under clear sky conditions the performance is better in the case of higher spatial resolutions, when the topographic features play a major role. This suggests an important role of clouds in this trend. On the other hand, in the second part of Chapter 5, an ad-hoc post-processing method is applied to the WRF raw outputs. It consisted of a smoothing filter based on spatial average –or spatial aggregation- of solar irradiance. Results clearly showed that the spatial averaging of solar irradiance notably reduces the forecasting errors, improving the WRF performance. Moreover, the results were more reliable for coarser resolutions and for a spatial aggregation covering around 100×100km. Similar results were obtained in other studies using other NWP models.

From the results of chapters 4 and 5 it was clearly concluded that a higher spatial resolution does not guarantee a better NWP model performance regarding solar irradiance forecasting. This does not mean that the representation of cloudiness is better at coarser model spatial resolutions. There are other aspects that have to be taken into account, such as the model representation of clouds itself, or even the evaluation process itself, since the modelled solar irradiance values represent the average over the entire grid cell extension, which is then compared to point-wise observations. Another important factor to be considered is the so-called double-penalty error, which accounts for the fact that clouds have to be correctly represented both in space and time in order to not produce penalties in the error scoring. That means that clouds should be in the exact position in the precise moment with respect to the observations gathered in the radiometric station. This is related with the effectiveness of spatial averaging at reducing the random

component of error -i.e. MAE and RMSE- since spatial averaging reduces the impact of double-penalty errors. This is because, somehow, it is as if increasing the size of the clouds prevents many errors in which the cloud is generated close to the validation point but not at the exact position. Alternatively, in general, the goodness of spatial averaging at reducing random errors is closely related with the spatial autocorrelation of solar radiation in the neighborhood of the evaluation site in such a way that the lower spatial autocorrelation, the smaller averaging size is required to achieve the same reduction of random errors. Additionally, there exists an averaging size for which this reduction is optimal (i.e., the maximum possible) (in our experiments is about 100x100 km) and it also depends on the spatial autocorrelation structure. Therefore, the optimal distance varies with the synoptic weather conditions and topographic configuration of the validation site -among other things- which are the factors determining the spatial autocorrelation structure of solar radiation. Finally, it is worth to mention that, although the spatial averaging of solar radiation may reduce the random errors of the model estimates, it may also distort the probability distribution of solar radiation as simulated by the model. This may be a limiting factor of this post-processing approach for those applications in which a good representation of the long-term data distribution is critical. Therefore, it can be concluded that the convenience of achieving high spatial resolutions –with the huge computational effort required- and/or using a spatial average post-processing depends on the final application. For instance, based on these results and those of Chapter 4, it can be concluded that the use of the WRF model is recommended for plant operation, instead of using GCMs. For these applications, higher temporal resolutions are more suitable. In addition, data variability is important for the modeling of power production. However, for energy trading, where the time resolution is typically 1 hour, it seems to be more convenient to use a GCM, as the IFS model of the ECMWF. This is because deviations –and consequent penalties- are linked to the forecast errors, particularly the MAE. Nonetheless, the potential of WRF model is much greater than any of the GCMs in the sense of having much more flexibility. In this regard, it should be mentioned the promising results recently

obtained by some authors with the implementation of a special version of WRF devoted to solar energy applications.

In summary, the aimed objectives for this part of the thesis work were attained. It was obtained that the WRF model is able to provide fair solar radiation forecasts that can be used in the solar energy industry. Nevertheless, it was also concluded that the model has a problem of misrepresentation of cloud fraction and/or cloud amount. In practice, this can be partially emended by the application of post-processing methods to the raw model outputs, which overall reduce or eliminate the systematic error. However, from the stand point of the scientific research, the next step should be the enhancement of the capabilities of WRF to improve the representation of clouds and aerosols. Ultimately, an important effort has to be done in order to improve the immense potential of this extraordinary tool that is the WRF model.

6.4 Future research

During the realization of this thesis, several motivating questions have been opened that merits a further in-depth research. Some of them are currently in progress.

On the one hand, regarding TSYs for improving solar resource assessment, there is an open issue concerning the detailed analysis and description of the uncertainty derived from the EVA method. Additionally, it would be important to carry out a practical case study in which a solar plant yield would be analyzed regarding the expected scenarios described by the TSYs. Both issues are currently under research. Nonetheless, it would be convenient the extension of the analysis presented in this thesis to new locations and, overall, to include satellite data, as these sources are the most used in the solar energy industry. These objectives can be merged in a comprehensive work that complete the core research work developed in this thesis.

Also concerning solar resource assessment, it merits further investigation the role of the climate change in the energy yield of the

solar plants during their life-time. The aim is to include its effects in the solar resource assessment studies. There are scarce publications on this question, most of them focused on the wind energy –usually one step further than solar energy-. Climate change effects on the solar resource may not be negligible during the estimated life of solar projects. Therefore, these effects should begin to be considered in the standardization of solar resource assessment studies, or at least as part of the catalog of best practices.

On the other hand, regarding the forecasting of solar resource, a comprehensive benchmarking of the most recent versions of the NWPs models is being conducted. In a first step, only most popular and accessible GCMs will be evaluated using data from 80 radiometric stations worldwide. This analysis will follow the guidelines established in previous benchmarking works, while new elements will be introduced to examine, for instance, the performance of the spatial aggregation. This work is devised to be a reference to compare other model performances. In a second step, the last version of the WRF model will be evaluated in a selected set of locations according to the results obtained in the previous work with the GCMs. Also the role of the horizontal spatial resolution will be evaluated again, in order to know the performance of the new parameterizations implemented in the new version of WRF model. To complete this work, the study will examine the model performance as a function of the boundary and initial conditions, following the conclusions of Chapter 4.

Finally, it would be very interesting to analyze the potential of the machine learning techniques in solar radiation applications. There are several studies about this issue, particularly for enhancing the model forecasting performance in a post-processing step. In a preliminary work, the gradient boosting regressor was applied to improve satellite solar irradiance estimation. This technique allowed expanding the amount of information with long-term time series obtained from model reanalysis, such MERRA2, MERRAero and CFSR. The results were promising, showing the extraordinary potential of these black-box methods. In the case of the solar irradiance forecasting, the preliminary

results show that the improvement obtained by means of this methods is also notable.

References

Aler R., Galván I. M., Ruiz-Arias J. A., Gueymard C. A. (2017). Improving the separation of direct and diffuse solar radiation components using machine learning by gradient boosting. *Solar Energy*, 150: 558-569.

Antonanzas J., Osorio N., Escobar R., Urraca R., Martinez-de-Pison F. J., Antonanzas-Torres F. (2016). Review of photovoltaic power forecasting. *Solar Energy*, 136: 78-111.

AQFMS (2010). Air Quality Forecast Modeling System. <http://asrc.albany.edu/research/aqf/>.

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

Aryaputera A. W., Yang D., Walsh W.M. (2015). Day-Ahead Solar Irradiance Forecasting in a Tropical Environment. *Journal of Solar Energy Engineering, Transactions of the ASME*, Vol. 137(5): 051009.

Badescu V., Gueymard C. A., Cheval S., Oprea C., Baciu M., Dumitrescu A., Iacobescu F., Milos I., Rada C. (2012). Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models. *Renew Sustain Energy Rev*, 16: 1636−1656, doi:10.1016/j.r.ser.2011.12.010.

Badosa J., Wood J., Blanc P., Long C. N., Vuilleumier L., Demengel D., Haeffelin M. (2014). Solar irradiances measured using SPN1 radiometers: uncertainties and clues for development, *Atmos Meas Tech*, 7: 4267–4283.

Bellouin N., Boucher O., Haywood J., Reddy M. S. (2005). Global estimate of aerosol direct radiative forcing from satellite measurements. *Nature*, 438: 1138–1141, doi:10.1038/ nature04348.

Bender G., Davidson F., Eichelberger F., Gueymard C. A. (2011). The road to bankability: improving assessments for more accurate financial planning. In: *Proceedings of Solar 2011 Conf. American Solar Energy Soc*., Raleigh, NC.

Blanc P., Wald L. (2012). The SG2 algorithm for a fast and accurate computation of the position of the Sun for multi-decadal time period. *Solar Energy*, 86: 3072−3083, doi:10.1016/j.solener.2012.07.018

Blanc P., Espinar B., Geuder N., Gueymard C. A., Meyer R., Pitz-Paal R., Reinhardt B., Renné D., Sengupta M., Wald L., Wilbert S. (2014). Direct normal irradiance related definitions and applications: The circumsolar issue. *Solar Energy*, 110: 561-577.

Blanc P. (2015). Statistical characterization of solar resource variability. *SolarPACES 2015 conference*, Cape Town (South Africa).

Blanco-Muriel M, Alarcón-Padilla D C, López-Moratalla T, Lara-Coira M (2001) Computing the solar vector. *Sol Energy*, 70:431−441

Boilley A., Wald L. (2015). Comparison between meteorological reanalyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface. *Renewable Energy*, 75: 135–143.

Boland J., Ridley B., Brown B. (2008). Models of diffuse solar radiation. *Renew Energy* 33: 575–584, doi:10.1016/j.renene.2007.04.012.

Breitkreuz H., Schroedter-Homscheidt M., Holzer-Popp T. (2007). A case study to prepare for the utilization of aerosol forecasts in solar energy industries. *Solar Energy*, 81, Issue 11: 1377-1385.

Breitkreuz H., Schroedter-Homscheidt M., Holzer-Popp T., Dech S. (2009). Short-range direct and diffuse irradiance forecasts for solar energy applications based on aerosol chemical transport and numerical weather modeling. *Journal of Applied Meteorology and Climatology*. 48: 1766–1779.

Cebecauer T., Suri M., Gueymard C. A. (2011). Uncertainty sources in satellite-derived direct normal irradiance: How can prediction accuracy be improved globally? *Procceedings of SolarPACES Conference*, Granada, Spain.

Cebecauer T., Suri M. (2015). Typical Meteorological Year data: SolarGIS approach. *Energy Procedia*, 69: 1958-1969.

Chow C. W., Urquhart B., Kleissl J., Lave M., Dominguez A., Shields J., Washom B. (2011). Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy*, http:// dx.doi.org/10.1016/j.solener.2011.08.025.

Climate Central (CC) (2016). Earth Flirts with a 1.5-Degree Celsius Global Warming Threshold. Scientific American, by Climate Central on April 20, 2016.

Coimbra C. F. M., Pedro H. T. C. (2013) Stochastic-Learning Methods. In: J. Kleissl (Ed.), *Solar Energy Forecasting and Resource Assessment*, pp. 383-406, Academic Press, Boston, doi: 10.1016/B978-0-12-397177-7.00015-2.

Crowther T. W., Todd-Brown K. E. O., Rowe C. W., Wieder W. R., Carey J. C., Machmuller M. B., Snoek B. L., Fang S., Zhou G., Allison S. D., Blair J. M., Bridgham S. D., Burton A. J., Carrillo Y., Reich P. B., Clark J. S., Classen A. T., Dijkstra F. A., Elberling B., Emmett B. A., Estiarte M., Frey S. D., Guo J., Harte J., Jiang L., Johnson B. R., Kröel-Dulay G., Larsen K. S., Laudon H., Lavallee J. M., Luo Y., Lupascu M., Ma L. N., Marhan S., Michelsen A., Mohan J., Niu S., Pendall E., Peñuelas J., Pfeifer-Meister L., Poll C., Reinsch S., Reynolds L. L., Schmidt I. K., Sistla S., Sokol N. W., Templer P. H., Treseder K. K., Welker J. M., Bradford M. A. (2016). Quantifying global soil carbon losses in response to warming. *Nature*, 540: 104–108.

Delignette-Müller M. L., Dutang C., Pouillot R., Denis J. B. (2015). Fitdistrplus: an R package for fitting distributions. *Journal of Statistical Software*, 64(4): 1-34.

Deng, A., Gaudet B. J., Dudhia J., Alapaty K. (2014). Implementation and evaluation of a new shallow convection scheme in WRF. *26th Conference on Weather Analysis and Forecasting/22nd Conference on Numerical Weather Prediction*, Atlanta, GA, Amer. Meteor. Soc., 12.5. [Available online at https://ams.confex.com /ams/94Annual/webprogram/Paper236925.html.]

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

Dudhia J. (1989). Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci*., 46: 3077-3107.

Dumortier D. *et al*. (2009). Management and Exploitation of Solar Resource Knowledge: Contract No. 038665—D 1.2.1 & D1.2.2 Description of Solar Resource Products: Summary of Benchmarking Results and Examples of Use. http://www.mesor.org/docs/MESoR_Handbook_on_best_practices.pdf.

Elminir H. K., Azzam Y. A., Younes F. I. (2007). Prediction of hourly and daily diffuse fraction using neural network, as compared to linear regression models. *Energy*, 32(8): 1513–1523, doi:10.1016/j.energy.2006.10.010.

Engerer N. A. (2015). Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia. *Solar Energy*, 116: 215-237.

EPRI (2003). Electricity technology roadmap, 2003 summary and synthesis. Report No. 1010929. Electric Power Research Institute, Palo Alto, California, USA.

Ernst B., Oakleaf B., Ahlstrom M. L., Lange M., Moehrlen C., Lange B., Focken U., Rohrig K. (2007). Predicting the wind. *Power and Energy Magazine*, IEEE 5 (6): 78–89. http://dx.doi.org/10.1109/MPE.2007.906306.

European Directive (EU) 2009/28 / EC. Parliament and the European Council of 23 April 2009.

European Directive (EU) 2009/29 / EC. Parliament and the European Council of 23 April 2009.

European Environment Agency (2017). Climate change, impacts and vulnerability in Europe 2016. Publications Office of the European Union.

Fernández Peruchena C. M., Ramírez L., Silva-Pérez M. A., Lara-Fanego V., Bermejo D., Gastón M., Moreno-Tejera S., Pulgar J., Liria J., Macías S., González R., Bernardos A., Castillo N., Bolinaga B., Valenzuela R. X., Zarzalejo L. F. (2016). A statistical characterization of the long-term solar resource: Towards risk assessment for solar power projects. *Solar Energy*, 123: 29-39.

Fernández-García A, Zarza E., Valenzuela L., Pérez M. (2010). Parabolic-trough solar collectors and their applications. *Renewable and Sustainable Energy Reviews*, 14: 1695-1721.

Gastón M., Lorenz E., Lozano S., Heinemann D., Blanco M., Ramírez L. (2009). Comparison of Global Irradiance Forecasting Approaches. *SolarPACES Conference Proceedings*; September 15–18, 2009, Berlin, Germany.

Geuder N., Affolter R., Kraas B., Wilbert S. (2014). Long-term behaviour, accuracy and drift of LI-200 pyranometers as radiation sensors in rotating shadowband irradiometers (RSI). *Energy Procedia*, 49: 2330–2339.

GFS (2010). Global Forecasting System. <http://www.emc.ncep.noaa.gov/index.php?branch=GFS>.

Grell G., Dudhia J., Stauffer D. (1998). A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note, NCAR/TN-398+STR, USA.

Grena R. (2008) An algorithm for the computation of the solar position. *Solar Energy*, 82: 462−470, doi:10.1016/j.solener.2007.10.001.

Gueymard C. A. (1998). Turbidity determination from broadband irradiance measurements: a detailed multicoefficient approach. *J. Appl. Meteorol*, 37(4): 414-435.

Gueymard C. A. (2001). Parameterized transmittance model for direct beam and circumsolar spectral irradiance. *Solar Energy*, 71: 325−346.

Gueymard C. A. (2003). Direct solar transmittance and irradiance predictions with broadband models. Part 1: Detailed theoretical performance assessment. *Solar Energy*, 74: 355–379; Corrigendum: *Solar Energy*, 76: 513 (2004).

Gueymard C. A. (2005). Interdisciplinary applications of a versatile spectral solar irradiance model: A review. *Energy*, 30: 1551–1576, doi:10.1016/j.energy.2004.04.032.

Gueymard C. A. (2008a). REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation—Validation with a benchmark dataset. *Solar Energy*, 82: 272−285, doi:10.1016/j.solener.2007.04.008.

Gueymard C. A., Myers D. R. (2008b). Solar radiation measurement: Progress in radiometry for improved modeling. In: Badescu V. (ed). Modeling solar radiation at the earth's surface: Recent advances. Springer, Berlin Heidelberg, pp. 1–27.

Gueymard C. A. (2009a). Direct and indirect uncertainties in the prediction of tilted irradiance for solar engineering applications. *Solar Energy*, 83(3): 432-444.

Gueymard C. A., Myers D. R. (2009b). Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling. *Solar Energy*, 83: 171–185.

Gueymard C. A. (2010). Progress in direct irradiance modeling and validation. American Solar Energy Soc., *Solar 2010 Conf*., Phoenix, AZ.

Gueymard C. A. (2012a). Clear-sky irradiance predictions for solar resource mapping and largescale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. *Solar Energy*, 86(6): 2145−2169.

Gueymard C. A. (2012b). Solar Radiation, Introduction. Encyclopedia of Sustainability Science and Technology, pp. 9740-9744. Editors: Robert A. Meyers. ISBN: 978-0-387-89469-0.

Gueymard C. A. (2012c). Temporal variability in direct and global irradiance at various time scales as affected by aerosols. *Solar Energy*, 86: 3544-3553.

Gueymard C. A. (2014). A review of validation methodologies and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects. *Renewable and Sustainable Energy Reviews*, 39: 1024-1034.

Gueymard C. A., Ruiz-Arias J. A. (2016). Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol. Energy, 128: 1-30.

Habte A., Sengupta M., Wilcox S. (2013). Validation of GOES-derived surface radiation using NOAA's physical retrieval method. National Renewable Energy Laboratory, Tech Report NREL/TP-5500-57442.

Habte A., Lopez A., Sengupta M., Wilcox S. (2014). Temporal and spatial comparison of gridded TMY, TDY, and TGY data sets.

Technical Report NREL/TP-5D00-60886. National Renewable Energy Laboratory (NREL), Golden, CO.

Hammer A., Heinemann D., Hoyer C., Kuhlemann R., Lorenz E., Müller R., Beyer H. G. (2003). Solar energy assessment using remote sensing technologies. *Remote Sensing of Environment*, 86(3): 423-432.

He X., Yuan C., Yang Z. (2016). Performance evaluation of Chinese solar radiation forecast based on three global forecast back ground fields, Taiyangneng Xuebao/Acta Energiae Solaris Sinica, 37(4), pp. 897-904.

Heinemann D., Lorenz E., Girodo M. (2006a). Forecasting of Solar Radiation. In: E. D. Dunlop, L. Wald, M. Suri (Eds.): Solar Energy Resource Management for Electricity Generation from Local Level to Global Scale. Nova Science Publishers, Hauppauge.

Heinemann D., Lorenz E., Girodo M. (2006b). Solar Irradiance Forecasting for the Management of Solar Energy Systems, Solar 2006, Denver, CO, USA (07.07.06).

HIRLAM (2010). High Resolution Limited Area Model. <http://hir-lam.org/>.

Hong S. Y., Noh Y., Dudhia J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Weather Rev*. 134: 2318–2341.

Hu Y. X., Stamnes K. (1993). An accurate parameterization of the radiative properties of water clouds suitable for use in climate models. *J. Climate*, 6(4): 728-742.

IEA (2006a). Global energy technology perspectives. International Energy Agency, OECD Publication Service, OECD, Paris.

IEA (2006b). World energy outlook 2006. International Energy Agency, OECD Publication Service, OECD, Paris.

IEA (2007). Energy technologies at the cutting edge. International Energy Agency, OECD Publication Service, OECD, Paris.

IEA (2008). Empowering Variable Renewables - Options for Flexible Electricity Systems. OECD/IEA, Paris.

IEA (2010). Electricity Information 2010. Published by the International Energy Agency. 762 pages. OECD/IEA, Paris.

IEA (2011). World Energy Outlook 2011. International Energy Agency, OECD Publication Service, OECD, Paris.

IEA (2014a). Technology Roadmap: Solar Photovoltaic Energy – 2014 edition. OECD/IEA, Paris.

IEA (2014b). Technology Roadmap: Solar Thermal Electricity - 2014 edition. OECD/IEA, Paris.

IEA (2015). Medium-Term Renewable Energy Market Report – 2015 edition. OECD/IEA, Paris.

IEA (2016a). World energy outlook 2016. OECD/IEA, Paris.

IEA (2016b). Medium-term renewable energy market report 2016. ISBN PRINT 978-92-64-26496-0 / PDF 978-92-64-26497-7. OECD/IEA, Paris.

IEA (2016c). Next Generation Wind and Solar Power. OECD/IEA, Paris.

IEA SHC Task 36 (2011). International Energy Agency, Solar Heating and Cooling Programme, Task 36 Solar Resource Knowledge Management, Subtask A, Model Benchmarking. <http://www.iea-shc.org/ task36/>.

Ineichen P. (2006). Comparison of eight clear sky broadband models against 16 independent data banks. *Solar Energy*, 80: 468−478, doi:10.1016/j.solener.2005.04.018.

Ineichen P. (2008a). A broadband simplified version of the Solis clear sky model. *Solar Energy*, 82: 758−762, doi:10.1016/j.solener.2008.02.009.

Ineichen P. (2008b). Comparison and validation of three global-to-beam irradiance models against ground measurements. *Solar Energy*, 82: 501-512.

Ineichen, P. (2011a). Global irradiance on tilted and oriented planes: Model validations. (333.7-333.9; 550). Genève: Retrieved from https://archive-ouverte.unige.ch/unige:23519.

Ineichen P. (2011b). Global irradiation: average and typical year, and year to year annual variability. (333.7-333.9; 550). Geneva: Retrieved from https://archive-ouverte.unige.ch/unige:23518.

Ineichen P. (2014). Long term satellite global, beam and diffuse irradiance validation. *Energy Procedia*, 48: 1586–1596.

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

IPCC (2014a). IPCC. Summary for Policymakers. In: Edenhofer O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B., Kriemann J. S., S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx editors. Climate Change 2014, Mitigation of Climate Change Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press, 2014.

IPCC (2014b). Climate change 2014: Impacts, adaptation and vulnerability. Part B: Regional aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge; New York.

Iqbal M. (1983). An Introduction to Solar Radiation. In: Academic Press, New York, 1983, pp. 188-195.

Isvoranu D., Badescu V. (2013). Comparison between measurements and WRF numerical simulation of global solar irradiation in Romania, Analele Universitatii de Vest din Timisoara, Seria Fizica, Vol. LVII, pp. 24-33.

Jensen T. L., Fowler T. L., Brown B. G., Lazo J. K., Haupt S. E. (2016). Metrics for Evaluation of Solar Energy Forecasts. NCAR Technical Note NCAR/TN-527+STR, 67 pp.

Jimenez P., Hacker J. Dudhia J., Haupt S. E., Ruiz-Arias J. A., Gueymard C. A., Thompson G., Eidhammer T., Deng A. (2015). 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.

Jones A. S., Fletcher S. J. (2013). Data assimilation in numerical weather prediction and sample applications, J. Kleissl (Ed.), Solar Energy Forecasting and Resource Assessment, pp. 319-355, Academic Press, Boston.

Kain J. S., Fritsch J. M. (1993). Convective parameterization for mesoscale models: The Kain-Fritcsh scheme. In: Emanuel K. A., Raymond D. J. (Eds). The representation of cumulus convection in numerical models. Amer. Meteor. Soc., pp. 165-170.

Kallos G. (1997). The regional weather forecasting system SKIRON. In: Proceedings, Symposium on Regional Weather Prediction on Parallel Computer Environments, pp. 15–17, October 1997, Athens, Greece.

Kasten F.; Young A. T. (1989). Revised optical air mass tables and approximation formula. *Applied Optics*, 28: 4735–4738.

Kim D., Ramanathan V. (2008). Solar radiation budget and radiative forcing due to aerosols and clouds. *J. Geophys. Res*., 113: D02203, doi:10.1029/2007JD008434.

Kleissl J. (2013). Solar Energy Forecasting and Resource Assessment, Academic Press, Boston, doi:10.1016/B978-0-12-397177-7.01001-9.

Kraas B., Schroedter-Homscheidt M., Madlener R. (2013) Economic merits of a stateof-the-art concentrating solar power forecasting system for participation in the Spanish electricity market. *Sol. Energy*, 93, pp. 244-255.

Kühnert J., Lorenz E., Heineman D. (2013). Satellite-based irradiance and power forecasting for the German Energy Market, In J. Kleissl (Ed.), Solar Energy Forecasting and Resource Assessment, pp. 267-297, Academic Press, Boston.

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

Larson V. E. (2013). Forecasting Solar Irradiance with Numerical Weather Prediction Models. In: J. Kleissl (Ed.), Solar Energy Forecasting and Resource Assessment, pp. 299-318, Academic Press, Boston.

Law E. W., Kay M., Taylor R. A. (2016) Evaluating the benefits of using short-term direct normal irradiance forecasts to operate a concentrated solar thermal plant, *Sol. Energy*, 140, 93-108.

Lefèvre M., Wald L., Diabaté L. (2007). Using reduced data sets ISCCP-B2 from the Meteosat satellites to assess surface solar irradiance. *Solar Energy*, 81: 240–253.

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., Morcrette J. J. (2013). McClear: A new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmos Meas Tech*, 6: 2403−2418, doi: 10.5194/amt−6−2403−2013.

Lima F. J. L., Martins F. R., Pereira E. B., Lorenz E., Heinemann D. (2016). Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. *Renewable Energy*, 87: 807-818.

Linares-Rodriguez A., Quesada-Ruiz S., Pozo-Vazquez D., Tovar-Pescador J. (2015). An evolutionary artificial neural network ensemble model for estimating hourly direct normal irradiances from meteosat imagery. *Energy*, 91: 264-273.

Long C. N., Shi Y. (2008). An automated quality assessment and control algorithm for surface radiation measurements. *Open Atmos. Sci. J.*, 2: 23–37.

Lorenz E., Hurka J., Heinemann D., Beyer H. G., (2009a). Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2(1).

Lorenz E., Remund J., Müller S. C., Traunmüller W., Steinmaurer G., Pozo D., Ruiz-Arias J. A., Lara-Fanego V., Ramirez L., Romeo M. G., Kurz C., Pomares L. M., Guerrero C. G. (2009b). Benchmarking of different approaches to forecast solar irradiance. In: Proceedings of the 24th European Photovoltaic Solar Energy Conference, pp. 21–25.

Lorenz E., Scheidsteger T., Hurka J., Heinemann D., Kurz, C. (2011) Regional PV power prediction for improved grid integration. *Prog. Photovolt: Res. Appl*., 19: 757–771.

Lorenz E., Heinemann, D. (2012) Prediction of Solar Irradiance and Photovoltaic Power. *Comprehensive Renewable Energy*, 1, 239 - 292.

Lorenz E., Kühnert J., Heinemann D., Nielsen K. P., Remund J., Müller S. C. (2015). Comparison of Irradiance Forecasts Based on Numerical Weather Prediction Models with different Spatio-Temporal Resolutions. 31st European Photovoltaic Solar Energy Conference and Exhibition. 5DP.1.3, pp. 1524 – 1537. ISBN: 3-936338-39-6, DOI: 10.4229/EUPVSEC20152015-5DP.1.3.

Lyamani H., Olmo F.J., Alados-Arboledas L. (2005). Saharan dust outbreak over southeastern Spain as detected by sun photometer. *Atmos. Environ*, 39(38): 7276-7284.

Mailhot J., Bélair S., Lefaivre L., Bilodeau B., Desgagné M., Girard C., Glazer A., Leduc A. M., Méthot A., Patoine A., Plante A., Rahill A., Robinson T., Talbot D., Tremblay A., Vaillancourt P. A., Zadra A. (2006). The 15-km version of the Canadian regional forecast system. *Atmosphere – Ocean*, 44: 133–149.

Manobianco J., Zack J. W., Taylor G. E. (1996). Workstation based real-time mesoscale modeling designed for weather support to operations at the Kennedy Space Center and Cape Canaveral Air Station. *Bulletin of the American Meteorological Society*, 77: 653–672.

Marcos et al. (2013) Attenuation of power fluctuations in large PV power plants: The use of prediction to optimize the storage system. In Proceedings of the 28th European Photovoltaic Solar Energy Conference (PVSEC), Paris, 2013.

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

Mayer (2009). Radiative transfer in the cloudy atmosphere. *Eur. Phys. J. Conferences*, 1: 75–99.

Mellit A., Pavan A. M. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar Energy*, 84(5): 807-821.

Meyer R., Beyer H. G., Fanslau J., Geuder N., Hammer A., Hirsch T., Hoyer-Klick C., Schmidt N., Schwandt M. (2009). Toward standardization of CSP yield assessments. In: Paper read at 15th SolarPACES Conference, at Berlin, Germany.

Michalsky J. J. (1988). The Astronomical Almanac's algorithm for approximate solar position (1950-2050). *Solar Energy*, 40: 227−235.

Michalsky J. J., Dolce R., Dutton E. G., Haeffelin M., Jeffries W., Stoffel T., Hickey J., Los A., Mathias D., McArthur B., Nelson D., Philipona R., Reda I., Rutledge K., Zerlaut G., Forgan B., Kiedron P., Long C., Gueymard C. (2005). Toward the development of a diffuse horizontal shortwave irradiance working standard. *J Geophys Res*, 110: D06107.

Michalsky J., Dutton E., Nelson D., Wendell J., Wilcox S., Andreas A., Gotseff P., Myers D., Reda I., Stoffel T., Behrens K., Carlund T., Finsterle W., Halliwell D. (2011). An extensive comparison of commercial pyrheliometers under a wide range of routine observing conditions. *J Atmos Oceanic Technol*, 28: 752–766.

Miller S. D., Heidinger A. K., Sengupta M. (2013). Physically based satellite methods. In: Kleissl J. (ed.) Solar energy forecasting and resource assessment. Academic Press, Oxford, pp. 49–80.

Mlawer E. J., Taubman S. J., Brown P. D., Iacono M. J., Clough S. A. (1997). Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the long-wave. *J. Geophys. Res*, 102(D14): 16663–16682.

Müller S. C., Remund J. (2013). Satellite based shortest term solar energy forecast system for entire Europe for the next hours, 29th European Photovoltaic Solar Energy Conference and Exhibition, doi: 10.4229/EUPVSEC20142014-5BV.1.6.

Natschläger T., Traunmüller W., Reingruber K., Exner H. (2008). Lokal optimierte Wetterprognosen zur Regelung stark umweltbeeinflusster Systeme; SCCH, Blue Sky. In: Tagungsband Industrielles Symposium Mechatronik Automatisierung. Clusterland Oberösterreich GmbH/ Mechatronik-Cluster, pp. 281–284.

NSRDB (2005). <http://rredc.nrel.gov/solar/old_data/nsrdb/1991-2005/>.

Pedro H., Coimbra C. (2012). Assessment of forecasting techniques for solar power production with no exogenous inputs. *Solar Energy*, 86 (2017): 2028.

Olmo F. J., Quirantes A.., Lara V., Lyamania H., Alados-Arboledas L. (2008). Aerosol optical properties assessed by an inversion method using the solar principal plane for non-spherical particles. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 109(8): 1504-1516.

Oreskes N. (1998). Evaluation (not validation) of quantitative models. *Environ Health Perspect.*,106(6): 1453–1460.

Oreskes N., Kristin Shrader-Frechette K., Kenneth Belitz K. (1994). Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science*, 263, 5147: 641-646.

Pachauri R. K., Allen M., Barros V., Broome J., Cramer W., Christ R., Church J., Clarke L., Dahe Q., Dasgupta P., *et al*. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Pelland S., Gallanis G., Kallos G. (2011). Solar and photovoltaic forecasting through postprocessing of the global environmental multiscale numerical weather prediction model. *Progress in Photovoltaics: Research and Applications* (November 22, 2011).

Pelland S., Galanis G., Kallos G. (2013). Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model. *Prog. Photovolt: Res. Appl*., 21: 284–296.

Perez R., Ineichen P., Seals R., Michalsky J., Stewart R. (1990a). Modelling daylight availability and irradiance components from direct and global irradiance. *Solar Energy*, 44(5): 271-289.

Perez R., Ineichen P., Seals R., Zelenka A. (1990b). Making full use of the clearness index for parameterizing hourly insolation conditions. *Solar Energy*, 45(2): 111-124.

Perez R., Ineichen P., Maxwell E. L., Seals R., Zelenka A. (1992). Dynamic global-to-direct irradiance conversion models. *ASHRAE Trans*., 98(1): 354–369.

Perez R., Seals R., Zelenka A. (1997). Comparing satellite remote sensing and ground network measurements for the production for site/time specific irradiance data. *Solar Energy*, 60: 89-96.

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

Perez R., Moore K., Wilcox S., Renné D., Zelenka A. (2007) Forecasting solar radiation – Preliminary evaluation of an approach based upon the national forecast database *Sol. Energy*, 81(6), 809-812.

Perez R. (2008). Making the case for solar energy. *D&A*. Issue 09.

Perez Y., Ramos-Real F. J. (2009a). The public promotion of wind energy in Spain from the transaction costs perspective 1986–2007. *Renew. Sust. Energ. Rev*., 13: 1058–1066.

Perez R., Kivalov S., Schlemmer J., Hemker K. Jr., Renné D., Hoff T. (2009b). Validation of short and medium term operational solar radiation forecasts in the US. Proceedings of the ASES Annual Conference, Buffalo, New York.

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

Perez R., Cebecauer T., Suri M. (2013a). Semi-empirical satellite models. In: Kleiss, Jan (Ed.), Solar Energy Forecasting and Resource Assessment. Elsevier (chapter 2).

Perez R., Hoff T. (2013b). SolarAnywhere forecasting, In: J. Kleissl (Ed.), Solar Energy Forecasting and Resource Assessment, pp. 233-

265, Academic Press, Boston, doi:10.1016/B978-0-12-397177-7.00010-3.

Polo J., Zarzalejo L. F., Ramirez L. (2008). Solar radiation derived from satellite images. In: Badescu, Viorel (Ed.), Modeling Solar Radiation at the Earth Surface. Springer-Verlag (chapter 18).

Polo J., Wilbert S., Ruiz-Arias J. A., Meyer R., Gueymard C. A., Suri M., Martín L., Mieslinger T., Blanc P., Grant I., Boland J., Ineichen P., Remund J., Escobar R., Troccoli A., Sengupta M., Nielsen K. P., Renne D., Geuder N., Cebecauer T. (2016). Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. *Solar Energy*, 132: 25-37.

Pozo-Vázquez D., Tovar-Pescador J., Gámiz-Fortis S. R., Esteban-Parra M. J., Castro-Díez Y. (2004). NAO and solar radiation variability in the European North Atlantic Region. *Geophysical Research Letters*, 31(5): 1-4.

Quaschning V. (2004). Technical and economical system comparison of photovoltaic and concentrating solar thermal power systems depending on annual global irradiation. *Solar Energy*, 77: 171-178.

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

Quesada-Ruiz S., Linares-Rodriguez A., Ruiz-Arias J. A., Pozo-Vazquez D., Tovar-Pescador J. (2015). An advanced ANN-based method to estimate hourly solar radiation from multi-spectral MSG imagery. *Solar Energy*, 115: 494-504.

Realpe A., Vernay C., Pitaval S., Lenoir C., Blanc P. (2016). Benchmarking of five Typical Meteorological Year datasets dedicated to Concentrated-PV systems. *Energy Procedia*, Elsevier, 97: 108 - 115.

Reda I., Andreas A. (2004). Solar position algorithm for solar radiation applications. *Solar Energy*, 76: 577–589, doi:10.1016/j.solener.2003.12.003.

Reda I., Andreas A. (2008). Solar position algorithm for solar radiation applications. NREL Technical Report NREL/TP-560–34302. National Renewable Energy Laboratory, Golden, CO. http://www.nrel.gov/docs/fy08osti/34302.pdf.

REE (2015). El sistema eléctrico español.

REE (2016). El sistema eléctrico español. Avance 2016.

Reindl D. T., Beckman W. A., Duffie J. A. (1990). Diffuse fraction correlations. *Solar Energy*, 45: 1–7.

Remund, Perez R., Lorenz E. (2008). Comparison of Solar Radiation Forecasts for the USA. Proceedings of the 23rd European Photovoltaic Solar Energy Conference, 1.9-4.9 2008, Valencia, Spain.

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

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

Roesch A., Wild M., Ohmura A., Dutton E. G., Long C. N., Zhang T. (2011). Assessment of BSRN radiation records for the computation of monthly means. *Atmos. Meas. Tech*., 4: 339–354.

Royal Decree, 2007. Royal Decree 661/2007, of 12 of March, regulating the production of electricity in the special regime in Spain.

Ruiz-Arias J. A., Pozo-Vázquez D., Sánchez-Sánchez N., Montávez J. P., Hayas-Barrú A., Tovar-Pescador J. (2008). Evaluation of two MM5-PBL parameterizations for solar radiation and temperature estimation in

the South-Eastern area of the Iberian Peninsula. *Il Nuovo Cimento C*, 31: 825–842.

Ruiz-Arias J. A. *et al*. (2009). Evaluation of two MM5-PBL parameterizations for solar radiation and temperature estimation in the South-Eastern area of the Iberian Peninsula. *Il Nuovo Cimento C*, 31: 825-842.

Ruiz-Arias J. A., Cebecauer T., Tovar-Pescador J., Suri M. (2010a). Spatial disaggregation of satellite-derived irradiance using a high-resolution digital elevation model. *Solar Energy*, 84(9): 1644-1657.

Ruiz-Arias J. A., Alsamamra H., Tovar-Pescador J., Pozo-Vázquez D. (2010b). Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. *Energy Convers Manage*, 51: 881–893.

Ruiz-Arias J. A., Pozo-Vázquez D., Lara-Fanego V., Tovar-Pescador J., (2011). A high-resolution topographic correction method for clear-sky solar irradiance derived with a numerical weather prediction model. *J.Appl. Meteorol. Climatol*., 50: 2460-2472.

Ruiz-Arias J. A., Gueymard C. A., Dudhia J., Pozo-Vazquez, D. (2012). Improvement of the Weather Research and Forecasting (WRF) model for solar re-source assessments and forecasts under clear skies. World Renewable Energy Forum, Denver, CO, 2012.

Ruiz-Arias J. A., Dudhia J., Santos-Alamillos F. J., Pozo-Vázquez D. (2013). Surface clear-sky shortwave radiative closure intercomparisons in the Weather Research and Forecasting model. *J Geophys Res-Atmos*, 118(17): 9901–9913.

Ruiz-Arias J. A., Dudhia J., Gueymard C. A. (2014). A simple parameterization of the short-wave aerosol optical properties for surface direct and diffuse irradiances as- sessment in a numerical weather model. *Geosci. Model Dev*., 7: 1159–1174, doi:10.5194/gmd-7-1159-2014.

Ruiz-Arias J. A., Gueymard C. A., Santos-Alamillos F. J., D. Pozo-Vázquez D. (2015a). Do spaceborne aerosol observations limit the accuracy of modeled surface solar irradiance? *Geophys. Res. Lett*., 42: 605–612, doi: 10.1002/2014GL062309.

Ruiz-Arias J. A., Gueymard C. A. (2015b). Solar Resource for High Concentrator Photovoltaic Applications. In: High Concentrator Photovoltaics: Fundamentals, Engineering and Power Plants, Pérez-Higueras P., Fernández E. F. (Eds) Green Energy and Technology Series, Springer. ISBN 978-3-319-15038-3.

Ruiz-Arias J. A., Arbizu-Barrena C., Santos-Alamillos F. J., Tovar-Pescador J., Pozo-Vázquez D. (2016). Assessing the surface solar radiation budget in the WRF model: A spatiotemporal analysis of the bias and its causes. *Monthly Weather Review*, 144: 703-711, doi:10.1175/MWR-D-15-0262.1.

Saintcross J., Piwko R., Bai X., Clara K., Jordan G., Miller N., Zimberlin J. (2005). The Effects of Integrating Wind Power on Transmission System Planning, Reliability, and Operations. The New York State Energy Research and Development Authority, Albany, New York, U.S.A.
http://www.nyserda.org/publications/wind_integration_report.pdf.

Salas V., Olias E. (2009). Overview of the photovoltaic technology status and perspective in Spain. *Renewable & Sustainable Energy Reviews*, 13(5): 1049-1057.

Sengupta M., Habte A., Kurtz S., Dobos A. P., Wilbert S., Lorenz E., Stoffel T., Renné D., Gueymard C. A., Myers D., Wilcox S., Blanc P., Perez R. (2015). Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications. National Renewable Energy Laboratory, Golden, CO. (No. NREL/TP-5D00-63112) (2015).

Skamarock W. C., Klemp J. B., Dudhia J., Gill D. O., Barker D. M., Duda M. G., Huang X. Y., Wang W., Powers J. G. (2008). A Description of the Advanced Research WRF Version 3, NCAR/TN-

475+STR, Mesoscale and Microscale Meteorology Division, National Centre for Atmospheric Research, Boulder, USA.

SolarAnywhere (2010). <www.solaranywhere.com>.

Sosa-Tinoco I., Peralta-Jaramillo J., Otero-Casal C., López- Agüera A., Miguez-Macho G., Rodríguez-Cabo I. (2016). Validation of a global horizontal irradiation assessment from a numerical weather prediction model in the south of Sonora-Mexico. *Renewable Energy*, 90: 105-113.

Stensrud D. J. (2007). Parameterization schemes: keys to understanding numerical weather prediction models. Cambridge University Press.

Stoffel T., Renné D., Daryl M., Wilcox S., Sengupta M., George R., Turchi C. (2010). Concentrating solar power: Best practices handbook for the collection and use of solar resource data (CSP). Report No. NREL/TP-550-47465. National Renewable Energy Laboratory (NREL), Golden, CO.

SURFRAD (2010). NOAA's Surface Radiation Network: https://www.esrl.noaa.gov/gmd/grad/surfrad/.

Suri M., Cebecauer T. (2011). Requirements and standards for bankable DNI data products in CSP projects. In: Proceedings of: SolarPACES 2011 Conf., Granada, Spain.

Szuromi P., Jasny B., Clery D., Austin J., Hanson B. (2007). Energy for the long haul. *Science*, 315(5813): 781.

Thompson G., Rasmussen R. M., Manning K. (2004). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis, *Mon. Weather Rev*., 132: 519–542.

Thompson G., Eidhammer T. (2014). A study of aerosol impacts on clouds and precipitation devel- opment in a large winter cyclone. *J. Atmos. Sci*., 71: 3636–3658, doi:10.1175/JAS-D-13-0305.1.

Thuman C., Schitzer M., Johnson P. (2012). Quantifying the accuracy of the use of Measure–Correlate–Predict methodology for long-term solar resource estimates. In: Proceedings of: American Solar Energy Society, Denver CO.

Tindal A., Parkes J., Munoz L., Wasey V. (2006). Wind energy trading benefits through short term forecasting. Presented at the European Wind Energy Conf., Athens, Greece. Available in: http://www.ewec2006proceedings.info/, 27.2-2.3.

Traunmüller W., Steinmaurer G. (2010). Solar Irradiance Forecasting, Benchmarking of Different Techniques and Applications of Energy Meteorology. Eurosun, Graz, Austria.

Urquhart B., Ghonima M., Nguyen D., Kurtz B., Chow C. W., Kleissl J. (2013). Sky-Imaging Systems for Short-Term Forecasting. In: J. Kleissl (Ed.), Solar Energy Forecasting and Resource Assessment, pp. 195-232, Academic Press, Boston.

Urquhart B., Kurtz B., Dahlin E., Ghonima M., Shields J. E., Kleissl J. (2015). Development of a sky imaging system for short-term solar power forecasting. *Atmospheric Measurement Techniques*, 8(2): 875-890.

van Zalinge B. C., Feng Q. Y., Dijkstra H. A. (2016). On determining the Point of no Return in Climate Change. *Earth Syst. Dynam*, in review.

Vignola F., Michalsky J., Stoffel T. (2012). Solar and infrared radiation measurements. CRC Press, Taylor & Francis Group, Boca Raton, FL.

Vuilleumier L., Hauser M., Félix C., Vignola F., Blanc P., Kazantzidis A., Calpini B. (2014). Accuracy of ground surface broadband shortwave radiation monitoring. *J Geophys Res Atmos.*, 119.

Wilcox S. (2012). National solar radiation database 1991–2010 update: User's manual. National Renewable Energy Laboratory Tech Report NREL/TP-5500-54824, Golden, CO.

Wilks, D. S. (1995). Statistical methods in the atmospheric sciences. Academic Press, New York.

Wittmann M., Breitkreuz H., Schroedter-Homscheidt M., Eck M. (2008). Case-Studies on the Use of Solar Irradiance Forecast for Optimized Operation Strategies of Solar Thermal Power Plants. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(1): 18-27.

Xue M., Droegemeier K. K., Wong V., Shapiro A., Brewster K., Carr F., Weber D., Liu Y., Wang D. H. (2001). The Advanced Regional Prediction System (ARPS) – a multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: model physics and applications. *Meteorology and Atmospheric Physics*, 76: 134–165.

Zamora R. J., Solomon S., Dutton E. G., Bao J. W., Trainer M., Portmann R. W., White A. B., Nelson D. W., McNider R. T. (2003). Comparing MM5 radiative fluxes with observations gathered during the 1995 and 1999 Nashville southern oxidant studies. *J. Geophys. Res*., 108(D2): 4050.

Zamora R. J., Dutton E. G., Trainer M., McKeen S. A, Wilczak J. M., Hou Y. T. (2005). The accuracy of solar irradiance calculations used in mesoscale numerical weather prediction. *Mon. Weather Rev*., 133: 783-792.

Zempila M. M., Giannaros T. M., Bais A., Melas D., Kazantzidis A. (2016). Evaluation of WRF shortwave radiation parameterizations in predicting Global Horizontal Irradiance in Greece. *Renewable Energy*, 86: 831-840.

Annex A

Analytical formulation of EVA method

Nomenclature

$i = 1, \ldots, 12$; month of the year.
$j = 1, \ldots, n$; year of the long-term time series.
x_i: monthly expected value (MEV) for month i; the unkonwns.
Pe_i^{50}: median of the available values of month i of the long-term time series.
Pe_y: annual probability of exceedance at the y level (Pexx).
w_i: weight.
MAD: median absolute deviation

$rad_{i,j}$: monthly irradiation of month i and year j of the long-term data.

$rad_{a,j}$: annual irradiation of year j of the long-term data.

$csrad_{i,j}$: clear-sky irradiation for month i and year j.

Determination of MEVs

The start point is the function defined by (equation 2.1 in Chapter 2):

$$f(x_1, \ldots x_{12}) = \sum_{i=1}^{12} \left(\frac{\sum_{i=1}^{12} w_i}{w_i} \right) \left(x_i - Pe_i^{50} \right)^2 \qquad \text{(A.1)}$$

The imposed constraint is that the sum of the monthly values (MEVs) should be equal to the annual probability of exceedance:

$$\sum_{i=1}^{12} x_i = Pe_y \tag{A.2}$$

To solve this minimization problem with constraint the method of the Lagrange multipliers can be applied. Thus, the following function is defined:

$$f(x_1, \dots x_{12}, \lambda) = \sum_{i=1}^{12} \left(\frac{\sum_{i=1}^{12} w_i}{w_i} \right) \left(x_i - Pe_i^{50} \right)^2 - \lambda \left(\sum_{i=1}^{12} x_i - Pe_y \right) \tag{A.3}$$

Partial differentials are applied in order to find the values that minimize the fuction A.3:

$$\frac{\partial f}{\partial x_i} = 0\,, \qquad \frac{\partial f}{\partial \lambda} = 0 \tag{A.4}$$

Thus, a system of 12+1 equations is obtained (A.5 + A.6):

$$2 \left(\frac{\sum_{i=1}^{12} w_i}{w_i} \right) (x_i - Pe_i^{50}) - \lambda = 0 \tag{A.5}$$

$$\sum_{i=1}^{12} x_i - Pe_y = 0 \tag{A.6}$$

Solvin for x_i in equations A.5:

$$x_i = \frac{\lambda}{2} \frac{w_i}{\sum_{i=1}^{12} w_i} + Pe_i^{50} \tag{A.7}$$

Replacing A.7 in A.6:

$$\sum_{i=1}^{12}\left(\frac{\lambda}{2}\frac{w_i}{\sum_{i=1}^{12}w_i}+Pe_i^{50}\right)=Pe_y \tag{A.8}$$

Operating:

$$\frac{\lambda}{2}\sum_{i=1}^{12}\frac{w_i}{\sum_{i=1}^{12}w_i}+\sum_{i=1}^{12}Pe_i^{50}=Pe_y \tag{A.9}$$

In equation A.9, the term of weights is equal to 1:

$$\sum_{i=1}^{12}\frac{w_i}{\sum_{i=1}^{12}w_i}=\frac{\sum_{i=1}^{12}w_i}{\sum_{i=1}^{12}w_i}=1 \tag{A.10}$$

And therefore:

$$\frac{\lambda}{2}=Pe_y-\sum_{i=1}^{12}Pe_i^{50} \tag{A.11}$$

Finally, replacing A.11 in A.7 the equation for the determination of the MEVs is obtained:

$$x_i=\frac{w_i}{\sum_{i=1}^{12}w_i}\left(Pe_y-\sum_{i=1}^{12}Pe_i^{50}\right)+Pe_i^{50} \tag{A.12}$$

Determination of weight factors

Variability factor: by definition it is the variability (measured with the MAD) of the normalized (de-seasonalised) monthly values.

$$f^1{}_i = MAD_i \left(\frac{rad_{i,j}}{csrad_{i,j}} \right) \tag{A.12}$$

Seasonality factor: by definition it is the contribution of each month to the total annual energy.

$$f^2{}_i = \frac{\sum_{j=1}^{n} rad_{i,j}}{\sum_{j=1}^{n} rad_{a,j}} \tag{A.13}$$

Weights are determined by the product:

$$w_i = f^1{}_i \cdot f^2{}_i \tag{A.14}$$

Equation A.13 can be simplified. Firstly, dividing by number of years (n):

$$f^2{}_i = \frac{mean_j(rad_{i,j})}{mean_j(rad_{a,j})} \tag{A.14}$$

Now considering the process of normalization of the weights, the mean of the annual irradiation becomes a common factor and therefore it can be canceled. In equation A.12, the normalized factor of weights is:

$$\frac{w_i}{\sum_{i=1}^{12} w_i} \tag{A.15}$$

Therfore:

$$\frac{f_i^1 \cdot f_i^2}{\sum_{i=1}^{12} f_i^1 \cdot f_i^2} = \frac{f_i^1 \cdot \frac{mean_j(rad_{i,j})}{mean_j(rad_{a,j})}}{\sum_{i=1}^{12} f_i^1 \cdot \frac{mean_j(rad_{i,j})}{mean_j(rad_{a,j})}} \tag{A.16}$$

And hence:

$$\frac{w_i}{\sum_{i=1}^{12} w_i} = \frac{f_i^1 \cdot mean_j(rad_{i,j})}{\sum_{i=1}^{12} f_i^1 \cdot mean_j(rad_{i,j})} \quad \text{(A.17)}$$

Therefore, when computing equation A.12, the factors and weights can be calculated by means of equations A.13, A.14 and:

$$f^2{}_i = mean_i\left(rad_{i,j}\right) \quad \text{(A.18)}$$

Annex B

Resumen

En la presente tesis doctoral se lleva a cabo un estudio de la evaluación y de la predicción del recurso solar para su aplicación en el campo de la industria de la energía solar. El objetivo principal es mejorar el conocimiento sobre varios aspectos de la radiación solar como fuente primaria de energía. El propósito es contribuir al desarrollo de esta energía renovable dentro del marco actual de transformación del sector energético a nivel mundial, fuertemente condicionado por el cambio climático y la necesidad de un desarrollo sostenible. Sin embargo, a pesar del incesante desarrollo tecnológico y el considerable abaratamiento de costes, su grado de introducción dentro de los sistemas eléctricos a gran escala está todavía lejos de su potencial real. Esto es debido en gran parte a que, a pesar de que la radiación solar es la fuente primaria de energía más abundante del planeta, presenta de forma natural una gran variabilidad espacio-temporal. Esta característica constituye la mayor fuente de incertidumbre en el desarrollo de los proyectos solares, tanto en la fase inicial de estudio de viabilidad como durante la fase de operación. Con el fin de contribuir a la reducción de dicha incertidumbre, en el trabajo de investigación llevado a cabo en esta tesis doctoral se han desarrollado y evaluado métodos para la caracterización y la estimación de la irradiancia solar en superficie, tanto para la componente global (GHI) como para la directa (DNI).

En primer lugar, dentro del ámbito de la evaluación del recurso se ha desarrollado un método novedoso para la obtención de un año representativo para la caracterización de la irradiancia solar a escalas multianuales en una determinada localización de interés. Este método,

denominado EVA, permite la generación de los llamados años solares representativos (TSY). Estos años artificiales son ampliamente usados dentro de la industria solar, principalmente en los análisis de bancabilidad de los proyectos solares. En particular, el uso de los TSY se ha convertido en un estándar para la estimación de la producción de las plantas solares bajo diferentes condiciones de disponibilidad del recurso. Sin embargo, hoy en día no existe un consenso científico respecto a la obtención de los TSY. Esto provoca que haya una mayor incertidumbre, dado que la aplicación de las distintas metodologías sobre un mismo conjunto de datos puede dar como resultado diferentes TSY. Por tanto la estandarización del método para la generación de los TSY es una demanda de la industria. En este sentido, el método desarrollado en esta tesis pretende contribuir a dicha estandarización. El método EVA tiene una fundamentación estadística y está completamente formulado de forma analítica. La característica más importante del método es que permite generar los TSY correspondientes a cualquier escenario anual de disponibilidad del recurso. Esto incluye las situaciones de baja cantidad de recurso, las cuales son fundamentales en los estudios de bancabilidad de los proyectos solares. Los resultados de la evaluación muestran que el método proporciona resultados coherentes para las dos componentes (GHI y DNI), con errores bajos para cualquier probabilidad de excedencia. Además, el método preserva la estadística de larga duración a alta resolución temporal.

Por otra parte, respecto a la predicción del recurso se ha analizado en profundidad el uso del modelo numérico de predicción meteorológica (NWP) Weather Research and Forecasting (WRF) para la predicción de la irradiancia solar en superficie. La predicción del recurso es esencial para la integración de la energía solar en la estructura de suministro eléctrico, el cual debe operar siempre bajo los principios de seguridad y estabilidad. Asimismo, es fundamental para la operación y gestión de las plantas, así como para la venta de la energía. Por tanto, es clave para mejorar decisivamente el nivel de competitividad de la energía renovable de origen solar. Dentro de las herramientas para la predicción de la radiación solar, los modelos NWP destacan como las herramientas

más potentes. En particular, el modelo WRF es uno de los más avanzados que existen actualmente. Sin embargo, en general los modelos NWP no están diseñados específicamente para las aplicaciones en energía solar. Así, en esta tesis se presenta primeramente una evaluación integral de las predicciones de GHI y DNI obtenidas con el modelo WRF. El análisis se lleva a cabo distinguiendo diferentes condiciones de cielo, estaciones del año y horizontes de predicción. En un segundo paso, las predicciones son evaluadas dentro del marco de referencia establecido en un ejercicio de intercomparación con otros modelos NWP. Finalmente, en base a los resultados obtenidos en los trabajos anteriores, se lleva a cabo un estudio para analizar el papel que juega la resolución espacial horizontal respecto a la calidad de las predicciones de radiación solar. También se evalúa dicha calidad después de aplicar a las salidas del modelo un pos-proceso basado en un promediado espacial. En general los resultados muestran que WRF tiende a sobreestimar la radiación. Además, la calidad de las predicciones, medida en términos de los errores estadísticos habituales, resulta mejor para resoluciones espaciales más bajas. Por su parte, los resultados de la intercomparación muestran que las predicciones de los modelos globales presentan errores más bajos que las de los modelos regionales como el WRF. No obstante, se concluye que la conveniencia del uso de un tipo de modelos u otros depende de la aplicación final.

Annex C

Contenidos

List of acronyms

Annex D

Capítulo 1. Antecedentes y marco de investigación

1.1 Introducción

Hoy en día el mundo está siendo testigo de una revolución en el sector energético. Ello obedece a varios factores entre los que destacan especialmente el fuerte incremento continuo de la demanda de energía, particularmente de energía eléctrica, y la concienciación social sobre el cambio climático debido al incremento de los gases de efecto invernadero como consecuencia directa de la actividad humana (IPCC , 2014a). Ambos elementos están parcialmente interrelacionados -la transformación de energía primaria es uno de los procesos más importantes de emisiones antropogénicas de gases de efecto invernadero-. Sin embargo, la percepción de la importancia de estos dos aspectos es desigual, ya que el aumento de la demanda de energía afecta directamente a la economía y sus consecuencias se perciben más rápidamente. Mientras que el cambio climático tiene efectos que aparecen de forma más gradual, pero son constantes y podrían ser catastróficos. Por lo tanto, estas dos motivaciones principales, junto con el incesante avance en el conocimiento y la tecnología, están favoreciendo definitivamente la transformación actual de los paradigmas energéticos del último medio siglo.

El cambio climático

El cambio climático es uno de los mayores retos a los que enfrenta el futuro de la humanidad. Con el fin de estudiar sus causas y prever sus posibles consecuencias, la comunidad internacional estableció el Grupo Intergubernamental de Expertos sobre el Cambio Climático

(IPCC) para llevar a cabo dichos estudios y cumplir el objetivo de proponer medidas para mitigar sus efectos. Dentro de esta organización y fuera de ella, la comunidad científica también ha estado realizando un importante esfuerzo para lograr este propósito. El IPCC considera el cambio climático como cualquier cambio en la media y/o la variabilidad de las propiedades del estado del clima durante un período prolongado debido a cualquier causa. La consideración de la Convención Marco de las Naciones Unidas sobre el Cambio Climático (UNFCCC) restringe las causas sólo a aquellas relacionadas, directa o indirectamente, con la actividad humana. En ambos casos el hecho es que las evidencias demuestran que la atmósfera inferior de la Tierra está sufriendo un calentamiento global promedio. En palabras del IPCC: "la evidencia científica del calentamiento del sistema climático es inequívoca". Esto se mide en todo el mundo como un incremento constante de las temperaturas medias globales del aire y de los océanos. Este calentamiento global promedio, junto con sus efectos, modifica el estado del clima y constituye lo que se llama cambio climático. Las primeras consecuencias inmediatas del mismo son actualmente visibles, como el derretimiento de vastas extensiones de nieve y hielo, el aumento del nivel medio del mar y las alteraciones en los patrones vitales de algunos animales y plantas. El IPCC describe que los efectos del cambio climático también están afectando a la agricultura, los ecosistemas terrestres y oceánicos y a los suministros de agua, y destaca que esto está ocurriendo en todo el mundo (IPCC, 2014b). Además, la Organización Mundial de la Salud (WHO) afirma que el cambio climático tiene un impacto en la salud humana, ya que afecta a la calidad del aire, al agua potable y a la agricultura. La WHO dice que "se prevé que el cambio climático causará aproximadamente 250.000 muertes anuales adicionales por malnutrición, enfermedades y estrés térmico" entre los años 2030 y 2050. En Europa, un informe reciente de la Agencia Europea del Medio Ambiente (EEA) advierte que se espera un aumento de la fenómenos meteorológicos extremos en el continente debido al cambio climático: olas de calor, inundaciones, sequías y tormentas más frecuentes e intensas (EEA, 2017). También señala que ya ha habido impactos negativos en la economía y la salud. En el sur de Europa, como es el caso de España, la EEA prevé un aumento de las

temperaturas máximas, sequías, inundaciones e incendios forestales, así como menos precipitaciones. Se espera también que las temperaturas más altas favorezcan la propagación de insectos que provoquen la redistribución e incremento de enfermedades. Para hacer frente a esta realidad en diciembre de 2015, la Conferencia de las Partes de la UNFCCC aprobó el Acuerdo de París -firmado por 195 países- para actualizar el Protocolo de Kioto de 1997. El objetivo del acuerdo es mantener " el aumento de la temperatura en este siglo muy por debajo de los 2 grados centígrados, e impulsar los esfuerzos para limitar el aumento de la temperatura incluso más, por debajo de 1,5 grados centígrados sobre los niveles preindustriales". La meta es mantener las condiciones climáticas relativamente estables y minimizar los impactos del cambio climático. El umbral de 2°C es el valor acordado por la comunidad científica para preservar tales condiciones. Con el fin de asegurar un escenario más seguro se propuso el valor de 1,5°C.

Existen diferentes factores que podrían contribuir, de manera individual o agregada, a la alteración de la temperatura global promedio. En este sentido, por ejemplo ha habido siete ciclos de glaciación en los últimos 650.000 años. Estos ciclos son producidos por cambios en la cantidad total de energía recibida en la Tierra desde el Sol debido a ligeras variaciones en la órbita terrestre. También hay otros factores naturales que pueden contribuir a cambios en el sistema climático, como la actividad del Sol, los volcanes y las fluctuaciones naturales de la concentración de gases de efecto invernadero (GHG) en la atmósfera. De manera natural, parte de la energía de la radiación solar recibida -aproximadamente el 30%- es reflejada de nuevo hacia el espacio exterior por las nubes, los aerosoles y la superficie de la Tierra. Sin el efecto invernadero de la atmósfera, la temperatura en la superficie sería aproximadamente -18 ° C, en lugar de los 15 ° C reales. Sin embargo, hoy en día las actividades humanas han emitido suficiente dióxido de carbono y otros gases de efecto invernadero como para incrementar artificialmente el efecto invernadero atmosférico natural, produciendo un calentamiento global más allá de estos 15 ºC, lo que evidencia una causa antropogénica determinante del cambio climático (IPCC, 2014a). En particular, el dióxido de carbono (CO_2) -uno de los

GHG que más contribuye al calentamiento global- ha aumentado notablemente su presencia en la atmósfera; Y continúa aumentando. Al comienzo de la era industrial, la concentración era de 278 ppm, lo que se considera un valor de referencia equilibrado del sistema climático. En 2016, la Organización Meteorológica Mundial (WMO) anunció que en 2015 se alcanzó el hito de una concentración media mundial de 400 ppm de CO_2, la cual se mantuvo incluso después del fin del fuerte fenómeno de El Niño. Actualmente, a principios de 2017, observatorios de referencia, como los de Manua Loa en Hawai (NOAA, EEUU) e Izaña en Tenerife (AEMet, España), están registrando valores cercanos a los 410ppm, estableciendo un nuevo récord. Según la Administración Nacional Oceánica y Atmosférica (NOAA) de los Estados Unidos, la tasa anual de aumento es de 1,92 ppm. Al mismo tiempo, observaciones recientes de anomalías de temperatura medidas por diferentes organismos científicos muestran que el valor de seguridad de 1,5 ºC podría estar más cerca de lo esperado. Actualmente, el incremento de la temperatura global durante el período preindustrial de referencia es de aproximadamente 1 ºC. Según Pachauri *et al*. (2014), en la situación actual, los escenarios proyectados muestran que es posible que la temperatura media global de la superficie sea 4 ° C más alta que en el período preindustrial para el final del siglo XXI. Además, una publicación reciente de Climate Central (CC) muestra que si las tendencias actuales de las emisiones de GHG continúan, es posible que el umbral de 1,5 ° C se cruce entre los años 2025-2030 (CC, 2016). Además, un trabajo reciente de investigación (Crowther *et al*., 2016) también ha concluido que el cambio climático podría estar ocurriendo mucho más rápido de lo que se pensaba, ya que la liberación de CO_2 almacenado en el suelo se ve reforzada por el calentamiento climático. Los autores concluyeron que, en un escenario conservador, el aumento de las emisiones de CO_2 de las reservas de suelo debido al calentamiento climático podría ser alrededor del 12 al 17% de las emisiones antropogénicas esperadas para 2050. Aunque hay una incertidumbre significativa, este trabajo demuestra que este aumento adicional sobre el CO_2 total emitido a la atmósfera estimulará un efecto de retroalimentación positiva y puede contribuir decisivamente a acelerar el cambio climático. Este resultado obliga a revisar las

proyecciones actuales del cambio climático para incorporar estos nuevos valores de emisiones no tenidos en cuenta previamente. Crowther *et al*. (2016) indicaron también que es posible que el aumento del efecto invernadero sea ahora irreversible, ya que el sistema podría haber pasado el punto de no retorno. En cualquier caso, el IPCC advierte que "sin mitigación adicional, e incluso con la adaptación, el calentamiento a fines del siglo XXI llevará a un riesgo muy alto de impactos severos, generalizados e irreversibles a nivel mundial". Por lo tanto, es imperativo llevar a cabo acciones decisivas para mitigar los efectos del cambio climático.

De acuerdo con este escenario descrito, la comunidad internacional se ha comprometido a promover políticas coordinadas para hacer frente a esta realidad. En este sentido, las estrategias eficaces se basan en la reducción de los impactos del cambio climático y también en la realización de esfuerzos específicos para adaptarse a él. Con este fin, la premisa es que el mundo tiene que acelerar intensamente la reducción de las emisiones de GHG. Entre las políticas específicas previstas por cada país, el Acuerdo de París está promoviendo la adopción de los denominados sistemas de fijación de precios del carbono (como los planes de comercio de emisiones y los impuestos sobre el carbono) por los países firmantes, en particular por las economías más grandes del mundo. Se espera que esta propuesta reduzca fuertemente las emisiones globales y así mitigue el cambio climático. Además, es un marco favorable para las energías renovables. Hoy en día, la mayoría de las emisiones antropogénicas están relacionadas con la actividad de la industria y la transformación de la energía. En particular, el 25% de las emisiones antropogénicas de GHG son generadas por los procesos de producción de electricidad y calor (IPCC, 2014a). Por lo tanto, las energías renovables –como la energía solar- son eminentes reductores activos de las emisiones de GHG. Entonces, la promoción de estas fuentes de energía es uno de los activos más importantes para mitigar los efectos del cambio climático a través de un desarrollo sostenible.

Escenario energético

A pesar de que los impactos del cambio climático deberían ser suficientemente fuertes para motivar la evolución del sector energético, el factor económico asociado a este campo tiene un peso decisivo, así como necesariamente determinante, porque es el motor que impulsa esta transformación. En este sentido, el marco actual de las energías renovables y su proyección futura son de una importancia clave. A este respecto, la Agencia Internacional de la Energía (IEA), en su Informe sobre las Perspectivas Energéticas Mundiales 2016 (IEA, 2016a), predice un incremento global de la demanda de energía primaria del 30% -sobre todo de países en desarrollo- para 2040. La electricidad aumentará más fuertemente que cualquier otra energía de uso final. En particular, en el caso de referencia, el 37% de la generación de energía será a partir de energías renovables en comparación con el 23% actual. La tasa esperada de aumento de las energías renovables no hidroeléctricas –principalmente el viento y la energía solar- es de un promedio anual del 2,9% desde 2012 a 2040, siendo la fuente de energía para la generación de electricidad que más rápido crece. De esta forma, las energías renovables no hidroeléctricas pasarán del 5% de la generación mundial total en 2012 al 40% en 2040. Representarán casi el 50% de la capacidad instalada en el período 2015-2020 (IEA, 2015). Al mismo tiempo, se espera un fuerte crecimiento del gas natural y –mucho menos pronunciado- del consumo de petróleo. En el caso de referencia, se calcula que las emisiones de carbono debidas al sector energético reducirán su tasa media anual del 2,4% desde 2000 al 0,5% en 2040. No obstante, la IEA reconoce que, a pesar de ser un logro significativo, esta tasa no permite alcanzar el objetivo del escenario de 2ºC del Acuerdo de París. Sólo limitará el aumento de las temperaturas globales medias a 2,7°C para el año 2100. Por lo tanto, cumplir con los objetivos climáticos para reducir las emisiones y mejorar la eficiencia será un reto muy difícil. No obstante, la IEA afirmó que, si bien requeriría mayores esfuerzos, es posible alcanzar el objetivo del Acuerdo de París mediante políticas adecuadas que aceleren el crecimiento de las tecnologías con baja emisión de carbono y la eficiencia energética. En este sentido, la IEA destaca que es muy importante intentar ampliar el uso de las energías renovables a otros

sectores clave desde el punto de vista del consumo de energía, como la industria, la construcción y el transporte. Por su parte, en la Unión Europea (EU) el compromiso con el desarrollo de energías limpias es muy importante. Un objetivo común, vinculante para todos los Estados miembros, se ha fijado para alcanzar un 20% de consumo de energía procedente de fuentes renovables para 2020, así como un acuerdo para reducir las emisiones de gases de efecto invernadero desde al menos un 20% (2009/28 / CE y 2009/29 / CE). Según Eurostat, la participación de las energías renovables en la producción total de energía primaria fue del 25,4% en 2014, un incremento del 73,1% desde 2004 (incremento medio del 5,6% anual). La contribución a la electricidad total generada a partir de fuentes de energías renovables fue del 27,5%. En este sentido, España es uno de los países líderes en la implantación de las energías renovables. Es el cuarto productor de energía renovable en la UE-28, con un 9,4% de la producción total. La contribución de estas fuentes de energía al *mix* eléctrico es ya muy importante. Según el operador de la red de transporte de España (Red Eléctrica de España, REE), el porcentaje de generación de electricidad a partir de fuentes renovables sobre la generación total en la España peninsular fue del 41,1% (REE, 2016). Además, la energía renovable instalada ha permitido que el saldo neto de intercambio de electricidad –con interconexión de alta tensión- entre España y sus países vecinos (Marruecos, Portugal y Francia) sea exportador (REE, 2015).

El marco descrito anteriormente pone de manifiesto el decidido impulso que las energías renovables están recibiendo globalmente, y en particular en España y el resto de Europa, por parte de los gobiernos y del sector energético. Así, la AIE señala que la generación de electricidad está entrando en un período de transformación hacia un entorno de desarrollo seguro y sostenible en el que las energías renovables juegan un papel importante, sobre todo la energía eólica y solar. Por un lado este tipo de energías está aumentando progresivamente su nivel de competitividad gracias al desarrollo de la tecnología, lo que se traduce en una reducción gradual de sus costes de generación e inversión. Por otro lado, ha permitido establecer un nuevo mapa de la distribución de los recursos energéticos mediante la introducción de fuentes abundantes y geográficamente distribuidas,

como la radiación solar y el viento. Esto permite suministrar parte de la demanda de energía a través de recursos propios y finalmente reducir el costo de la factura energética de los países y su dependencia externa. Esto es de particular interés en un país como España, cuya dependencia externa del suministro de energía es de alrededor del 80% (20 puntos por encima de la media europea). Por lo tanto, teniendo en cuenta las expectativas de un consumo creciente, un despliegue adecuado de energías renovables es una estrategia conveniente y un elemento clave en las proyecciones económicas de cualquier país. A este respecto, la IEA subraya que, a pesar de los actuales precios más bajos del petróleo –algunos expertos señalan que los precios nunca rebotarán más allá de los 100 dólares por barril-, la producción de energía de fuentes renovables se ha expandido a su tasa más rápida en 2015 (IEA, 2016b). Esto ha sido posible gracias en parte a las políticas energéticas favorables de los gobiernos. Pero también en gran parte gracias a un importante desarrollo tecnológico y, en consecuencia, a una fuerte reducción de costes en los procesos de fabricación de bienes de equipo. Por ejemplo, los costes de producción de tecnología solar han disminuido un 80% en el período 2008-2015. Por lo tanto, el desarrollo tecnológico de las energías renovables las hace económicamente más rentables, aumentando su competitividad frente a otras fuentes de energía más maduras, como los combustibles fósiles. Sin embargo, la IEA hace hincapié en que hay riesgos a considerar. La financiación es un parámetro clave para proporcionar una inversión sostenida en el sector de las energías renovables. Por lo tanto, es crucial proporcionar un marco regulatorio seguro y estable a largo plazo y reducir las incertidumbres legales. Las políticas poco fiables pueden comprometer el ritmo de despliegue de las energías renovables socavando la confianza de los inversores. Además, la IEA destaca que los problemas asociados a la integración de las energías renovables variables en los sistemas eléctricos y en el mercado son una prioridad crítica para la política energética. En particular, la IEA destaca como una acción estratégica el mejorar la predicción avanzada de los recursos renovables como parte de las estrategias de operación.

Energía solar

En este escenario, la energía solar jugará un papel fundamental. Dentro de las energías renovables es probablemente la única con potencial suficiente para cubrir de manera sostenible las expectativas de energía a largo plazo del planeta (Perez, 2008). La energía radiante que llega a la Tierra desde el Sol es el principal motor del sistema climático que impulsa la circulación atmosférica (Pozo-Vázquez *et al*., 2004). Esta fuente de energía primaria supera las necesidades energéticas actuales en aproximadamente 1.500 veces (Perez, 2008), siendo el recurso energético más abundante del planeta. Por esta razón, una parte importante del esfuerzo dedicado al desarrollo de las energías renovables se centra en la explotación de este recurso (Szuromi *et al*., 2007). La generación de energía solar ha aumentado enormemente en los últimos años gracias a una mayor eficiencia de conversión de energía y a los menores costes de producción e instalación. Se espera que la energía renovable basada en tecnologías que exploten la irradiancia solar en superficie sea un elemento central en la futura generación eléctrica. Por ejemplo, este es el caso de la EU. Según Eurostat, la producción de energía solar en la EU sigue siendo relativamente baja, representando una cuota del 6,1% de la producción total de energía renovable en 2015. Sin embargo, el crecimiento de la energía solar ha sido el mayor de todas las renovables; **s**u contribución a la generación eléctrica total de renovables aumentó del 0,1% al 10,0%, en el período 2004-2014. Así, por ejemplo, Europa ha alcanzado recientemente el valor de 100GW de potencia instalada de tecnología fotovoltaica (PV). Por su parte, España es un referente en el uso de la energía solar, ya que es uno de los países pioneros en la producción eléctrica de origen solar en todo el mundo (IEA, 2016a). En particular, España es líder mundial en el desarrollo y operación de plantas basadas en la denominada tecnología de concentración de energía solar (CSP), con aproximadamente 2300MWe en operación (Fernández-García *et al*., 2010), seguidos por los Estados Unidos de América**, c**on 1700MWe. En cuanto a la tecnología fotovoltaica, el país líder es China, con una capacidad instalada de 78GWe. España es el décimo en el ranking con una capacidad instalada de 5.5GWe. Para alcanzar el objetivo acordado por los países de la EU sobre el consumo

de energía, España ha desarrollado un nuevo Plan Nacional de Energías Renovables para la próxima década (2011-2020). Este plan estima una contribución de las energías renovables al consumo energético en España del 22,7% en 2020, con una contribución a la producción de electricidad del 42,3%. Ambos porcentajes superan ampliamente los objetivos fijados por la EU. Se espera que una parte importante del aumento estimado sea de origen solar.

La irradiancia solar que llega a la superficie terrestre es muy variable en el espacio y en el tiempo, esencialmente debido a factores geométricos de la posición relativa entre el Sol y la Tierra y a las condiciones meteorológicas. Desde el punto de vista de la capacidad de ciertas tecnologías para suministrar electricidad bajo demanda, la energía solar se considera una energía renovable variable (VRE), como la energía eólica, la energía de las olas y de las mareas y la energía hidroeléctrica de los ríos (IEA, 2008). En contraste, hay energías renovables que pueden ser clasificadas como tecnologías estables, a saber: energía hidroeléctrica de embalse, biomasa, geotérmica y, en un grado menor, algunas tecnologías de CSP que tienen almacenamiento térmico de sal fundida. Esta característica natural de la radiación solar, junto con su relativo primer estadio de desarrollo, reduce su competitividad frente a otras fuentes de energía. Como se mencionó anteriormente, el creciente desarrollo tecnológico ha reducido notablemente los costes de la tecnología solar. No obstante, los problemas asociados con la incertidumbre de la fuente de energía primaria siguen siendo un factor determinante en el coste final de los proyectos solares. En este sentido, un concepto fundamental para los proyectos solares es la bancabilidad. De forma concisa, un proyecto es bancable si los inversionistas, públicos o privados, consideran que hay suficientes garantías de inversión, por lo que están dispuestos a financiarla. Esto significa que es probable que un proyecto bancable asegure el éxito financiero con el mayor grado de confiabilidad posible. En un proyecto solar uno de los elementos principales para asegurar el éxito del proyecto es la capacidad o nivel de producción -hay otros como el precio final de la energía-, que depende directamente del recurso solar. De esta manera, hay dos aspectos a considerar en relación

con el recurso solar: la evaluación realizada durante las primeras etapas para evaluar la viabilidad del proyecto y la previsión durante la fase operativa.

En cuanto a la evaluación de los recursos solares, la energía solar esperada en la ubicación de interés durante el tiempo de vida de la instalación se evalúa cuidadosamente, porque es la fuente más alta de incertidumbre –este parámetro se traduce en riesgo y este en interés financiero, dicho de forma simplificada-. Para ello, se analizan las series temporales de largo plazo de datos históricos y se contemplan diferentes escenarios de disponibilidad de energía basados en consideraciones estadísticas. Hoy en día esto se hace mediante la generación de series temporales artificiales de datos de recurso solar que tienen como objetivo recopilar toda la información estadística de la serie histórica de largo plazo en un período de un solo año, llamados años meteorológicos típicos (*vide infra* sección 1.2 *Revisión del estado del arte*). Por lo tanto, la importancia de la fiabilidad de estos datos y los escenarios generados es incuestionable.

Por otra parte, mientras que los costes de producción son cercanos a cero y muy estables –la irradiancia solar en superficie es mucho más abundante de lo que se puede utilizar y está disponible libremente-, el carácter variable de la energía solar complica significativamente su integración en los sistemas de suministro eléctrico a gran escala, aumentando los costes asociados. El origen del problema reside en el hecho de que la energía eléctrica no puede ser almacenada a gran escala. Por lo tanto la generación, el transporte y el consumo de electricidad deben ser coordinados -por los TSO- y llevados a cabo al mismo tiempo. En el caso de una VRE, como la energía solar, la intermitencia en la generación constituye un desafío notable. La integración exitosa de la energía solar debe abordarse de dos maneras. La primera se refiere a los sistemas de suministro de electricidad, que deben adaptarse para obtener flexibilidad para poder responder con fiabilidad y rapidez a las fluctuaciones de la oferta y la demanda (IEA, 2008). Además, deben seguir evolucionando desde las redes aisladas a los mercados nacionales e internacionales para construir un sistema

energético más sostenible y seguro (IEA, 2016c). La segunda forma es conocer de antemano –con la máxima certeza posible- la producción de energía solar, para que los TSO puedan programar y controlar el funcionamiento del sistema salvaguardando las máximas exigencias de estabilidad y seguridad. Para ello, una predicción fiable de la disponibilidad del recurso solar a corto y medio plazo -horas a días- es clave para las operaciones de los TSO, así como para los productores de energía solar para el funcionamiento y la gestión de las instalaciones solares, así como para la comercialización de la energía.

En lo que respecta a los sistemas de conversión de energía solar, hoy en día existen esencialmente dos tecnologías principales con diferentes características que operan integradas en los sistemas de suministro de energía (Quaschning, 2004): plantas fotovoltaicas (PV) (IEA, 2014a) y plantas de concentración (CSP) –media y alta temperatura- (IEA, 2014b). Los sistemas fotovoltaicos transforman directamente la irradiancia solar global en electricidad a través de dispositivos semiconductores -células fotovoltaicas-. Las plantas CSP transforman la energía de la componente normal directa de la radiación (DNI) en calor mediante el uso de dispositivos que enfocan la irradiancia solar en los receptores, iniciando así un ciclo termodinámico que transforma el calor en energía mecánica y luego en electricidad –electricidad solar térmica (STE)-. El uso del calor tiene una ventaja desde el punto de vista del almacenamiento de energía. Algunas plantas CSP permiten tener un sistema de almacenamiento de energía térmica basado en sales fundidas, que es capaz de producir hasta alrededor de 10 horas de capacidad de generación a plena carga. Las tecnologías empleadas podrían ser: colectores cilindro parabólicos, receptor central, reflector Fresnel lineal y disco Stirling. Hoy en día la mayor planta CSP en operación es Ivanpah Solar Power Facility en EE.UU., pero las potencias instaladas de 50MW y 100MW son bastante habituales. Por otro lado, los parques fotovoltaicos más grandes actualmente alcanzan potencias nominales instaladas de más de 500MW, siendo Kurnool Ultra Mega Solar Park en India actualmente la mayor planta fotovoltaica del mundo, con 900MW en operación. Finalmente se debe mencionar una tecnología que tiene un despliegue mucho menor pero

que alcanza la mayor eficiencia en la transformación de energía fotovoltaica, el concentrador fotovoltaico (CPV). Esta tecnología utiliza DNI como fuente de energía primaria, pero a diferencia de la CSP, se concentra la luz solar en una célula fotovoltaica. Esta tecnología parece tener un futuro prometedor. Además de los altos valores de eficiencia que genera, la necesidad de sistemas fotovoltaicos más pequeños reduce los costes del sistema, lo que mejora su competitividad. Sin embargo, el CPV todavía tiene que resolver desafíos tecnológicos importantes para ser competitiva frente a la PV o la CSP.

En resumen, la investigación aplicada sobre la irradiancia solar como fuente de energía primaria para aplicaciones de energía solar es de máxima importancia para cumplir el objetivo de aumentar su competitividad y favorecer un mayor despliegue. Para ello se destacan dos aspectos: i) la reducción de la incertidumbre en la evaluación del recurso solar y ii) la mejora de la previsibilidad del recurso solar para facilitar la integración en las estructuras de suministro de energía a gran escala. Así, además de las políticas favorables a su desarrollo dentro de un marco jurídico estable y la adaptación de las estructuras de suministro energético para asimilar las VRE, el estudio del recurso solar y su aplicación a la industria solar es sin duda fundamental para favorecer la expansión de esta fuente de energía, ya que la explotación de su enorme potencial será decisiva para alcanzar los objetivos del Acuerdo de París en la lucha contra el cambio climático y el logro de un desarrollo más sostenible.

1.2 Revisión del estado del arte

La investigación en el campo de la radiación solar y su interacción con los componentes del sistema climático está en constante desarrollo. A pesar de que existen muchos conceptos sólidos, el desarrollo de las aplicaciones solares ha traído nuevas mejoras, actualizaciones, suplementos y la creación de otros nuevos, cubriendo todos los aspectos relativos a las dos dimensiones del recurso solar desde el punto de vista de las aplicaciones solares: la evaluación y la predicción. Así, los fundamentos básicos de las investigaciones de recurso solar

abarcan diferentes aspectos interrelacionados, tales como instrumentación, medidas, control de calidad de datos, modelado y post-procesado. Hay otros conectados más directamente con la industria solar, como el desarrollo de productos solares para los usuarios finales y la bancabilidad. A continuación se presenta una descripción general del actual estado del arte de los elementos esenciales de este campo en continuo progreso.

1.2.1 Consideraciones generales

Radiación solar en la superficie de la tierra

La luz solar en la superficie terrestre presenta una distribución espectral variable de energía que cubre longitudes de onda que oscilan entre 0,3 y 4 μm. Esta gama contiene la mayor parte de la energía electromagnética total del Sol que llega a la superficie terrestre, y es referida habitualmente como radiación solar de onda corta o, simplemente, radiación solar. Por lo general, la radiación solar se describe en términos de su proyección sobre la superficie horizontal, que es el parámetro conocido como irradiancia global horizontal (GHI). Representa el flujo de energía total recibido por unidad de área. La GHI es utilizada por las tecnologías fotovoltaicas, pero proyectada sobre la superficie del panel; En ese caso es usual denominarla como irradiancia global inclinada (GTI) (Perez *et al*., 1990a; Gueymard, 2009a; Ineichen, 2011a). La irradiancia directa es la radiación solar procedente del disco solar. En la práctica, sin embargo, su medición también contiene la radiación solar procedente de un ángulo sólido centrado en el disco solar, llamada radiación circunsolar (Blanc *et al*., 2014). La irradiancia directa se describe generalmente sobre una superficie normal a la dirección de propagación, y por lo tanto se denomina irradiancia normal directa, DNI. Esta es la fuente de energía primaria de interés para las tecnologías de concentración, a saber: CSP y CPV. Una tercera componente es la irradiancia difusa, que es la irradiancia procedente de la luz solar dispersa por los constituyentes de la atmósfera. Por lo general se nombra en términos de su proyección horizontal, es decir, la irradiancia horizontal difusa. La DNI junto con la irradiancia difusa están relacionadas con la GHI por la ecuación 1.1,

donde Z es el ángulo cenital solar y dhi es la irradiancia horizontal difusa. Finalmente, cabe mencionar que las consideraciones espectrales, que son de importancia significativa para tecnologías como la CPV, no se incluyen aquí, ya que quedan fuera del alcance de esta investigación. Por lo tanto, en el contexto de este trabajo se define el recurso solar en términos de la radiación solar espectral de banda ancha de interés para las aplicaciones en energía solar, a saber: GHI y DNI.

$$GHI = DNI \cdot cosZ + dhi \tag{1.1}$$

La irradiancia solar que llega a la superficie terrestre es muy variable en el espacio y el tiempo. Su distribución es función de varios factores. Principalmente, hay factores deterministas, que dependen de las características geométricas de la posición relativa de la Tierra en su órbiata alrededor del Sol. Las variaciones de estos parámetros modifican la insolación recibida por la Tierra en cualquier momento, siendo responsables del día y la noche, las estaciones y otras variaciones a largo y muy largo plazo que ocurren durante miles de años –la interacción gravitacional de la Tierra con otros cuerpos del sistema solar causa cambios en su excentricidad orbital, inclinaci y precesión-. Durante períodos geológicos cortos, la irradiación recibida en la parte superior de la atmósfera (TOA) –la llamada irradiancia extraterrestre- permanece aproximadamente constante, determinada por el espectro solar que cambia con la actividad del Sol. Tiene un valor recientemente actualizado de 1361.2 W/m^2 –la integral de la irradiancia solar sobre todo su espectro-, conocida como la Irradiancia Solar Total (TSI) o también conocida anteriormente como la constante solar (Gueymard, 2006; Gueymard, 2012b). Sin considerar la atmósfera, la distribución de esta energía sobre las tierras y los océanos tiene lugar según la latitud y los efectos topográficos locales (Ruiz-Arias *et al*., 2010a). Sin embargo, la existencia de la atmósfera introduce adicionalmente un conjunto de elementos que interactúan con la radiación solar como una especie de filtro dinámico que altera drásticamente el campo de radiación inicialmente homogéneo. Estas interacciones tienen lugar naturalmente por medio de procesos físicos

de dispersión y absorción de la radiación solar por los constituyentes atmosféricos, a saber: moléculas de gas, partículas de aerosol y gotas y partículas de nubes. De los gases presentes en la atmósfera, el 98% corresponde a N_2 y O_2 y el 2% restante está constituido por Ar, vapor de agua y gases traza: CO_2, CH_4, O_3, N_2O, CO y compuestos clorofluorocarbonados (CFC). El efecto en la radiación solar de este 2% de los gases constituyentes es muy importante. Por su parte, las partículas de aerosol tienen una amplia gama de tamaños –de menos de 0,1 μm hasta más de 20 μm- y distribución de formas (Olmo, 2008). Las nubes están formadas mayoritariamente por gotitas de agua y cristales de hielo. Las fluctuaciones naturales en la cantidad y distribución de estos elementos en la atmósfera son responsables de la fuerte variabilidad en el espacio y tiempo de la irradiancia solar en la superficie terrestre. En particular, entre los procesos de extinción –absorción y dispersión- producidos por estos constituyentes atmosféricos, no hay uno más relevante para las aplicaciones de energía solar que el producido por las nubes. Esto se debe a la rápida variabilidad espacial y temporal de las nubes, que no es igualada por ninguna otra especie atmosférica (Mayer, 2009). Por lo tanto, las nubes son la mayor fuente de incertidumbre para la evaluación y previsión de los recursos solares (Kim y Ramanathan, 2008; Ruiz-Arias *et al*., 2016). A continuación, los aerosoles son la segunda fuente de incertidumbre más grande, siendo la primera bajo condiciones de cielo despejado. Además, los aerosoles actúan como núcleos de condensación para las gotas de nubes, lo que puede aumentar la cantidad de cobertura nubosa y también influir en el tiempo de vida de las nubes. Entre los componentes de la radiación solar, la DNI es en gran medida la más afectada por su interacción con las nubes y los aerosoles, por lo que es más difícil de estimar y típicamente presenta mayores valores de incertidumbre que la GHI (Bellouin *et al*., 2005; Ruiz-Arias *et al*., 2015a). En consecuencia, esto afecta más a las tecnologías basadas en la energía solar concentrada, que utilizan la DNI como fuente de energía primaria. En resumen, los constituyentes atmosféricos, particularmente las nubes, introducen una variabilidad extremadamente alta en el campo de radiación solar que finalmente alcanza la superficie terrestre, tanto en el espacio como en el tiempo,

con variaciones que pueden ser ya significativas para aplicaciones solares en rangos de metros y minutos. La contribución final de todos estos factores, tanto deterministas como estocásticos es generalmente, la atenuación de la TSI inicialmente disponible; aunque a veces puede incluso ser un efecto de aumento temporal debido a fenómenos de reflexión sobre nubes muy blancas y brillantes.

Modelización de la radiación solar

La modelización de la radiación solar es esencial para las aplicaciones en energía solar. Existe una amplia variedad de métodos que van desde enfoques estadísticos simples hasta sofisticadas técnicas físicamente fundamentadas. La adecuación del método depende de la aplicación concreta y de la información disponible (Ruiz-Arias y Gueymard, 2015b). En este sentido, por ejemplo, la posición solar es un problema importante en las aplicaciones de radiación solar. Para los sistemas de seguimiento solar, la ubicación precisa del Sol a lo largo de su trayectoria diaria determinista en el cielo –definida por sus ángulos azimutal y cenital- siempre debe ser conocida con alta precisión; en particular la tecnología CPV requiere una precisión extremadamente alta. Con este fin, actualmente se dispone de diferentes algoritmos, que difieren en su nivel de exactitud y complejidad –usualmente traducidos en un mayor coste computacional- (Michalsky, 1988; Blanco-Muriel *et al*., 2001; Reda y Andreas, 2004, 2008; Grena, 2008; Blanc y Wald, 2012). De estos algorítmos, el de Reda y Andreas (2008) es el más preciso, pero también es el más costoso desde el punto de vista computacional. Otro parámetro geométrico clave para los cálculos de la radiación solar es la llamada masa de aire, que explica la longitud de la trayectoria a través de la atmósfera que es atravesada por los rayos solares con respecto a una trayectoria vertical. Para altas elevaciones del Sol, la aproximación de una atmósfera plano-paralela podría ser suficiente. Sin embargo, para grandes ángulos del cenit solar, la aproximación no resulta apropiada, por lo que luego se deben aplicar correcciones empíricas para tener en cuenta la refracción de los rayos solares (Kasten y Young, 1989; Gueymard, 2003).

Dos de las aproximaciones más útiles en la modelización de la radiación solar para aplicaciones solares son los modelos de cielo despejado y los llamados modelos de separación para todas las condiciones de cielo. El primer grupo de modelos, llamado también modelos de transferencia radiativa de cielo despejado (CSRT), pretende representar fielmente la curva diaria de las componentes de la radiación solar en ausencia de nubes en cualquier lugar sobre la superficie terrestre. Para este fin actualmente existen varios enfoques. Una de las principales diferencias entre ellos son los parámetros de entrada necesarios para calcular la irradiancia solar, que en todos los casos es un número reducido de parámetros tales como el agua precipitable, la profundidad óptica de aerosol (AOD) en determinadas longitudes de onda, el exponente de Ångström, la presión en la superficie o el coeficiente de turbiedad de Linke (TL). Ejemplos destacables de estos modelos de banda ancha son: el modelo ESRA (Rigollier *et al*., 2000), la Versión Simplificada de SOLIS (Ineichen, 2008a) y el REST2 (Gueymard, 2008a). Diversos estudios exhaustivos han evaluado de forma comparativa diferentes modelos respecto a medidas de estaciones en tierra (Ineichen 2006). Estudios más recientes han demostrado que el REST2 obtiene los mejores resultados, seguido de cerca por la versión simplificada del SOLIS (Gueymard, 2012a; Badescu *et al*., 2012). Los resultados de estos modelos son extremadamente precisos, incluso en el rango de incertidumbre de los instrumentos, pero siempre que los parámetros de entrada sean correctos. Más recientemente se ha implementado un nuevo modelo, llamado modelo McClear (Lefèvre *et al*., 2013). A diferencia de otros modelos de banda ancha, este se basa en la interpolación de la irradiancia solar a partir de tablas pre-calculadas. Por otra parte, debido a su prominencia debe mencionarse también un ejemplo de modelo espectral, denominado SMARTS (Gueymard, 2001, 2005). La aplicación de estos modelos en la evaluación y previsión de recursos solares es fundamental para la estimación de la GHI y la DNI en condiciones de cielo despejado, cuando las instalaciones solares pueden alcanzar los mayores valores de producción. También son muy útiles para la evaluación de la calidad de los datos, verificación de la referencia temporal, para interpolación de datos, generación de modelos de persistencia inteligente e incluso para

la verificación preliminar del funcionamiento de la instrumentación, entre muchas otras aplicaciones.

Respecto a los modelos de separación para todas las condiciones de cielo, debe decirse que actualmente son de una importancia clave para las aplicaciones de la energía solar. Estos modelos separan la GHI en sus componentes directa y difusa. Esto es muy útil para aprovechar las mediciones mucho más abundantes de GHI. Además, en la actualidad los métodos más establecidos para estimar la irradiancia solar superficial a partir de mediciones de satélite se han desarrollado para obtener la GHI. Por lo tanto, todos los modelos de separación para todas las condiciones de cielo desempeñan un papel fundamental para proporcionar una DNI basada en las mediciones de satélite. De hecho, estos modelos son en realidad los métodos empleados por algunos de los proveedores de datos de irradiancia solar más reputados en la industria solar. No obstante, estos modelos no son la única solución para la estimación de DNI a partir de medidas de satélite (*vide infra* Sección 1.2.2 *Evaluación de la radiación solar*). Además, también se utilizan para proporcionar predicciones de DNI cuando los modelos operados no proporcionan esta variable entre sus salidas. Además, debido a que se basan en enfoques empíricos, los modelos de separación son relativamente fáciles de implementar y ejecutar. En resumen, estos modelos, a pesar de que no son indispensables, siguen siendo un elemento fundamental para las aplicaciones de energía solar. Por lo tanto, evaluar el rendimiento de estos modelos es de gran interés. Un trabajo importante fue realizado en este sentido en primer lugar por Ineichen (2008b). Sin embargo, más recientemente, Gueymard y Ruiz-Arias (2016) han realizado un enorme esfuerzo de evaluación en una trabajo publicado donde analizan en profundidad 140 modelos de separación, y que se continúa en posteriormente en el trabajo de Aler et al. (2017). Los resultados muestran que este tipo de modelos depende en gran medida de las condiciones particulares del lugar donde se lleva a cabo la evaluación. Esto se debe al hecho de que cada modelo de separación se ajusta con un conjunto particular de datos y, por lo tanto, están condicionados por la información contenida en dichos conjuntos de datos. Así pues, es bastante difícil aventurar cuál sería el

rendimiento esperado de cualquier modelo en cualquier lugar. De todos modos, el estudio muestra que algunos modelos se destacan sobre otros. Tal es el caso de los modelos de Pérez (1992, 2002) y Engerer (2015). Otros ejemplos de diferentes enfoques son: Reindl *et al*. (1990), Elminir *et al*. (2007), Boland *et al*. (2008) y Ruiz-Arias *et al*. (2010b).

Evaluación de la irradiancia solar

Por último, tanto para la evaluación de la radiación solar como para la predicción es esencial analizar y contrastar los valores de la radiación solar modelada frente a las observaciones en tierra. En esencia, el procedimiento de evaluación consiste en una comparación de los datos que se están examinando con una referencia equivalente obtenida mediante mediciones apropiadas, que se consideran como la "verdad" y que comúnmente se denominan valores reales u observaciones (Oreskes, 1994, 1998). Cualquier evaluación debe hacerse bajo las consideraciones necesarias que aseguren la confiabilidad de los resultados. Esto significa esencialmente que las mediciones de referencia deben tener propiedades de fiabilidad, exactitud y estabilidad. De lo contrario degradan el valor de los resultados y, en consecuencia, la validez de las conclusiones. Por ejemplo, se pueden obtener mediciones de alta calidad de GHI mediante un *piranómetro*. Sin embargo, este tipo de instrumentos no pueden evitar un error asociado con el propio diseño del dispositivo, que aparece a ángulos zenitales solares elevados – error instrumental de coseno-. Por lo tanto, si la evaluación no se limita a ángulos de elevación solar por encima de cierto umbral –que depende del instrumento-, la incertidumbre en la GHI observada se agrega a la evaluación, afectando a la fiabilidad de los resultados. En ese caso, es mucho más conveniente limitar el alcance de la evaluación a cambio de obtener resultados más fiables. Por lo tanto, se entiende fácilmente que las mediciones de radiación solar son de una importancia clave en las aplicaciones de la energía solar. Por un lado, durante la fase de planificación, los datos de alta calidad son indispensables para apoyar el proyecto. En este sentido, por lo general, al menos un año de mediciones locales es requerido por las instituciones financieras para corregir las series temporales necesarias a largo plazo, que normalmente se derivan con datos satelitales y pueden

presentar altos valores de incertidumbre (*vide infra* Sección 1.2.2 *Evaluación del recurso solar*). Por otro lado, las bases de datos de calidad ganan un valor mucho mayor con el tiempo. Así, durante la fase operativa de la planta, estos datos son de gran importancia para llevar a cabo un control continuo de su rendimiento. En consecuencia, la selección del instrumento adecuado es fundamental, así como la selección de la mejor ubicación (Stoffel *et al*., 2010; Vignola *et al*., 2012). La mejor elección del instrumento depende de las características de la aplicación para la que se vaya a utilizar. También está subordinada a las limitaciones impuestas por las condiciones de instalación y la posibilidad de llevar a cabo las actividades necesarias de operación y mantenimiento. A este respecto, existen varias publicaciones especializadas de estudios comparativos que analizan el rendimiento de diferentes tipos de radiómetros (Badosa *et al*., 2014; Gueymard y Myers, 2009b; Michalskby *et al*., 2005, 2011; Vuilleumier *et al*., 2014; Geuder *et al*., 2014). Idealmente, los instrumentos deben ser confiables, duraderos, fáciles de mantener y baratos. La operación y mantenimiento de la instrumentación de alta calidad para medir la irradiancia solar es una tarea muy exigente si se pretende mantener toda la calidad potencial de los instrumentos. Esto debe incluir un programa de calibraciones periódicas, que son esenciales para asegurar la calidad de los datos a lo largo del tiempo (Gueymard y Myers, 2008b). Además, es una práctica recomendada, cuando es posible, tener mediciones redundantes de instrumentos independientes en la misma estación. Esto es muy útil para la verificación de la calidad de los datos y para recuperar los datos que pudieran perderse. En última instancia, los datos obtenidos en el tiempo siguiendo las directrices que garantizan la calidad son extraordinariamente valiosos. Sin embargo, a pesar de todos los esfuerzos, la medición de la irradiancia solar es una tarea compleja y delicada que no está exenta de posibles errores –la suciedad, la pérdida de seguimiento, etc.-, que afectan a la calidad final del conjunto de datos. En este sentido, la verificación de la calidad de los datos debe ser siempre un procedimiento estándar a realizar cuando se trabaja con datos de irradiancia solar, no sólo para analizar periodos prolongados de datos, sino también para llevar a cabo comprobaciones en tiempo casi real. Este conjunto de procedimientos tiene como

objetivo detectar posibles errores en los datos, estableciendo el indicador correspondiente asociado a cada uno, el cual describe un nivel de confianza en la medida, desde completamente correcto hasta ciertamente erróneo (Wilcox *et al*., 2012). Posiblemente, el conjunto más popular de controles de calidad es el propuesto por la Red de Referencia de Radiación en Superficie (BSRN) (Long y Shi, 2008; Roesch *et al*., 2011). Sin embargo, hay otras posibilidades que añaden límites más o menos restrictivos para validar los datos (Ineichen, 2014; Gueymard y Ruiz-Arias, 2016). Entre estos controles automáticos, siempre se debe realizar una inspección visual. En cualquier caso, para la comprobación de la calidad de los datos, una práctica muy recomendable es salvar los datos con un muestreo de muy alta frecuencia, entre 1 y 5 minutos.

Una vez que las observaciones están caracterizadas correctamente, se puede realizar el procedimiento de evaluación. Para ello, se establece un conjunto de parámetros objetivos para evaluar cuantitativamente las diferencias entre los datos estimados y las mediciones reales. Entre los parámetros estadísticos disponibles, hay tres que son de uso estándar, a saber: error de desviación medio (MBE), error absoluto medio (MAE) y error cuadrático medio (RMSE). Todos describen una característica importante de los errores en un solo valor promediado. Así, el MBE explica el error sistemático, permitiendo saber si los datos sobreestiman o subestiman la irradiancia solar. El MAE añade siempre la diferencia entre los datos estimados y los observados, siendo el más útil para dar cuenta de las desviaciones esperadas en la predicción de la irradiancia solar y, por tanto, para las posibles sanciones consecuentes. El RMSE, relacionado con el MBE a través de la desviación estándar, explica las desviaciones, pero penalizando sobre todo las más grandes. Por lo tanto, estos parámetros sólo proporcionan una visión promedio de los errores, pero permite una evaluación fácil. Además, proporcionan una manera sencilla de establecer inter-comparaciones entre diferentes fuentes de datos frente a las mismas observaciones en un marco común, lo que permite conocer el rendimiento de tales fuentes de información de manera relativa. Además de estos parámetros populares, existe un conjunto completo de métodos para evaluar la irradiancia solar,

diseñado para analizar la información relativa al conjunto de datos examinado. Recientemente, Gueymard (2014) y Jensen *et al*. (2016) han llevado a cabo una revisión completa de todos estos métodos. Por último, cabe mencionar que es una práctica común para evaluar el pronóstico de irradiancia solar utilizar un modelo de referencia simple, como el modelo de persistencia o el modelo de persistencia inteligente. Dado que este tipo de modelo es el más simple de usar, se espera siempre que sea superado por un método más sofisticado.

1.2.2 Evaluación del recurso solar

La evaluación del recurso solar implica una colección de métodos para caracterizar la radiación solar normalmente disponible en un lugar de interés. En este sentido, no es una simple cuantificación de la irradiancia solar disponible, sino una descripción exhaustiva que incluye un detallado análisis estadístico y un estudio exhaustivo de la incertidumbre asociada. Además, es un campo interdisciplinario en el que varias disciplinas deben trabajar juntas, como la modelización de la irradiancia solar, la radiometría, la metrología, la meteorología, la climatología, la geografía, la ingeniería, la teledetección, la estadística y la economía. En este sentido, una valoración fiable del recurso solar junto con un análisis correcto de la incertidumbre es fundamental para favorecer la bancabilidad del proyecto. Además, la reducción de la incertidumbre confiere al promotor confianza en el proyecto, ya que proporciona estimaciones más fiables de la producción anual, lo que facilita la búsqueda de inversores y el hacer frente a los procesos de licitación. Sin embargo, la evaluación del recurso solar ha tenido un desarrollo lento hasta los últimos años, cuando el impulso de la industria para satisfacer sus necesidades ha conducido a un progreso más rápido y más consistente.

Para caracterizar el recurso solar con la menor incertidumbre, se requieren mediciones a largo plazo de irradiancia solar de alta calidad. Sin embargo, estas mediciones nunca están disponibles en la práctica. Por lo tanto, los modelos basados en satélites (Perez *et al*., 2013a; Miller *et at*., 2013) son ampliamente utilizados hoy en día, ya que

pueden proporcionar tanto una cobertura espacial completa como series de datos históricos a largo plazo. Alternativamente, es posible utilizar métodos físicos más sofisticados que utilizan la transferencia radiativa y la modelización atmosférica (Jones y Fletcher, 2013). Este es el caso, por ejemplo, de los reanálisis atmosféricos, tales como el MERRA2 de la NASA, el ERA-Interim del ECMWF o el CFSR de la NOAA. Sin embargo, estos modelos son generalmente menos precisos que los basados en datos de satélite (Boilley y Wald, 2015). No obstante, incluso estos últimos modelos presentan con más frecuencia una mayor incertidumbre que las observaciones directas, con un amplio rango de variación, dependiendo de las características locales del sitio de interés. Por lo tanto, para caracterizar mejor el recurso solar, se utiliza una variedad de métodos, combinando mediciones *in situ* y datos modelados a largo plazo derivados de satélites. Varios factores influyen en la selección de los métodos a aplicar, como el propósito del estudio, los requisitos y la disponibilidad de datos, entre otros. Ruiz-Arias y Gueymard (2015b) han propuesto un conjunto de mejores-prácticas para la evaluación del recurso solar. Estas son: la selección de una ubicación adecuada para la instalación de la estación radiométrica, la recogida de al menos 12 meses de datos, la adquisición de datos de satélite históricos a largo plazo de irradiancia solar, el solventar las lagunas debido a datos erróneos y a la carencia de datos, la evaluación de la calidad de los datos medidos y la incertidumbre, la evaluación de las observaciones satelitales y terrestres, la corrección a largo plazo de los datos satelitales y la derivación de información útil para la caracterización del recurso solar como parámetros estadísticos del conjunto de datos a largo plazo, el TMY y la incertidumbre total.

Estimaciones basadas en satélites

Para cubrir periodos de largo plazo (más de 15 años) y casi cualquier parte del planeta, los satélites son la mejor opción para estimar la irradiancia solar superficial. Transportan instrumentos radiométricos que detectan la radiancia multiespectral reflejada y emitida por la Tierra, tanto de la superficie como de la atmósfera, con altas resoluciones espaciales y temporales. Para derivar la irradiancia solar recibida por la superficie terrestre a partir de la reflectancia

medida por el sensor de satélite, se han propuesto varios enfoques a lo largo de los años. Hoy, los métodos más comunes se pueden clasificar según su naturaleza. Por lo tanto, existen métodos físicamente fundamentados y métodos semi-empíricos –también conocidos como métodos de índice de nubosidad (Ruiz-Arias y Gueymard, 2015b). El primer tipo de métodos tiene probablemente un mayor potencial para ser mejorado en los próximos años, aunque es más complejo y requiere una mayor exigencia computacional. Además, al igual que los modelos CSRT, requiere información precisa del estado atmosférico actual, que no siempre está disponible con la exactitud requerida (Habte *et al*., 2013; Miller *et al*., 2013). Por otra parte, los métodos semi-empíricos siguen siendo los más utilizados para las aplicaciones en energía solar (Perez *et al*., 2002, 2013a; Polo *et al*, 2008; Rigollier *et al*, 2004; Lefèvre *et al*., 2007). En estos métodos, la GHI se estima superponiendo la cantidad de de las propiedades ópticas de la nube derivada de la radiancia medida por el satélite sobre la GHI de cielo despejado obtenida con un modelo CSRT. La DNI se obtiene entonces por medio del modelo de separación para todas las condiciones de cielo. Por lo general, estos métodos son adaptados por los investigadores y los proveedores de datos de recurso solar de acuerdo con sus propios criterios con el fin de reducir la incertidumbre de las estimaciones. Ineichen (2014) presentó un estudio comparativo de siete métodos semi-empíricos. De todas formas, todavía subsisten varias fuentes de incertidumbre (Cebecauer *et al*., 2011). Más recientemente, se han propuesto otros enfoques, basados en técnicas de inteligencia artificial (AI), como las redes neuronales artificiales (ANN) (Quesada-Ruiz *et al*., 2015, y Linares-Rodríguez *et al*., 2015). A diferencia de los métodos semi-empíricos, los métodos basados en AI pueden proporcionar directamente tanto GHI como DNI, y están limitados por la cantidad de información disponible para llevar a cabo el entrenamiento del modelo.

Adaptación al sitio

También denominada unión de conjuntos de datos o extensión de registros medidos, el término adaptación al sitio designa un conjunto de métodos aplicados para corregir el error sistemático en el conjunto de

datos a largo plazo, mediante un conjunto de medidas realizadas en un perido relatiamente corto (Suri y Cebecauer, 2011; Bender *et al*., 2011, Thuman *et al*., 2012). El objetivo es reducir la incertidumbre original de la estimación del satélite. Con este fin, los diferentes métodos calibran los datos del satélite a largo plazo frente a las observaciones locales durante el período de superposición. Esta estrategia reduce los errores aleatorios y, en general, los sesgos o errores sistemáticos. Actualmente, este es un requisito estándar para la bancabilidad de grandes proyectos solares.

Hoy en día, existen varios enfoques para llevar a cabo la adaptación al sitio. Todos ellos se basan en el uso de conjuntos de datos de observaciones de tierra de irradiancia solar a corto plazo –generalmente se requiere por lo menos un año-. Estos métodos pueden ser estadística o físicamente fundamentados. Polo *et al*. (2016) presentan una revisión del estado del arte de los métodos de adaptación al sitio.

Año meteorológico típico

Los años meteorológicos típicos (TMY) son series temporales anuales artificiales –normalmente de valores horarios o intra-horarios- de variables útiles para los sistemas de aprovechamiento de la energía solar, cuyo objetivo es condensar toda la información de largo plazo en un solo año. De esta manera, este año artificial tiene las variaciones naturales diurnas y estacionales, y teóricamente representa un año de condiciones climáticas típicas en la ubicación de interés. En el contexto de las aplicaciones del recurso solar, el objetivo del TMY es preservar las características estadísticas que caracterizan al recurso solar en la ubicación de interés, para describir el comportamiento esperado del sistema de energía solar, idealmente durante la vida útil de la instalación (Stoffel *et al*., 2010, Sengupta *et al*., 2015). En este sentido, la representatividad del TMY se ve beneficiada directamente por el uso de datos mejorados a largo plazo. Originalmente fue introducido en 1978 por el laboratorio Sandia National Lab y posteriormente por el NREL, como una solución a las limitaciones de las capacidades computacionales que hacía difícil relaizar complicadas simulaciones en energía solar. Desde entonces se ha convertido en una herramienta de

uso común en la industria solar. Los TMY son necesarios para complementar el análisis de bancabilidad de proyectos solares, como en plantas CSP y PV. Sin embargo, hoy en día no se recomienda su uso para el diseño de las plantas.

Por lo tanto, los TMYs son fundamentales en estudios de evaluación integral del recurso solar. Las características teóricas de este año sintético facilitan la obtención de la estimación cuantitativa objetiva de la producción esperada de energía solar en el lugar de interés. En este sentido, es una propiedad clave de los TMYs, ya que permite a los especialistas estimar los valores de la irradiación solar anual esperada en diferentes escenarios de probabilidad de la cantidad de energía solar disponible. Tales escenarios son usualmente medidos en términos de la probabilidad de excedencia –habitualmente también denominada por el nombre de su concepto complementario, el percentil- y la incertidumbre estimada correspondiente, que incluye la incertidumbre asociada a la estimación de los datos a largo plazo. Estas valiosas propiedades son de gran interés para la industria de la energía solar, ya que son información clave para evaluar el riesgo económico del proyecto solar, que finalmente se traduce en interés financiero. Por lo tanto, se suelen requerir los TMY de varios escenarios para llevar a cabo el análisis de viabilidad, usualmente el promedio –probabilidad de superación o excedencia del 50%- y los casos pesimistas o de baja energía –probabilidad de excendecia del 75%, 90%, 95% o incluso 99% (Cebecauer y Suri, 2015).

El TMY no corresponde a un determinado año de un determinado período, sino que es un año del calendario construido artificialmente como una composición ponderada, estadísticamente fundamentada, de doce meses seleccionados de la serie temporal histórica de largo plazo de la ubicación de interés. Habitualmente, los pesos se utilizan para tener en cuenta no sólo la variable principal de interés para generar el TMY –la GHI o la DNI, dependiendo de la aplicación-, sino también variables meteorológicas auxiliares, tales como la temperatura, la velocidad del viento, la humedad relativa, etc., que son de interés para el análisis de la producción. Sin embargo, los expertos no coinciden en

la conveniencia de tener en cuenta las variables meteorológicas y cómo hacerlo para ajustar la composición de pesos, ya que no hay pruebas concluyentes de tal conveniencia. Así, Habte *et al*. (2014) del NREL han propuesto una nueva configuración de pesos en la que las principales variables para aplicaciones PV o CSP, es decir, GHI y DNI respectivamente, tienen el 100% del peso. Ellos nombran el TMY resultante como año típico de global (año de irradiancia global horizontal, TGY) y año típico de directa (irradiancia normal directa, TDY). En el contexto de este trabajo de investigación, para referirse a ambos al mismo tiempo se utiliza el término año solar típico (TSY). Cabe señalar que el TMY no tiene por qué tener necesariamente más información que los TSY, sino ligeramente distinta, ya que estos últimos pueden ser completados simplemente agregando la información meteorológica correspondiente. La diferencia es que el TMY tiene en cuenta –de forma no consensuada- las variables meteorológicas para generar el año, mientras que los TSY solo consideran la componente de irradiancia solar de interés para formarlo.

Por lo tanto, desafortunadamente no se ha alcanzado aún un consenso científico sobre un método estándar para generar los TMY o los TSY. Por el contrario, hay un conjunto variado de métodos y estrategias seguidas por diferentes autores y proveedores para generar sus propios TMY. Por lo tanto, ambos métodos y pesos pueden resultar en diferentes TMY para la misma ubicación y el conjunto de datos a largo plazo, lo que es una situación no deseada. En este contexto, algunos esfuerzos de investigación analizan los resultados de diferentes métodos para generar TMY en aplicaciones para la energía solar (Ineichen, 2011b; Realpe *et al*., 2016).

Incertidumbre

La incertidumbre es probablemente el concepto individual más importante en la evaluación del recurso solar. Es esencial para evaluar la calidad de las estimaciones y la clave para la bancabilidad de los proyectos solares. La incertidumbre es, por lo tanto, indispensable para cualquier análisis riguroso en aplicaciones de energía solar para obtener conclusiones exhaustivas. La incertidumbre en el recurso solar es

equivalente a la incertidumbre en el rendimiento esperado de la planta solar. Además, la incertidumbre en la estimación del recurso solar es la fuente más grande de la incertidumbre total de la producción final de energía.

Sin embargo, es un parámetro que a veces no está completamente claro en la industria de la energía solar. No debe confundirse con la variabilidad de los datos. En realidad, la incertidumbre total es una composición de las contribuciones individuales de todos los elementos involucrados en la caracterización de la irradiancia solar: medidas, modelado, variabilidad interanual, variabilidad espacial, representatividad del período y método de adaptación al sitio (Meyer *et al*., 2009). Además de todo esto, para el TMY también se debe tener en cuenta la incertidumbre derivada del propio método (Fernández-Peruchena *et al*., 2016). Las características de cada componente de la incertidumbre dependen del aspecto específico de la evaluación del recurso solar. Para añadir todas estas contribuciones (u_i) para obtener la incertidumbre total (U) es común utilizar la ley de Gauss de propagación de errores (ecuación 1.2).

$$U = \sqrt{\sum (u_i)^2} \tag{1.2}$$

1.2.3 Predicción de la radiación solar

La predicción de la radiación solar para aplicaciones en energía solar es un tema relativamente reciente. El progreso desde los primeros estudios hasta el estado actual del arte ha sido considerable, cubriendo actualmente las dos componentes –GHI y DNI- y las diferentes dimensiones espaciales y temporales de la materia. De todos modos, es una investigación en curso de capital importancia para la integración a gran escala de la energía solar en los sistemas de suministro de energía. Hoy en día, las predicciones fiables de la producción de energía solar son una herramienta básica para la gestión de la red en países con un nivel de penetración considerable, como es el caso de España. El Operador del Sistema Eléctrico (TSO) español utiliza actualmente

previsiones de producción para las tecnologías PV y CSP con el fin de mejorar el control del sistema en términos de estabilidad de la red. También las usa para llevar a cabo la gestión de la complementariedad con otras fuentes variables de energía. En general, los TSO son responsables del equilibrio seguro de la red. Además, durante la fase operativa de las plantas solares, las previsiones de producción son de una importancia clave para la operación y la gestión eficientes de las instalaciones, así como para el venta de energía. Dado que la incertidumbre de la producción de energía solar se debe principalmente a la incertidumbre en la previsión de la irradiancia solar, la importancia de un pronóstico confiable del recurso solar es incuestionable. En este sentido, varios estudios han evaluado el valor de la predicción operacional de la irradiancia solar para aplicaciones energéticas (Dumortier *et al*., 2009). Algunos de estos trabajos de investigación se centran en una tecnología particular, como los estudios de Pérez *et al*. (2007) y Marcos *et al*. (2013) para la PV, y los de Wittmann *et al*. (2008); Kraas *et al*. (2013) y Law *et al*. (2016) para la CSP.

Una forma conveniente de describir básicamente la predicción de la radiación solar es tener en cuenta su dimensión temporal. En este sentido, se deben considerar tres momentos principales con respecto a la predicción, a saber: horizonte de predicción, resolución temporal de la predicción y frecuencia de actualización de la predicción. Hay otro tiempo que debe tenerse en cuenta en situaciones prácticas, que es el tiempo de entrega previsto. En el uso operacional y para el comercio de energía, el tiempo en que la previsión debe estar disponible es una limitación importante a considerar. En cuanto al horizonte de predicción, que está relacionado con los requisitos de aplicación en los que se va a utilizar la predicción, este influye decisivamente en las metodologías más apropiadas a aplicar. De esta manera, a lo largo de la dimensión temporal, la predicción puede dividirse en tres partes principales, con límites difusos. Los primeros 10 a 20 minutos, de 4 a 6 horas, y hasta varios días. En Inman *et al*. (2013), Diagne *et al*. (2013) y Kleissl (2013), se puede encontrar una revisión del actual estado del arte de los métodos de predicción. Antonanzas *et al*. (2016) llevó a

cabo una revisión de los métodos de predicción para la producción de la PV.

Predicción intra-horaria

La predicción intra-horaria se utiliza para la operación de la planta solar. Por lo general, esta predicción se realiza mediante cámaras de todo el cielo y mediciones en tierra. El horizonte de predicción depende en realidad del tipo de nubes, la velocidad de desplazamiento y la altura de la base de la nube. Estas previsiones normalmente alcanzan resoluciones espaciales y temporales muy altas, de metros y minutos. También la actualización es muy alta, del orden de 1 minuto. Este método se basa en el supuesto de un predominio de la advección horizontal de las nubes. Es decir, durante un período de minutos muy corto la forma de la nube y sus propiedades ópticas no cambian significativamente, mientras que la velocidad horizontal, detectada por intercomparaciones de dos imágenes consecutivas del cielo, se supone constante y por lo tanto el movimiento de la nube es extrapolado hacia adelante. Así, se estima la posición futura de las nubes (Urquhart *et al*., 2013, 2015; Quesada-Ruiz *et al*., 2014). Este método, basado en la determinación de los llamados vectores de movimiento de la nube (CMV), es también la base del siguiente paso en el horizonte de tiempo de la predicción.

Nowcasting

La predicción a corto plazo con horizonte temporal de hasta varias horas, también conocida como *nowcasting* para diferenciarla del tercer tipo de horizonte de predicción, también se utiliza principalmente para el funcionamiento de la planta. Como se basa en imágenes de satélites geoestacionarios, la resolución temporal nominal es la del satélite, con actualizaciones que van desde 15 a 30 minutos. Teóricamente, el horizonte de predicción puede extenderse más allá de 6 horas, pero en la práctica después de 4 a 6 horas la fiabilidad de la predicción de este método disminuye por debajo de la fiabilidad de los modelos numéricos de predicción meteorológica (NWP). Este método también utiliza la misma estrategia de detección de los vectores de movimiento de las nubes (CMV), pero, en este caso, a partir de imágenes de satélite. Esta

diferencia es esencial, lógicamente, cambiando las características del método, tales como la cobertura espacial, las resoluciones y el intervalo de tiempo. Las propiedades de las nubes y la estimación de la irradiancia se obtienen normalmente, como se ha expuesto antes, por medio de métodos semi-empíricos (*vide supra* sub-sección 1.2.2 *Evaluación de la radiación solar*) (Hammer *et al*., 2003; Hoff, 2013b). Müller y Remund (2013) proponen un método combinado de satélite y NWP. Utilizan la forma de la nube y la información de la posición del satélite y utilizan la información del viento proporcionada por el modelo NWP. Un enfoque similar es el propuesto por Miller *et al*. (2013), pero, en este caso, utilizan un método físico para recuperar las propiedades de la nube y la irradiancia solar. Por último, los métodos más recientes aprovechan las ventajas de la información mejorada disponible sobre la nube que puede ser advectada y difundida por un NWP. Este método híbrido presenta mejores resultados que los tradicionales basados en CMV (Arbizu-Barrena *et al*., 2017).

Predicción

Los pronósticos a corto y medio plazo –simplemente denominados predicción- se llevan a cabo generalmente mediante modelos de NWP. Los ejercicios comparativos han demostrado que esta técnica es la más eficaz para los horizontes de predicción más allá de 4 a 6 horas, cuando la correlación del estado real de la atmósfera y el estado futuro es baja. Entonces, un método físicamente fundamentado como el modelo NWP funciona mejor.

Los modelos NWP son, sin duda, la herramienta más poderosa para la predicción de la irradiancia solar, porque están construidos sobre principios físicos y son capaces de resolver el estado completo de la atmósfera en cada paso temporal en el futuro. Son un instrumento muy complejo, cuyas limitaciones son continuamente superadas gracias a los esfuerzos de investigación para mejorar la modelización física y una potencia computacional cada vez mayor. Hoy en día, su potencial parece incomparable. Para las aplicaciones de energía solar, no sólo pueden proporcionar previsiones de los principales componentes de la irradiancia solar –la GHI directamente y la DNI puede ser derivada-,

sino también para un conjunto completo de variables meteorológicas de interés para la producción de energía solar, tales como la temperatura, la humedad relativa o la velocidad del viento y su dirección. Por otra parte, ya que cubren los horizontes horarios más grandes, son la herramienta adecuada para el gestión de la planta y el comercio de la energía.

Básicamente, los modelos de NWP son una implementación software de los procesos físicos que tienen lugar en la atmósfera y el resto de elementos del sistema climático, así como la interacción entre ellos. Por ejemplo, esto incluye la formación, advección, difusión y disipación de las nubes y la transferencia de la irradiancia solar desde la parte superior de la atmósfera hasta la superficie terrestre. Las ecuaciones diferenciales de los modelos de NWP son capaces de resolver el estado de la atmósfera para determinadas resoluciones espaciales –grid- y temporales. Sin embargo, hay muchos procesos clave que ocurren a escalas subgrid. Por ejemplo, algunas estructuras de nubes pueden tener fácilmente tamaños más pequeños que las típicas resoluciones del grid. Entonces se dice que no pueden ser vistos por el modelo, o en otras palabras, no pueden ser resueltos. Por lo tanto, se espera que mayores resoluciones reduzcan el error en la representación de los procesos modelados. Sin embargo, todavía hay detalles desconocidos sobre el rendimiento del modelo respecto a la resolución espacial para ciertos procesos. De todos modos, para tener en cuenta los fenómenos subgrid, los modelos se complementan constitutivamente con los esquemas de parametrizaciones (Stensrud, 2007). Se trata de implementaciones software que describen fenómenos físicos que se producen a resoluciones superiores a las nominales del modelo. Por ejemplo, la radiación solar en los modelos NWP es un ejemplo de proceso parametrizado.

Los modelos de NWP requieren conocer el estado inicial de la atmósfera y las condiciones de los límites para integrar las ecuaciones diferenciales y así lanzar la atmósfera hacia delante en la dimensión temporal. Esta información se basa en observaciones mundiales que se recogen y procesan rutinariamente para producir lo que se denomina

análisis. Hoy en día, este proceso suele tardar entre 6 y 12 horas en completarse. Con este estado inicial de la atmósfera, los grandes centros de predicción global hacen uso de sus propios modelos –llamados modelos de circulación global (GCM)- para generar el pronóstico del tiempo mundial. Algunos ejemplos de GCM son: el Sistema de Predicción Global (GFS) producido por el Centro Nacional para la Predicción del Medio Ambiente (NCEP, EEUU), el Sistema de Predicción Integrado del Centro Europeo de Pronóstico del Tiempo a Medio Plazo (ECMWF, UE), el Sistema de Predicción (GDPS) del Centro Meteorológico Canadiense (CMC) y el Sistema Goddard de Observación de la Tierra Versión 5 (GEOS-5) desarrollado conjuntamente por la NOAA, el NCEP y el EMC (EEUU). La resolución temporal en la que se difunden actualmente sus productos oscila entre 1 y 3 horas. Las resoluciones espaciales típicas son de 25 km, a excepción del IFS que es de aproximadamente 12 km. Estos modelos producen predicciones fiables de la irradiancia solar, con bias bajos. Sin embargo, la resolución temporal nominal, entre 1 a 3 horas, es baja para aplicaciones en energía solar. Normalmente se requiere una una resolución temporal horia para la integración en los sistemas de suministro de energía, pero las resoluciones intra-horarias son las preferidas por los productores para estimar mejor la producción, particularmente en las plantas CSP. Por tanto, se requiere llevar a cabo una interpolación temporal cuando se emplean los GCM para aplicaciones en energía solar (Lorenz *et al*., 2009a). Además, por lo general, este tipo de modelos no proporciona la DNI como salida. Por lo tanto, se debe derivar por medio de un modelo externo, por ejemplo, los modelos de separación para todas las condiciones de cielo. Esta podría ser una de las razones por las que esta componente ha sido menos estudiada que la GHI. Otra razón importante es que es mucho más difícil de tratar. Además la tecnología fotovoltaica está mucho más extendida que la basada en concentración. Asimismo, disponer de medidas fiables de DNI es más complicado y son más escasas.

Utilizando el análisis y las predicciones globales proporcionadas por los GCM como condiciones iniciales y de contorno, se pueden ejecutar modelos regionales NWP para proporcionar una predicción de un

conjunto completo de variables meteorológicas con resoluciones espaciales y temporales muy altas, típicamente en el orden de pocos kilómetros y unos pocos minutos. Entre todos los modelos regionales, el modelo WRF (Skamarock *et al*., 2008) es uno de los más avanzados y ampliamente utilizados en todo el mundo. Se apoya en un esfuerzo colaborativo del NCAR y otras instituciones de los EEUU. También se desarrolla gracias a las contribuciones de la comunidad científica en todo el mundo. El WRF tiene un extenso conjunto de parametrizaciones que dotan al modelo de gran flexibilidad y adaptabilidad a las condiciones meteorológicas y geofísicas específicas de una región en particular.

Sin embargo, los modelos NWP están destinados a producir predicción del tiempo. Por lo tanto, no se diseñan específicamente para aplicaciones en energía solar. Esto es evidente por las limitaciones expuestas anteriormente con respecto a la DNI. Pero también hay otros aspectos relacionados con la irradiancia solar que no se consideran plenamente en los NWP. Esto se debe esencialmente a dos circunstancias principales. Por un lado, la falta de información de parámetros determinados que son esenciales para las interacciones con la radiación solar, como los aerosoles. Por otro lado, debido a que la radiación solar en el modelo es muy costosa en términos de cálculo y por lo tanto los cálculos radiativos tienden a ser simplificados tanto como sea posible. Sin embargo, la creciente importancia del recurso solar ha promovido iniciativas para avanzar en este campo. De esta manera, el modelo WRF ha evolucionado recientemente con un compendio de actualizaciones dirigidas a mejorar la predicción solar para aplicaciones energéticas. En conjunto, estas actualizaciones constituyen una configuración particular del modelo WRF conocido como WRF-Solar (Jimenez *et al*., 2016). En concreto, WRF-Solar está diseñado para mejorar la representación de las nubes (Thompson y Eidhammer, 2014; Deng *et al*., 2014) y aerosoles (Ruiz-Arias *et al*., 2014), así como las interacciones de estos elementos clave con la radiación solar. Los autores han evaluado el modelo bajo condiciones de cielo despejado y han obtenido resultados prometedores, mejorando significativamente aquellos obtenidos para las tres componentes de la

radiación solar estimadas con la versión estándar del WRF. También demuestran que es necesario incorporar la influencia de los aerosoles para obtener estimaciones precisas de la irradiancia solar. Esta conclusión también ha sido encontrada por Ruiz-Arias *et al*. (2013).

Actualmente, se han realizado numerosos estudios para evaluar el rendimiento de varios modelos bajo diferentes condiciones y regiones. Algunos ejemplos de estos trabajos de investigación son: Gaston *et al*. (2009), Lorenz et al. (2011, 2012), Mathiesen y Kleissl (2011), Pelland *et al*. (2013), Isvoranu y Badescu (2015), Aryaputera *et al*. (2015), Zempila *et al*. (2016), He *et al*. (2016), Lima *et al*. (2016) y Sosa-Tinoco *et al*. (2016).

Existen también otros métodos posibles para la predicción de la radiación solar. Se asientan esencialmente en enfoques estadísticos. Estos métodos se basan en el supuesto de que los patrones en los conjuntos de datos históricos se repiten en el futuro. Por lo tanto, requieren series temporales históricas de datos para analizar y recuperar tales patrones. Se pueden aplicar varias técnicas para aprovechar esta persistencia, como por ejemplo los modelos autorregresivos integrados de media móvil (ARIMA). Los métodos más modernos se basan en técnicas de inteligencia artificial, tales como las redes neuronales artificiales, los *k-neighbours* o *support vector machines*. Una descripción de varios métodos estadísticos utilizados para la predicción de la irradiancia solar ha sido presentada por Coimbra y Pedro (2013) y Diagné *et al*. (2013). La combinación de los métodos de predicción estadística y física es de especial interés, ya que proporciona herramientas potentes que se aplican a las salidas proporcionadas por los modelos físicos como los NWP. Estos métodos permiten reducir los posibles errores sistemáticos del modelo físico. Pueden tener en cuenta los efectos locales y también incluyen información adicional sobre parámetros externos. Además, se pueden utilizar para derivar información que no está incluida en la salida original del modelo. El conjunto de técnicas que combina métodos estadísticos y modelos físicos usualmente se denomina técnicas de post-procesado o estadísticas de la salida del modelo (MOS). Estos pueden incluir

técnicas simples tales como filtros de suavizado por medio del promediado espacial (Lorenz *et al*., 2009a).

Por último, es importante mencionar la conveniencia de realizar intercomparaciones entre las previsiones de diferentes fuentes. Esto se adapta para establecer un marco de referencia general en el que se puedan evaluar diferentes métodos frente a otros. En este sentido, los ejercicios de evaluación comparativa son una práctica muy conveniente. Larson (2013) y Lorenz *et al*. (2015) han llevado a cabo un estudio de las intercomparaciones entre varios modelos NWP.

1.3 Motivaciones

La energía solar renovable está aumentando su presencia en el suministro de energía eléctrica a gran escala. Sin embargo, hoy en día su importancia ponderada en el mix energético sigue estando lejos de su potencial teórico. Sin embargo, como se mencionó anteriormente, sus expectativas de crecimiento para las próximas décadas son las más altas entre todas las fuentes de energía renovable. Este crecimiento prospectivo estará impulsado por dinámicas tecnológicas y económicas, que están indefectiblemente interconectadas. La tasa de penetración estará asociada con la capacidad de proporcionar soluciones que puedan alinear los avances de los conocimientos técnicos con los beneficios económicos, más allá de las iniciativas de los distintos gobiernos e instituciones públicas. De esta manera, hoy en día muchos de los esfuerzos en el campo de la energía solar se centran en la reducción de costes, mientras protegen la seguridad del suministro de electricidad. El presente trabajo de investigación se centra en aquellos relativos al recurso solar.

Dentro del ámbito de la energía solar, además de la tecnología de la explotación misma, hay un aspecto eminentemente esencial, a saber, el recurso solar. Esta fuente primaria de energía –la irradiancia solar que llega a la superficie terrestre- y los factores meteorológicos que le afectan determinan naturalmente de forma directa los aspectos más importantes asociados a la energía solar, que son el desarrollo e

integración de esta fuente renovable en el sistema eléctrico, tanto en los niveles de generación como de distribución. Además, el recurso solar condiciona también la tecnología de explotación a emplear, siendo determinante para todos ellas. En los últimos años se han incrementado el interés y, paralelamente, la investigación asociada a este tema en sus dos dimensiones principales: la evaluación y la previsión del recurso solar. La mayor parte de esta investigación sigue una trayectoria claramente práctica, proporcionando soluciones y respuestas a cada problema particular planteado por la industria solar. Sin embargo, hay muchas preguntas que están lejos de ser resueltas, a pesar de que se han propuesto varias alternativas. Al mismo tiempo, hay soluciones que necesitan ser desarrolladas con mayor profundidad para que puedan alcanzar el potencial que se espera de ellas. En ambas circunstancias, la elaboración de trabajos rigurosos y científicamente validados es de gran importancia, ya que contribuyen al desarrollo e integración de la energía solar renovable de manera real –en un sentido práctico- cuando son asimilados y aplicados por la industria solar. En este sentido, es la propia industria la que demanda este tipo de trabajos de gran valor que ayudan al crecimiento del sector. La elaboración del estudio presentado en esta tesis sigue decididamente esta orientación, aportando nuevas soluciones y profundizando en otras, tanto en los ámbitos de la evaluación del recurso solar como en la predicción de éste.

Respecto a la evaluación de la irradiancia solar, uno de los principales problemas actuales de la industria es la necesidad de disponer de la información más fiable posible sobre el recurso solar, para llevar a cabo la toma de decisiones en proyectos solares. En particular, las instituciones financieras exigen a los promotores que realicen un análisis objetivo de viabilidad de sus proyectos, con una alta fiabilidad del rendimiento esperado de las instalaciones basadas en la energía solar durante su vida útil. Estos análisis de rendimiento se llevan a cabo apoyados fundamentalmente en series temporales de datos de irradiancia solar. Estas series de tiempo deben ser capaces de caracterizar objetivamente el rendimiento a largo plazo de las plantas solares. Estos estudios de rendimiento se realizan en términos de la capacidad de producción estimada de la instalación, de acuerdo con su

comportamiento esperado bajo ciertos escenarios de estrés que suelen considerar períodos anuales de baja disponibilidad de energía solar. La definición de estos escenarios, junto con el parámetro de incertidumbre derivado de la evaluación de la energía solar, es de capital importancia, ya que determina aspectos esenciales de la viabiliad de los proyectos, como el retorno de la inversión y los costes financieros. Por lo tanto, la calidad de la información contenida en los conjuntos de datos utilizados y su posterior procesamiento para obtener dicha información valiosa son decisivos. Es precisamente la definición de estos escenarios lo que representa una de las dificultades actuales en el campo de la evaluación del recurso solar. Como se explicó en la sección anterior (*vide supra* Sección 1.2 *Revisión del estado del arte*), estos escenarios se establecen sobre la base de series temporales históricas a largo plazo de irradiancia solar superficial –y otras variables meteorológicas de interés- y se describen en términos del denominado año meteorológico típico (TMY) –o del concepto más reciente de año solar típico (TSY)-. El problema se basa en el hecho de que no existe un método unificado común para generar los TMYs que sea ampliamente aceptado por la comunidad científica. Además, la existencia de diferentes metodologías para la generación del TMY evidencia la falta de consenso científico sobre esta cuestión. Asimismo, los métodos actuales son escasos. De ellos, no todos permiten generar el TMY para cualquier escenario de probabilidad de disponibilidad del recurso solar (Cebecauer y Suri, 2015). Además, algunos de ellos están en discusión y se han detectado algunos problemas (Blanc, 2015). Finalmente, no todos son públicos, sino que forman parte del conocimiento privado de ciertas empresas que prestan este tipo de servicios de consultoría. Sin embargo, desde el punto de vista de los usuarios finales de la industria solar, la posibilidad de obtener dos resultados diferentes para la misma fuente de datos introduce un factor extra de incertidumbre que permanece incontrolado. A pesar de que el TMY es una herramienta que no cuenta con el favor de la comunidad científica, es tan utilizado en la industria solar que sigue siendo una herramienta clave para los proyectos solares. Por lo tanto, debido a su amplio uso práctico, realizar un esfuerzo para dar una respuesta científica unificada es por lo tanto ineludible. El objetivo es desarrollar un método estándar para la generación del TMY que pueda

generar confianza por tener un amplio consenso científico y contribuya a mejorar la bancabilidad de los proyectos.

Por otro lado, integrar una fuente variable como la energía solar en la estructura de suministro energético a gran escala compromete la operatividad del sistema, que debe funcionar bajo el principio de seguridad y estabilidad, mientras que la producción debe equilibrarse con la demanda esperada. Además, los productores deben conocer el recurso esperado que estará disponible para programar la operación y gestión de la planta, así como para planificar la mejor estrategia para participar en el mercado de la electricidad. En este sentido, es habitual aplicar una política de sanciones contra las desviaciones anunciadas de antemano por los productores a los operadores del sistema (TSO). Estas sanciones reducen los ingresos potenciales de la planta respecto al caso del 100% de la fiabilidad prevista, es decir, sin ninguna desviación. Además, conocer la irradiancia solar con cierta previsión ampliaría las posibilidades de mejorar la rentabilidad de la planta, por ejemplo, mediante su participación en mercados especiales mejor remunerados, como los llamados mercados de ajuste. Por lo tanto, la predicción de la radiación solar será una herramienta clave para la programación eficaz, mejorando decisivamente la competitividad de la energía solar. Por lo tanto, se puede concluir que el problema final no es tanto la variabilidad de la radiación solar, sino su predictibilidad (IEA, 2008). Por lo tanto, la industria solar exige un pronóstico especializado confiable del recurso solar para todos los períodos útiles en la fase de operación de la planta. Entre las metodologías para la predicción de la radiación solar, los modelos numéricos de predicción meteorológica (NWP) se destacan como la herramienta más potente. Proporcionan pronósticos completos sobre una base física para rangos cortos y medios –desde minutos hasta días- para localizaciones puntuales y regiones extendidas, con alta resolución espacial y temporal. En particular, el modelo regional del WRF es uno de los NWP más avanzados actualmente en el mundo. Está ampliamente apoyado y desarrollado por la comunidad científica, estando en el estado del arte de la predicción numérica. Sin embargo, los modelos NWP no han sido específicamente diseñados para la aplicación de la predicción de

irradiancia solar en el sector de la energía solar. Aún más, el conocimiento sobre la habilidad del modelo para predecir la irradiancia solar está lejos de ser ampliamente estudiado. Los principales problemas siguen siendo desconocidos. Es necesario entender el comportamiento específico del modelo al prever la irradiancia solar. En particular, la predicción de la componente DNI sigue estando poco investigado. Por lo tanto, resulta fundamental conocer el comportamiento del modelo bajo diferentes situaciones de nubosidad, estabilidad de las predicciones respecto al horizonte de predicción y cómo es la capacidad del modelo para predecir la GHI y la DNI. Además, es muy importante comparar los pronósticos con otros modelos, con el fin de establecer un marco de referencia que permita comprender no sólo el rendimiento del modelo en términos absolutos, sino también en los relativos. Por último, es esencial saber si alcanzar altas resoluciones, que a priori significan una mejor representación de los fenómenos físicos, proporciona mejores resultados al evaluar las predicciones de radiación solar (Stensrud, 2007). Esto proporciona una mejor visión sobre el rendimiento del modelo para predecir la irradiancia solar. En conjunto, la presente investigación proporciona una mejor comprensión de las capacidades del modelo WRF para proporcionar predicciones de irradiancia solar y su aplicabilidad a la industria solar con respecto a la integración de esta fuente de energía en las estructuras del suministro.

Para resumir de una manera general, el trabajo de investigación de esta tesis está motivado por la necesidad de responder a varias preguntas relevantes cuyas respuestas siguen siendo desconocidas o incompletas, pero que resultan de una importancia clave para la aplicación del conocimiento más avanzado al sector de la energía solar. En particular, su interés está en mejorar la competitividad de la energía solar a través de la mejora de la evaluación y predicción de los recursos solares en los aspectos específicos descritos anteriormente.

1.4 Objetivos

El propósito de esta tesis doctoral es desarrollar y evaluar métodos para la caracterización y estimación de la irradiancia solar en superficie para su aplicación práctica en el campo de la energía solar. El alcance de este objetivo aborda dos dimensiones de la radiación solar como fuente primaria de energía: su evaluación y su predicción. En esta tesis se tratan las dos componentes principales de la irradiancia solar para las tecnologías de conversión de energía solar, a saber, la irradiancia horizontal global (GHI) y la irradiancia directa normal (DNI) (esta última mucho menos investigada que el anterior).

Los objetivos específicos cubiertos en esta tesis se detallan a continuación:

1. Desarrollar y evaluar un método para la generación de años representativos de irradiancia solar para la caracterización del recurso solar a escalas plurianuales en cualquier lugar de interés.
2. Evaluar la fiabilidad de la predicción de irradiancia solar proporcionada por el modelo numérico de predicción meteorológica (NWP) regional de mesoescala Weather Research and Forecasting (WRF).
3. Evaluar las predicciones de irradiancia solar proporcionadas por el modelo WRF frente a otros modelos de referencia NWP.
4. Analizar el efecto de la resolución espacial horizontal del modelo WRF en la fiabilidad de las predicciones de irradiancia solar.
5. Estudiar el efecto de la agregación espacial de la irradiancia solar predicha en la fiabilidad de la predicción.

El primer objetivo tiene que ver con la evaluación del recurso solar. Está relacionado con el requerimiento de la industria solar de un método científicamente fundamentado que se convierta en un estándar para la generación de años solares típicos (TSY) enfocados particularmente en la descripción del recurso solar. Los TSY se utilizan

convencionalmente para analizar aspectos clave de la viabilidad financiera de los proyectos solares –normalmente dentro del ámbito de los estudios de bancabilidad-. El logro de este objetivo puede conducir a una herramienta común ampliamente utilizada para el cálculo de los TSY. Esto también permitiría unificar la parte de la incertidumbre total asociada con el propio método en los estudios de evaluación del recurso solar. Además, su carácter abierto y la propiedad de ser fácilmente implementado en un algoritmo podrían favorecer su adopción por la comunidad solar.

Los objetivos restantes se refieren a la predicción de la radiación solar, que es esencial para la integración de la energía solar en los sistemas de suministro de energía. En particular, el segundo objetivo analiza el uso del modelo NWP WRF para su aplicación en la predicción de la irradiancia solar superficial, con el fin de determinar su idoneidad como herramienta de soporte para la integración de la energía solar. El WRF es un modelo comunitario de mesoescala puntero, disponible públicamente y con una amplia comunidad de usuarios en todo el mundo. El tercer objetivo se refiere a la necesidad de evaluar la fiabilidad del WRF frente a otros modelos NWP. Esto permite determinar una escala de referencia en un marco comparativo y así conocer la habilidad del modelo en términos relativos. El cuarto y quinto objetivos pretenden investigar a fondo la influencia de la resolución espacial en el modelo WRF al estimar la irradiancia solar superficial. Esto está directamente relacionado con las conclusiones de los resultados obtenidos en los objetivos segundo y tercero. Es importante porque, precisamente, se supone que la capacidad de los modelos regionales como el WRF para alcanzar altas resoluciones espaciales es a priori una ventaja, ya que se considera que esta propiedad proporciona una mejor representación de los fenómenos físicos. Estos objetivos deben conducir a una mejor comprensión del uso del modelo WRF como una herramienta potente para la predicción de la irradiancia solar, lo que permitirá mejorar la competitividad de la energía solar. Además, deben servir como un trabajo de apoyo para futuras investigaciones en esta área.

Recapitulando, estos objetivos resumen una tesis cuyo propósito general está fuertemente orientado hacia las necesidades relacionadas con la industria solar. Con este fin, este trabajo de investigación intenta alinear los avances científicos con los intereses particulares de la industria solar desde el punto de vista de la fuente de energía primaria. Así, esta tesis contribuirá a mejorar el nivel de progreso e integración de las actuales tecnologías de explotación de la energía solar renovable.

1.5 Estructura de la tesis

Esta tesis se organiza de la siguiente manera. El Capítulo 1 presenta los antecedentes y el marco en el que se lleva a cabo la investigación. A continuación, se presenta el núcleo del trabajo de investigación. Está compuesto por un corpus de cuatro obras publicadas – capítulos 2 a 5-. La organización de los capítulos no sigue el orden cronológico de la investigación, sino que ha sido estructurado persiguiendo las preferencias acostumbradas en la comunidad solar, que siguen la cronología normal de desarrollo de los proyectos solares. De esta manera, el Capítulo 2, que se refiere a la evaluación de los recursos solares, se coloca antes de los capítulos dedicados al pronóstico de los recursos solares (capítulos 3 a 5). A continuación, se presenta una descripción resumida de estos capítulos. Finalmente, la tesis termina con una discusión de las principales conclusiones derivadas del trabajo de investigación, junto con una explicación de las futuras investigaciones propuestas para continuar las líneas de investigación aquí iniciadas.

Capítulo 2 - *Un nuevo procedimiento para generar TSY de irradiancia solar*

Los años solares típicos (TSY) son herramientas clave para la industria de la energía solar. En particular, los TSY se utilizan principalmente para el diseño y análisis de la bancabilidad de los proyectos de energía solar. En esencia, un TSY tiene la intención de describir el comportamiento esperado a largo plazo del recurso solar (irradiancia directa y / o global) en un período condensado de un año en la ubicación específica de interés. Un TSY difiere de un Año

Meteorológico Típico convencional (TMY) por su ausencia de variables meteorológicas distintas de la radiación solar. En cuanto a la probabilidad de excedencia (Pe) necesaria para la bancabilidad, se utilizan corrientemente varios escenarios, siendo Pe90, Pe95 o incluso Pe99 los escenarios más desfavorables, junto con el escenario medio más utilizado (Pe50). No existe un consenso en la comunidad científica con respecto a la metodología para generar los TSY para cualquier escenario de Pe. Además, la aplicación de dos métodos de construcción diferentes al mismo conjunto de datos original podría producir TSY diferentes. En este marco, la Asociación Española de Normalización y Certificación (AENOR) ha establecido un grupo de expertos para proponer un método que pueda ser estandarizado. El método desarrollado por este grupo de trabajo, denominado método EVA, se presenta en esta contribución. Su evaluación muestra que proporciona resultados razonables para los dos componentes principales de la irradiancia (directa y global), con errores bajos en las estimaciones anuales para cualquier Pe dada. El método EVA también preserva las estadísticas a largo plazo cuando el TSY calculado para un Pe específico se expanden a partir de la base mensual utilizada en la generación del TSY a resoluciones de tiempo más altas, como 1 hora, que son necesarias para la simulación precisa de la producción de energía de los sistemas solares.

Capítulo 3 - *Evaluación de la predicción de la irradiancia solar del modelo WRF en Andalucía (sur de España)*

En este trabajo se evalúa la fiabilidad de la predicción de la irradiancia horizontal global (GHI) y de la irradiancia directa normal (DNI) con tres días de anticipación, proporcionada por el modelo atmosférico de mesoescala WRF para Andalucía (sur de España). Las predicciones de la GHI fueron producidas directamente por el modelo, mientras que las predicciones de la DNI se obtuvieron sobre la base de un procedimiento de post-procesado físico utilizando las salidas del WRF y datos de satélite. Las estimaciones con una resolución horaria y una resolución espacial de 3 km se analizaron frente a las medidas de tierra recogidas en cuatro estaciones radiométricas a lo largo de los años 2007 y 2008. La evaluación se llevó a cabo independientemente

para diferentes horizontes de previsión (1, 2 y 3 días en adelante), las diferentes estaciones del año y tres condiciones diferentes del cielo: claro, nuboso y completamente cubierto. Los resultados mostraron que el modelo WRF presenta una habilidad importante en la predicción de la GHI y la DNI, en general, mejor que el trivial modelo de persistencia. Sin embargo, los valores del MBE y el RMSE presentaron una marcada dependencia con las condiciones de cielo y con la estación del año. En particular, durante las primeras 24 h de predicción, el MBE de la GHI predicha fue del 2% para cielos despejados y del 18% para condiciones nubladas. Sin embargo, el MBE de la DNI predicha aumentó hasta aproximadamente el 10% y el 75% para condiciones claras y nubladas, respectivamente. Con respecto a los valores RMSE, en el caso de la GHI predicha, los resultados oscilaron entre menos del 10% bajo cielos claros y 50% para condiciones nubladas. En el caso de la DNI predicha, el RMSE varió de 20% a 100% para cielos despejados y nublados, respectivamente. Esto demostró la mayor sensibilidad de la DNI a las condiciones del cielo. En general, se observó un incremento de los valores del MBE y el RMSE con la nubosidad. Esto refleja una habilidad todavía limitada del modelo WRF para predecir adecuadamente condiciones nubladas en comparación con cielos despejados. Sin embargo, el modelo fue capaz de predecir con exactitud los cambios abruptos en las condiciones del cielo (nubosidad). Por último, el rendimiento del WRF fue considerablemente mejor que el modelo de persistencia para cielos despejados tanto para la GHI como para la DNI, con valores relativos RMSE alrededor de la mitad. Sin embargo, para condiciones nubladas, el rendimiento de ambos fue similar.

Capítulo 4 - *Comparación de predicciones de irradiancia solar de modelos numéricos de predicción meteorológica en EEUU, Canadá y Europa*

Este artículo combina y discute tres validaciones independientes de la irradiancia horizontal global (GHI) de los modelos de predicción a varios días que se llevaron a cabo en los EEUU, Canadá y Europa. Todos los modelos de predicción se basan directa o indirectamente en modelos numéricos de predicción meteorológica (NWP). Dos modelos

son comunes a los tres esfuerzos de validación –el modelo global del ECMWF y el modelo de mesoescala WRF con condiciones iniciales del GFS- y permiten las siguientes observaciones generales: (1) las predicciones del modelo WRF alimentado con GFS no funcionan tan bien como las predicciones basadas en modelos globales, tales como el ECMWF y (2) el promedio simple de la predicción de los modelos tiende a funcionar mejor que los modelos individuales.

Capítulo 5 - *Evaluación de la predicción de la DNI basada en el modelo atmosférico de mesoescala WRF para aplicaciones en CPV*

La integración de la producción de la energía solar a gran escala en las estructuras de suministro de energía depende esencialmente del conocimiento previo preciso del recurso disponible. Los modelos numéricos de predicción meteorológica (NWP) proporcionan una herramienta fiable y completa para las predicciones de radiación solar de corto y medio alcance. La metodología seguida aquí se basa en el modelo WRF. Para los sistemas CPV, la fuente de energía primaria es la irradiancia normal directa (DNI), que se ve dramáticamente afectada por la presencia de nubes. Por lo tanto, la fiabilidad de las predicciones de DNI está directamente relacionada con la exactitud de la información de la nube. Dos aspectos de esta cuestión se discuten aquí: (i) el efecto de la resolución espacial horizontal del modelo; y (ii) el efecto de la agregación espacial de la irradiancia predicha. Los resultados muestran que no hay mejoría en la habilidad del pronóstico de la DNI en altas resoluciones espaciales, excepto en condiciones de cielo despejado. Además, el promedio espacial de la irradiancia predicha reduce notablemente su error inicial.

Annex E

Capítulo 6. Resumen y conclusiones

6.1 Introducción

La investigación realizada en este trabajo tiene por objetivo la mejora del conocimiento actual sobre el recurso solar para el avance y desarrollo de aplicaciones en la industria solar. En particular, se llevan a cabo dos tipos investigaciones. En primer lugar, el foco se centra en mejorar la evaluación del recurso solar para los estudios de viabilidad en proyectos de centrales de energía solar. Con este fin, en el Capítulo 2 se presenta y evalúa un nuevo método para obtener un año representativo para la caracterización de la irradiación solar a escalas multianuales. En segundo lugar, la investigación se centra en los problemas de integración de la energía solar, una vez que la planta de energía solar está en operación. Para ello, es obligatorio el desarrollo de sistemas precisos de predicción de la radiación solar. En el Capítulo 3, se evalúa el rendimiento del modelo WRF para la predicción de la irradiancia solar. A continuación, el Capítulo 4 presenta una comparación de las predicciones de radiación solar obtenidas con WRF y con otros modelos numéricos de predicción meteorológica (NWP). Finalmente, a partir de los resultados de los estudios anteriores, se presenta un análisis de la influencia de la resolución espacial horizontal en la habilidad del modelo para la predicción de la radiación solar (Capítulo 5).

6.2 Generación de TSY para una evaluación mejorada del recurso solar

La evaluación del recurso solar puede considerarse como el corpus de métodos, técnicas y sus aplicaciones para la caracterización a largo

plazo de la radiación solar en un lugar de interés. Un ingrediente clave de la evaluación del recurso para la industria de la energía solar es el llamado Año Meteorológico Típico (TMY). A pesar de que los fundamentos teóricos del TMY no son plenamente aceptados por la comunidad científica especializada, son ampliamente utilizados y requeridos por la industria solar, principalmente para el diseño y análisis de bancabilidad de los proyectos solares. En particular, el uso de TMYs se ha convertido en un estándar para tipificar la producción de energía esperada de las plantas solares. Sin embargo, los métodos para obtener TMY son variados, debido fundamentalmente a la falta de consenso científico. Esto causa una inconsistencia no deseada, ya que la aplicación de diferentes metodologías al mismo conjunto de datos puede resultar en diferentes TMY. En última instancia, esto aumenta la incertidumbre en la evaluación del recurso solar, que es justo lo contrario de lo que se pretende. Al aumentar la fiabilidad del recurso solar esperado, se pueden reducir los recargos asociados al riesgo y, por lo tanto, se reducen los costes efectivos de producción. En este sentido, la estandarización de los métodos para derivar TMY es una demanda de la industria solar, ya que establece una referencia para la calidad de los datos y ayuda a reducir la incertidumbre.

En este contexto, el Capítulo 2 presenta un nuevo método destinado a contribuir a la estandarización de la generación del Año Solar Típico (TSY). El TSY, al igual que el más común TMY, es una serie temporal anual formada artificialmente para caracterizar el recurso solar de un lugar determinado. A pesar de que ambos -TMY y TSY- pueden proporcionar la misma cantidad de información (véase la discusión más adelante) en el Capítulo 2 el TSY se considera de acuerdo con su definición más sustancial, es decir, como una serie temporal de una sola variable de irradiancia solar. El objetivo principal del método, denominado método EVA –un acrónimo de las palabras estacionalidad y variabilidad-, es determinar los valores mensuales acumulados de energía solar cuya suma anual es igual al valor objetivo anual para una probabilidad determinada de excedencia. Así, es posible seleccionar los meses correspondientes de la serie temporal a largo plazo más cercanos a los meses estimados. El método EVA proporciona la selección de los

meses basándose en una composición de pesos teniendo en cuenta: i) la variabilidad de las distribuciones mensuales desestacionalizadas, ii) la contribución energética mensual individual de cada mes del calendario respecto a la energía total anual. La novedad de este método se basa, en primer lugar, en su base estadística –a través de un enfoque original- y, en segundo lugar, en su sólida formulación analítica. Esto proporciona robustez y fiabilidad al método, al tiempo que facilita su implementación algorítmica (*vide supra* Annex A). Además, el método EVA tiene la propiedad de ser válido para el análisis de cualquier componente de la radiación solar indistintamente, GHI o DNI dependiendo de la tecnología solar. Sin embargo, su característica más destacada es su capacidad para generar TSY para cualquier escenario esperado de disponibilidad del recurso solar. Esto incluye los casos más desfavorables de energía extremadamente baja, cuya importancia es clave para la bancabilidad de los proyectos solares. Actualmente, hay pocos métodos que permiten este tipo de análisis, y algunos de ellos no son públicos. Por último, el método parece ser consistente respecto a la determinación de la incertidumbre. En particular, esta incertidumbre puede ser fácilmente cuantificada y combinada con el valor de incertidumbre total de la evaluación del recurso solar por medio de la ley de Gauss de propagación de errores. Sin embargo, se necesitan investigaciones adicionales para analizar más a fondo esta cuestión. La determinación de la incertidumbre es fundamental para la fiabilidad de los proyectos solares y, por tanto, para su bancabilidad.

En el Capítulo 2 también se presentan los principales resultados de la evaluación del método EVA. En general, se muestra que el método EVA funciona adecuadamente. Las diferencias entre el valor objetivo anual para una cierta probabilidad de excedencia y el valor anual correspondiente de la irradiancia generada por el TSY obtenido con el método EVA son bajas. Las diferencias entre los valores de irradiación anual de los TSY y los valores percentiles anuales objetivo son generalmente inferiores al 1%, tanto para la GHI como para la DNI. Además, el método conserva las estadísticas a largo plazo cuando el TSY construido para una probabilidad de excedencia determinada se

recompone a la resolución original más alta de sus meses constituyentes -típicamente 1 hora.

Sin embargo, se deben realizar investigaciones adicionales para ampliar el conjunto de ubicaciones utilizadas para evaluar el método. Además, también sería importante analizar la aplicación del método EVA a datos satelitales. Esto se debe principalmente a que, en la práctica, la información por satélite es la fuente más común de series temporales de largo plazo de datos de irradiancia para aplicaciones en la industria solar. En este sentido, es oportuno comparar los resultados de la aplicación del método utilizando series temporales a largo plazo de datos medidos y los derivados de satélites.

Además, sería muy importante extender esta investigación analizando la producción de energía de una planta solar en función de diferentes escenarios esperados de recurso solar, como los descritos por los TSY derivados del método EVA. Esto permitiría estimar la respuesta típica de la planta bajo las condiciones establecidas con los TSY. En cualquier caso, se puede concluir que el objetivo particular propuesto en esta tesis sobre la evaluación de los recursos solares se logró con éxito.

Por último, debe aclararse que el método EVA no proporciona el TMY habitual, sino el llamado TSY. Es importante señalar que el TSY no tiene menos información que el TMY, sino ligeramente diferente. En este sentido, el TSY puede complementarse directamente con las variables meteorológicas requeridas simplemente añadiendo sus valores a los registros de tiempo correspondientes. La diferencia es que la generación tradicional de los TMY hace uso de información meteorológica auxiliar mediante pesos preestablecidos no consensuados, mientras que el TSY solo usa la variable principal, es decir, da el 100% del peso a la variable principal de irradiancia -GHI o DNI-. En cierto sentido, el uso del adjetivo "meteorológico" en TMY puede resultar un poco ambicioso. Principalmente, porque no existe una evidencia clara sobre la conveniencia de combinar -con determinado grado de influencia- las variables secundarias con las principales variables de irradiancia -GHI y/o DNI-. En consecuencia, no hay

consenso sobre las variables externas requeridas y el peso que cada una debe tener. Por lo tanto, para aplicaciones de energía solar parece aconsejable adoptar un método que considere solamente las variables de irradiación fundamentales. Esto reduciría los grados de libertad en la determinación de la incertidumbre, como por ejemplo aquellos asociados a la determinación de las variables meteorológicas. En este sentido, a veces esto puede conducir a problemas en la aplicación del concepto TMY en la industria solar. A pesar de que los TSY no se recomiendan para el análisis de la producción, realmente se utilizan para inferir el rendimiento esperado de la planta de acuerdo con los escenarios definidos por los TSY. De esta manera, en situaciones prácticas el interés de un promotor de plantas solares es presentar el proyecto más atractivo para ganar un proceso de licitación. Con este fin, el rendimiento energético esperado es determinante y, por tanto, la cantidad estimada de energía. Puesto que las tablas de pesos a utilizar para la combinación de las variables para generar los TSY forman parte del conocimiento de la empresa, pueden ser ligeramente modificadas para favorecer un resultado más conveniente para los intereses del promotor. Por ejemplo, añadiendo un pequeño incremento al porcentaje del peso asociado a la temperatura, o a cualquier otra variable; Lógicamente con mucho cuidado de no superar las expectativas excesivamente. Al final, se puede obtener un conjunto de diferentes TSY para el mismo conjunto de datos a largo plazo, y así el promotor puede elegir el más conveniente. En este sentido, el uso del TSY en lugar del TMY permite eludir estas situaciones. Además, la estandarización de los métodos también ayudará a evitar estas imprecisiones innecesarias, proporcionando una solución única para su uso común.

6.3 Rendimiento de la predicción de radiación solar con WRF para la integración de la energía solar

Dada la naturaleza caótica del clima, la radiación solar es muy variable en el espacio y en el tiempo, y una predicción confiable de la

producción de energía solar es de suma importancia para la integración a gran escala de mayores proporciones de energía solar en los sistemas de generación y distribución de energía. La predicción de la energía solar contribuye a minimizar los riesgos del sistema de suministro asociados a las fluctuaciones de la producción de energía solar y a maximizar los ingresos económicos programando el suministro de energía de acuerdo con la producción esperada y la situación del mercado. En este contexto, hoy en día, la previsión de la producción de energía solar es una herramienta básica de los operadores de redes de transporte (TSO) para la gestión de la red. También es fundamental para los operadores de centrales solares para el comercio de energía y el funcionamiento y la gestión de la planta. Entre las metodologías para predecir la radiación solar superficial, los modelos NWP se destacan como la herramienta más poderosa para pronosticar horizontes más allá de unas 5 horas. Su base física les permite proporcionar pronósticos meteorológicos integrales, manteniendo la coherencia espacial y temporal sobre regiones extensas y períodos de tiempo de corto a mediano plazo. En particular, estos modelos pueden proporcionar predicción de irradiancia solar, así como de variables auxiliares de interés para las aplicaciones en energía solar. Las escalas espacial y temporal de los pronósticos proporcionados por estos modelos están en el rango de 10-25 km y 1-3 horas, en el caso de los modelos de circulación global, y sólo unos pocos kilómetros e infra-horarias, en el caso de los modelos regionales. Dentro de este último tipo de modelos NWP, el modelo WRF destaca como uno de los más avanzados. Cuenta con un amplio conjunto de parametrizaciones que proporcionan al modelo una gran flexibilidad para adaptarse a una tarea específica y a las características geofísicas de una región particular. También es uno de los modelos más utilizados en todo el mundo, teniendo en cuenta que se trata de un modelo comunitario en continua evolución y desarrollo. Por lo tanto, es una elección adecuada para los propósitos de esta investigación.

En general, los modelos de NWP no están específicamente diseñados para aplicaciones en energía solar. Como consecuencia, todavía hay pocos trabajos de investigación, como el que aquí se

presenta, que evalúen el rendimiento de este tipo de modelos en la tarea específica de predicción de la radiación solar.

El punto de partida de esta investigación se aborda en el Capítulo 3. En este estudio inicial se realiza una evaluación exhaustiva de la fiabilidad de las predicciones de GHI y DNI obtenidas con el modelo WRF. El análisis se realiza comparando los pronósticos del modelo con las mediciones en tierra de varias estaciones radiométricas ubicadas en la región de Andalucía (sur de España). El período de análisis es de un año. La resolución temporal es la base horaria habitual, mientras que la resolución espacial es de 3 km. Esta alta resolución espacial es notablemente diferente en comparación con las resoluciones espaciales más gruesas logradas por los GCM -normalmente mayores de 10 km-. El estudio incluye diferentes aspectos que afectan a la predicción de la radiación solar (nubes, aerosoles, etc.). Además, no se aplicó pos-procesado con el fin de enfocar el análisis sólo en el rendimiento del modelo. En particular, se evaluó la calidad de la predicción para diferentes condiciones del cielo, a saber: completamente nublado, nuboso, cielo despejado y todas las condiciones de cielo. Se realizaron análisis independientes a nivel estacional y anual. Además, se evaluó la habilidad del modelo para diferentes horizontes de predicción (24h, 48h y 72h). Por otra parte, a diferencia de la GHI, que es proporcionada directamente por el modelo, la DNI se derivó de los resultados del modelo ya que no era una variable de salida incluida en la versión de WRF utilizada en este estudio, versión 3.2.1. Las versiones actuales de WRF ya proporcionan la DNI y la componente difusa. No obstante, este fue el primer estudio que analizó la predicción de DNI derivada de WRF en el contexto de las aplicaciones en el ámbito de la energía solar. Sin embargo, todavía no se ha estudiado adecuadamente la fiabilidad relativa de las predicciones de DNI de los modelos NWP respecto a los pronósticos de DNI obtenidos mediante un modelo externo, como los modelos de separación para todas las condiciones de cielo o los modelos de transferencia radiativa como el utilizado aquí. Desde un punto de vista práctico, esto podría ser importante, tanto para usar los pronósticos más fiables, como para tener una valiosa referencia para la

evaluación de predicciones de DNI. No obstante, el objetivo debe ser mejorar el rendimiento del modelo.

Como resultado de la evaluación se encontró que, en general, el modelo WRF tiende a sobrestimar la irradiación solar, para todas las condiciones del cielo y para todos los períodos analizados. Para condiciones de nubosidad, se encontró que esto es debido principalmente al hecho de que el modelo WRF subestimaba la cantidad de nubes. En el caso de condiciones de cielo despejado, este resultado sugiere que los valores de AOD utilizados en el modelo subestiman los valores reales. Como era de esperar, los errores para DNI eran marcadamente superiores a los obtenidos para la GHI. Esto se debe a la mayor sensibilidad de la DNI, en general a la presencia de nubes, pero también a la incertidumbre de la carga de aerosol. Los resultados también mostraron que el modelo en general funciona mejor que el modelo trivial de persistencia, excepto en los períodos donde la presencia de nubes es más significativa, donde el rendimiento fue similar. De esta manera, el verano presenta mejores resultados, mostrando que la presencia de nubes es, con diferencia, el factor más importante en la estimación de la irradiancia solar. En resumen, se llegó a la conclusión de que la predicción de nubes sigue siendo un gran problema con respecto a la predicción de radiación solar. Por lo tanto, se deberían obtener mejoras sustanciales en la representación de las nubes en los modelos NWP en el futuro. Esto ayudará a alcanzar los estrictos requisitos de la industria solar.

El siguiente paso en el análisis dela predicción de la radiación solar fue la intercomparación del método propuesto aquí, basado en el modelo WRF, frente a otros métodos basados en diferentes modelos meteorológicos regionales y globales. Este análisis se realizó en un extenso ejercicio de evaluación comparativa descrito en el Capítulo 4. Este trabajo estableció un marco de referencia que permitió evaluar las previsiones de radiación solar del WRF con respecto a las previsiones de otros modelos NWP. El estudio se realizó sólo para la GHI debido principalemente a la falta de observaciones de DNI. Cabe señalar que, a diferencia del enfoque seguido aquí, algunos modelos que participaron

en este ejercicio utilizaron pos-procesados para mejorar sus resultados. En este sentido, este estudio de *benchmarking* también es interesante para conocer las posibilidades de mejorar las estimaciones del modelo WRF. De hecho, los resultados muestran que el pos-procesado tiene un gran potencial para mejorar las previsiones con grandes desviaciones sistemáticas, como es el caso de WRF según los resultados anteriores obtenidos en el Capítulo 3. En general, se concluyó que el rendimiento de los modelos globales (GCM) fue mejor que el de los modelos regionales. En general, los errores del modelo WRF fueron significativamente mayores que los errores del modelo ECMWF, que obtuvo los mejores resultados. Se llegó a la conclusión de que una de las razones eran los bajos errores de sistemáticos proporcionados por los GCM; a diferencia de WRF, que mostró una sobreestimación significativa. Otro resultado interesante del estudio de *benchmarking* sugirió que las diferencias en la predicción entre los modelos regionales y los GCM tienen más que ver con los propios modelos regionales que con las condiciones iniciales y de contorno utilizadas -proporcionadas en este estudio por el modelo GFS-. A este respecto, se requieren estudios adicionales detallados para confirmar esta afirmación. Sin embargo, contrariamente a lo esperado, se encontró que la resolución espacial horizontal no desempeñó un papel importante en la fiabilidad de las predicciones de GHI. Esta cuestión motivó una investigación adicional, que se presenta en el Capítulo 5.

Siguiendo los resultados obtenidos en el Capítulo 4, el siguiente paso en el plan de investigación fue analizar el papel de la resolución espacial horizontal. Este es un aspecto importante para la predicción del recurso solar desde el punto de vista de la optimización de la salida del modelo. Además, se evaluó también el efecto del promediado espacial de la radiación solar en la fiabilidad de la predicción de la irradiación solar de WRF. Este método simple de post-procesado es ampliamente utilizado, ya que generalmente reduce el error absoluto, que está directamente relacionado con las desviaciones de la producción de energía esperada. Ambos aspectos fueron analizados en el Capítulo 5. Este estudio se realizó para ambas componentes: GHI y DNI, aunque

sólo se han mostrado los resultados para la DNI. Sin embargo, las conclusiones principales son válidas para ambas variables.

Por un lado, se examinó el papel de la resolución espacial horizontal para aclarar su efecto sobre la fiabilidad de la predicción de irradiancia solar. Contrariamente a lo que cabría esperar, los resultados mostraron que un incremento en la resolución espacial no aumenta necesariamente la fiabilidad de las predicciones de DNI basadas en WRF. Sólo en condiciones de cielo despejado el rendimiento es mejor en el caso de mayores resoluciones espaciales, cuando las características topográficas son más relevantes. Esto sugiere un papel importante de las nubes en esta tendencia. Por otro lado, en la segunda parte del Capítulo 5, se aplica un método de post-procesado sobre las salidas del WRF. Consistió en un filtro de alisado basado en un promediado espacial -o agregación espacial- de la irradiancia solar. Los resultados mostraron claramente que el promedio espacial de la irradiación solar reduce notablemente los errores de predicción, mejorando la habilidad del WRF. Además, los resultados fueron más fiables para resoluciones más gruesas y para una agregación espacial que abarca alrededor de 100 × 100 km. Se han obtenido resultados similares en otros estudios utilizando otros modelos de NWP.

A partir de los resultados de los capítulos 4 y 5, se concluyó claramente que una mayor resolución espacial no garantiza un mejor rendimiento del modelo con respecto a la predicción de la irradiancia solar. Esto no significa que la representación de la nubosidad que hace el modelo sea mejor en resoluciones espaciales más gruesas. Hay otros aspectos que hay que tener en cuenta, como la propia representación de las nubes por parte del modelo, o incluso el propio proceso de evaluación, ya que los valores de irradiancia solar modelada representan el promedio sobre toda la extensión de la celda del grid, las cuales son comparadas con observaciones puntuales. Otro factor importante a considerar es el denominado error de doble penalización, que está motivado por el hecho de que las nubes tienen que estar representadas correctamente tanto en el espacio como en el tiempo para no producir penalizaciones en el error. Esto significa que las nubes

deben estar en la posición exacta en el momento preciso con respecto a las observaciones recogidas en la estación radiométrica. Esto está relacionado con la efectividad del promediado espacial para reducir la componente aleatoria del error –MAE y RMSE- ya que el promedio espacial reduce el impacto de los errores de doble penalización. Esto es así porque, de alguna manera, es como si el aumento del tamaño de las nubes evitara muchos errores en los que la nube se genera cerca del punto de validación pero no en la posición exacta. Alternativamente, en general, la bondad del promediado espacial en la reducción de errores aleatorios está estrechamente relacionada con la autocorrelación espacial de la radiación solar en las proximidades del sitio de evaluación, de tal manera que a menor autocorrelación espacial, menor es el tamaño promedio que se requiere para lograr la misma reducción de los errores aleatorios. Además, existe un tamaño promedio para el cual esta reducción es óptima (es decir, el máximo posible) (en nuestros experimentos es aproximadamente 100x100 km) y también depende de la estructura de autocorrelación espacial. Por lo tanto, la distancia óptima varía con las condiciones meteorológicas sinópticas y la configuración topográfica del sitio de validación -entre otras cosas- que son los factores que determinan la estructura de autocorrelación espacial de la radiación solar. Por último, es necesario mencionar que, aunque el promedio espacial de la radiación solar puede reducir los errores aleatorios de las estimaciones del modelo, también puede distorsionar la distribución de probabilidad de la radiación solar simulada por el modelo. Esto puede ser un factor limitante de este pos-procesado para aquellas aplicaciones en las que una buena representación de la distribución de datos a largo plazo sea crítica. Por lo tanto, se puede concluir que la conveniencia de lograr altas resoluciones espaciales -con el enorme esfuerzo computacional requerido- y/o el uso de un post-procesado basado en el promediado espacial dependen de la aplicación final. Por ejemplo, sobre la base de estos resultados y los del Capítulo 4, se puede concluir que el uso del modelo WRF se recomienda para el funcionamiento de la planta, en lugar de usar un modelo global. Para estas aplicaciones, resoluciones temporales más altas son más adecuadas. Además, la variabilidad de los datos es importante para el modelado de la producción de energía. Sin

embargo, para la venta de la energía, donde la resolución termporal suele ser de 1 hora, parece más conveniente usar un modelo global, como el modelo IFS del ECMWF. Esto se debe a que las desviaciones -y las sanciones consiguientes- están vinculadas a los errores de predicción, particularmente el MAE. Sin embargo, el potencial del modelo WRF es mucho mayor que cualquiera de los GCM en el sentido de tener mucha más flexibilidad. A este respecto, cabe mencionar los prometedores resultados obtenidos recientemente por algunos autores con la implementación de una versión especial del WRF dedicada a aplicaciones de energía solar.

En resumen, se han alcanzado los objetivos perseguidos para esta parte de la tesis. Se ha comprobado que el modelo WRF es capaz de proporcionar predicciones de radiación solar que pueden ser empleadas en la industria solar. Sin embargo, también se concluyó que el modelo tiene un problema en la representación de la fracción de nubes y/o de la cantidad de nubes. En la práctica, esto puede ser parcialmente modificado por la aplicación de métodos de pos-procesado a los resultados iniciales del modelo, que en general reducen o eliminan el error sistemático. Sin embargo, desde el punto de vista de la investigación científica, el siguiente paso debe ser aumentar las capacidades de WRF para mejorar la representación de las nubes y los aerosoles. En última instancia, hay que hacer un importante esfuerzo para mejorar el inmenso potencial de esta extraordinaria herramienta que es el modelo WRF.

6.4 Trabajo de investigación futuro

Durante la realización de esta tesis, se han abierto varias preguntas interesantes que merecen una investigación más profunda. Algunas de estas investigaciones se encuentran ya actualmente en curso.

Por un lado, en lo que respecta a los TSY para mejorar la evaluación del recurso solar, existe una cuestión abierta relacionada con el análisis detallado y la descripción de la incertidumbre derivada del método EVA. Adicionalmente, sería importante llevar a cabo un estudio de

caso práctico en el que se analizaría el rendimiento de una planta solar en relación con diferentes escenarios esperados descritos por los TSY. Ambos temas están actualmente bajo investigación. No obstante, sería conveniente ampliar el análisis presentado en esta tesis a nuevas ubicaciones y, en general, incluir datos satelitales, ya que estas fuentes son las más utilizadas en la industria solar. Estos objetivos se pueden combinar en un trabajo integral que complete el trabajo de investigación básico desarrollado en esta tesis.

También en lo que respecta a la evaluación de los recursos solares, merece una mayor investigación el papel del cambio climático en el rendimiento energético de las plantas solares durante su vida útil. El objetivo es incluir sus efectos en los estudios de evaluación del recurso solar. Hay escasas publicaciones sobre esta cuestión, la mayoría de ellas centradas en la energía eólica -habitualmente un paso por delante de la energía solar-. Los efectos del cambio climático en el recurso solar pueden no ser insignificantes durante la vida estimada de los proyectos solares. Por lo tanto, estos efectos deberían comenzar a ser considerados en la estandarización de los estudios de evaluación del recurso solar, o al menos formar parte del catálogo de buenas prácticas.

Por otro lado, con respecto a la predicción del recurso solar, se está llevando a cabo un benchmarking integral de las versiones más recientes de algunos modelos NWP. En un primer paso, sólo se evaluarán los GCM más populares y accesibles utilizando datos de 80 estaciones radiométricas de todo el mundo. Este análisis seguirá las pautas establecidas en trabajos de benchmarking anteriores, mientras que se introducirán nuevos elementos para examinar, por ejemplo, el rendimiento de la agregación espacial. Este trabajo está diseñado para ser una referencia para comparar el rendimiento de otros modelos. En un segundo paso, la última versión del modelo WRF se evaluará en un conjunto seleccionado de ubicaciones de acuerdo con los resultados obtenidos en el trabajo anterior con los GCM. También se evaluará de nuevo el papel de la resolución espacial horizontal, con el fin de conocer el rendimiento de las nuevas parametrizaciones implementadas en la nueva versión del modelo WRF. Para completar este trabajo, el

estudio examinará el funcionamiento del modelo respecto a las condiciones iniciales y de contorno, siguiendo las conclusiones del Capítulo 4.

Por último, sería muy interesante analizar el potencial de las técnicas de *machine learning* para aplicaciones en radiación solar. Existen varios estudios sobre este tema, particularmente para mejorar la fiabilidad de la predicción de los modelos mediante un pos-proceso. En un trabajo preliminar, el método llamado *gradient boosting regressor* se aplicó para mejorar la estimación de irradiancia solar estimada con satélite. Esta técnica permitió ampliar la cantidad de información con series temporales a largo plazo obtenidas de los reanálisis de modelos, como MERRA2, MERRAero y CFSR. Los resultados fueron prometedores, mostrando el extraordinario potencial de estos métodos de caja-negra. En el caso de la predicción de la irradiancia solar, los resultados preliminares muestran que la mejora obtenida por medio de este método es también notable.