

# **Características espacio-temporales de la producción fotovoltaica en el área Euro-Mediterránea: análisis climático**

Claudia Gutiérrez Escribano





---

## Preface

Applied science has made societies evolved through the link between the basic science and the 'ingenios' that implement the structural knowledge to the individual daily life. The bridge between the fundamental studies and its evolution or transformation into something applied (useful) for the societies development is been a matter of interpretation or 'translation' (so interpreters/translator) that have been able to find the accurate language to exchange this knowledge.

Transversality between the fundamental disciplines of physics, chemistry or biology and their different applied branches, most of them under the umbrella of engineering, end unavoidably into a transformation from the abstract ion to the tangible with the premise or ideal of improving wellbeing of present and future societies.

However, it is very likely that this conception of the scientific application had lead to the most challenging problem that humanity has to face from its existence, climate change, being this anthropocentrism part of the problem and not of the solution.

In his book "El pensador intruso", Jorge Wasengber wrote about the evolution of science and the underlying progress what follows:

"There are above all two vices that tend to set 'pre-cooked' ideology into science. First is based on different kinds of anthropocentrism and consist in putting the knowledge subject into the cosmos's center. The history of knowledge is the witness: each time we 'barremos' the 'I' from the spot, the knowledge progresses and only because of it.'



---

# Contents

<b>Contents</b>	<b>ii</b>
<b>I Introduction</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 A changing world . . . . .	3
1.2 Renewable Energies . . . . .	3
1.3 Links between climate and renewable energy . . . . .	5
1.4 Photovoltaic Energy . . . . .	5
1.5 Climate change and the Mediterranean area . . . . .	5
1.6 General scientific question: spatiotemporal behaviour of solar resource and photovoltaic production . . . . .	5
1.7 Organization of the manuscript . . . . .	5
<b>2 State of knowledge</b>	<b>7</b>
2.1 Intermittency of PV . . . . .	7
2.2 Variability sources in PV . . . . .	7
2.3 From short to long term issues . . . . .	7
2.4 Climate Change perspectives for solar resource and photovoltaic po- tential . . . . .	7

<b>II Data &amp; Methods</b>	<b>9</b>
<b>3 Data</b>	<b>11</b>
3.1 Solar radiation measurements: <i>Baseline Surface Radiation Network, BSRN</i> . . . . .	11
3.2 Modelling of solar radiation . . . . .	12
3.3 Satellite data: <i>Climate Monitoring Satellite Application Facility, CM-SAF</i> . . . . .	14
3.4 Photovoltaic production data . . . . .	15
3.5 Temperature data . . . . .	16
<b>4 Methods</b>	<b>19</b>
4.1 Clustering algorithm applied to climate data . . . . .	20
4.2 Simulating a photovoltaic system . . . . .	24
4.3 Using Regional Climate Models . . . . .	27
4.4 Future projections and scenarios . . . . .	30
<b>III Results</b>	<b>33</b>
<b>5 Long-term characteristics of solar resource over the Iberian Peninsula</b>	<b>35</b>
5.1 Introduction . . . . .	40
<b>6 Impact of aerosols in photovoltaic energy production over the Euro-Mediterranean area</b>	<b>55</b>
6.1 Introduction . . . . .	58
6.2 Data and Methods . . . . .	59
6.3 Results . . . . .	66
6.4 Discussion . . . . .	77
6.5 Conclusions . . . . .	79
<b>7 Future projections of solar resource for photovoltaic applications</b>	<b>81</b>
7.1 Introduction . . . . .	84
7.2 Climate data . . . . .	85
7.3 Methods . . . . .	86
7.4 Results . . . . .	87
7.5 Discussion . . . . .	90
7.6 Conclusion . . . . .	90
<b>Bibliography</b>	<b>93</b>



**Part I**

**Introduction**



---

# Introduction

*"Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."*

— Lewis Carroll, *Alice in Wonderland*

## 1.1 A changing world

Ideas a desarrollar:

\* Un mundo en constante evolución. \* Cambio climático \* Transición energética global \* Cambios sociales/políticos. \* Migraciones

\* Entrelazamiento de todas estas cuestiones. \* El conocimiento puro + el conocimiento aplicado: obligación de la ciencia a ser en la medida de lo posible aplicada a solucionar problemas de las distintas sociedades. Justicia social. Más si estos problemas han sido creados por el ser humano.

## 1.2 Renewable Energies

Podemos definir el concepto físico de energía como la capacidad de un sistema para realizar un trabajo, clasificando las distintas formas de energía en dos grandes grupos: la energía cinética, relacionada con el movimiento del sistema, o la energía potencial, que atiende a la posición de mismo.

En esta clasificación fundamental encontramos que la energía mecánica, la energía electromagnética o la energía térmica son formas de energía que se engloban

## 1. INTRODUCTION

---

dentro de la energía cinética, mientras que la energía química (de la cual es un subtipo la energía nuclear) o la energía potencial gravitacional (como es el caso de las centrales hidroeléctricas) se encontrarían dentro de la energía potencial.

Es elemental para nuestro estudio delimitar la diferencia entre formas y fuentes de energía. Estas últimas son definidas como aquellas a partir de las cuales la energía 'útil', aquella que es directamente aplicable para su finalidad sea esta el movimiento, la producción de electricidad o los distintos procesos metabólicos en el ser humano, puede ser extraída directamente o mediante algún proceso de transformación. Así, los combustibles fósiles son una fuente de energía, puesto que su energía 'útil' debe ser extraída. Es decir, extraemos la energía química contenida en los combustibles fósiles y necesitamos ciertos procesos de transformación entre la energía potencial de una masa de agua elevada que termina pasando a través de una turbina para obtener electricidad. Algunos ejemplos de fuentes de energía son los combustibles fósiles, la energía solar o la biomasa.

Aparece de manera natural la clasificación de estas fuentes de energía según su origen renovable, como la energía solar, eólica o geotérmica y no renovable como los combustibles fósiles. La energía renovable, es por lo tanto, aquella que tiene su origen en estas fuentes inagotables de energía. El sol, a pesar de su indiscutible finitud, es considerado una fuente inagotable de energía debido a la diferencia entre la escala temporal de la vida humana y la vida de la estrella a la que nos referimos, muchos órdenes de magnitud mayor.

El uso de la energía renovable ha acompañado el desarrollo de la humanidad desde tiempos remotos. Desde el uso de la biomasa para generar energía térmica para calentarse, hasta la utilización de la energía eólica transformada en energía mecánica en los molinos tradicionales o en la navegación. Sólo hacia la mitad del siglo 19, con la invención de la máquina de vapor, comenzó la utilización de los combustibles fósiles y con ello lo que se denominó la primera revolución industrial.

A pesar de constituir un periodo de gran desarrollo en términos tanto tecnológicos como de las sociedades y su bienestar, este lapso de tiempo es pequeño comparado con la historia de la humanidad. Además, este desarrollo, a pesar de haber conseguido que los niveles de vida alcancen los más altos estándares en nuestra historia, este hecho no ha sido ni mucho menos global ni 'isótropo' (homogéneo).

**1.3 Links between climate and renewable energy**

**1.4 Photovoltaic Energy**

**1.5 Climate change and the Mediterranean area**

**1.6 General scientific question: spatiotemporal behaviour of solar resource and photovoltaic production**

**1.7 Organization of the manuscript**



CHAPTER

**2**

---

## **State of knowdlege**

**2.1 Intermittency of PV**

**2.2 Variability sources in PV**

**2.3 From short to long term issues**

**2.4 Climate Change perspectives for solar resource and photovoltaic potential**



**Part II**

**Data & Methods**



---

## Data

In the present chapter, the databases used in the manuscript are described. Different types of data are used in each of the studies due to the fact that different problems are addressed. A more specific description is included in later chapters.

There are many applications from meteorology to agriculture or even health sciences, that would benefit from an accurate station network that provide high quality radiation measurements at the surface. However, it is well-known that the lack of well spread solar radiation measurements has been a constrain not only for the development of solar forecasting or resource assessment techniques, but also for the study of the whole atmospheric/climatic processes in which solar radiation takes places. The progress and improvement in the satellite-based products has helped to overcome some of these issues providing gridded data at a high spatial and temporal resolution.

### **3.1 Solar radiation measurements: *Baseline Surface Radiation Network, BSRN***

Solar radiation data from ground stations are not common at the publicly available metheorological stations. One of the main sources of high quality data are from the 'Baseline Surface Radiation Network', BSRN [KL+13].

The aim of the BSRN project is to detect changes in the Earth's radiation field that could be a consequence of climatic changes [ref web]. The monitoring network provides high-quality and high-frequency data of short and long-wave radiation fluxes from different stations around the globe, which correspond to different

### 3. DATA

---

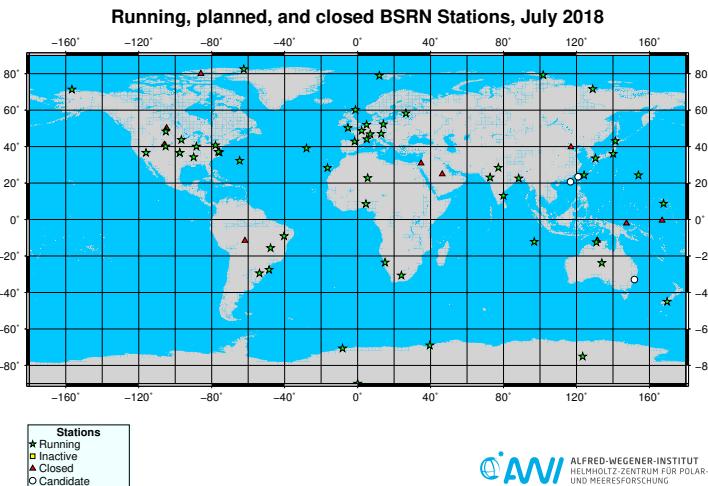
climatic zones. In figure 3.1 all the available stations from BSRN are displayed.

The World Climate Research Programme (WCRP) Radiative Fluxes Working Group initiated the Baseline Surface Radiation Network (BSRN) to support the research projects of the WCRP and other scientific programs related to solar radiation. By the time this manuscript is written, 52 BSRN are in operation. There are some stations with 'candidate' status and some that will be close in 2019. Among these stations there are different levels of data provided, from the basic measurements that include the components of solar radiation, air temperature and pressure to other variables and synoptic observations.

## 3.2 Modelling of solar radiation

In order to understand the processes that take places when solar radiation goes through the atmosphere and reaches the Earth's surface it is necessary a modelling approach. This modelling process can be made by a physical approach or using statistical methods (Festa and Ratto, 1993).

The physical modeling of solar radiation is based on physical equations of the interaction between atmospheric components and solar radiation (aerosols, water vapour, clouds...). When solar radiation goes through the atmosphere, it can interact



---

FIGURE 3.1: Solar radiation stations from BSRN.

with its components being absorbed by molecules, backscattered to space or scattered in any other direction. Thus, only part of the solar radiation at the top of the atmosphere (TOA) will reach the Earth's surface.

This process is the physical process of energy transfer described by the radiative transfer equation (RTE) and its solution needs a radiative transfer model (RTM). The radiative transfer is a very complex process that needs to be simplified to be solved numerically and some parametrizations are needed. However, from these models it is possible to reproduce the solar radiation behaviour across the atmosphere and at the surface. RTMs are in the core of numerical weather prediction models and climate models and sometimes they are used to obtain solar radiation from satellite images, although in this case empirical or semi-empirical approaches are also commonly applied.

Secondly, the solar radiation at the surface can be estimated using statistical methods. In this case it is necessary to have enough data to validate the model and likely these models will not be able to reproduce solar behaviour universally. However, they can be very useful for local applications.

In chapters 6 and 7 different climate models are used as main source of solar radiation data. The output of each model will be the result of the radiative scheme inside their codes and it can be used as the input variable for a photovoltaic production model to analyse photovoltaic potential under different climate conditions.

#### 3.2.1 CORDEX initiative

Under the acronym of CORDEX (Coordinated Regional Downscaling Experiment) there have been developed a wide range of regional climate models<sup>1</sup> that provide climate downscaled simulations for different areas around the globe<sup>2</sup>. Their importance arise with the advance in knowledge of climate change provided by global models. Whereas those models give an overview of the evolution of global conditions, their coarse resolution is not enough to understand climate change impacts at a local scale. The higher resolution of the regional climate models are needed for supporting adaptation and mitigation plans.

Two different domains from CORDEX are used in this study both focus on the Mediterranean area: EURO-CORDEX and MED-CORDEX (in figure 3.2). The hori-

---

<sup>1</sup>The chapter ?? explain the origin of regional climate modelling and the simulations used in the thesis

<sup>2</sup>Cordex website:<http://www.cordex.org/>. This international framework is under the umbrella of the world climate research program, WRCP

### 3. DATA

---

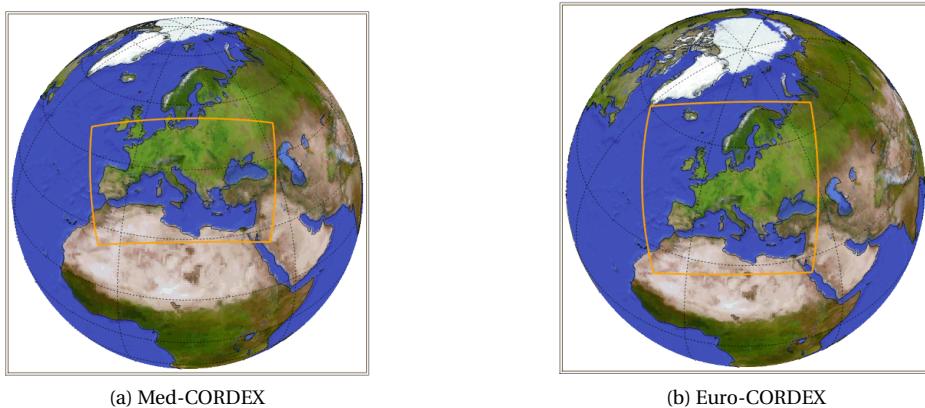
zontal resolution of the simulations included in EURO/MED-CORDEX go from 44° to 11° (from 50km to 12km).

*¿Incluir una tabla con los diferentes modelos regionales que participan en MED cordex y EURO cordex? Los que se usan en la tesis están explicados después, pero este es un punto general de dónde se han obtenido los datos.*

### 3.3 Satellite data: *Climate Monitoring Satellite Application Facility, CM-SAF*

It has been previously commented that the scarcity of available ground-based solar radiation measurements is a well-known problem. Satellite datasets have become one of the main sources of solar radition data because of their high spatial and temporal resolution as well as their increasing accuracy.

There are different methods to retrieve solar radiation from satellite images that goes from physical models to empirical ones. In the first case, the models try to explain the radiance observed by the satellite instrumentation with a radiation transfer model (RTM). In order to do that, it is necessary to know the composition of the atmosphere. On the other hand, empirical models are based in simple regression models between the visible-channel's recorded intensity and grounded measurements. There are also some approaches that are in between theses two sides: the semi-empirical models. They use a simpler radiative-transfer scheme and some statistical regressions between data from satellite sensors and observed data (Schmetz (1989), Noia et



---

FIGURE 3.2: Domains from Med-CORDEX and Euro-CORDEX. Images from wRCP-CORDEX website.

al. (1993), Pinker et al. (1995), Zelenka (2001), and Hammer et al (2003).<sup>3</sup>

In this work some satellite-data-based products from CM-SAF (The Satellite Application Facility on Climate Monitoring) have been used. The aim of CM-SAF consortium from EUMETSAT [ref] is to develop satellite-data-based-products for climate monitoring since in 2000 it was recognised the importance of using satellite for this purpose.

The CM SAF products are derived from several instruments on-board operational satellites in geostationary and polar orbits. There are two main types of products that can be obtained from the CM-SAF: operational products and climate data records (CDR). The main purpose of the CDR is to provide a high-quality database to monitor climate variability and changes (ref), as well as the detection of trends. Operational products, on the other hand, are not accurate enough for this purpose because some errors like inter-satellite biases or sensors degradation are not corrected (ref: PUM product user manual).

The SARAH dataset from CM-SAF has been used in chapters 6 and 7 of this manuscript for the analysis of shortwave solar radiation at the surface. The data is based on the records from Meteosat images, first and second generation, using the on-board MVIRI and SEVIRI instruments respectively. As the purpose of the CDR is to provide long time series covering more than 20 years, it is necessary a retrieval algorithm that can be applied to SEVIRI instruments as well as to the older MVIRI. This algorithm has been called MAGICSOL and has two parts: first, the modified Heliosat method is used to obtain the cloud effective albedo (CAL) and second, the MAGIC approach is used to obtain all sky surface radiation based on CAL. (ref PUM)

No sé si merece la pena poner algo más acerca del método. Puede ser de ayuda también para ver como se incluyen los aerosoles. Incluyo una imagen del esquema que hice del proceso de obtención de la radiación con los datos del satélite y podría estar bien incluir.

## 3.4 Photovoltaic production data

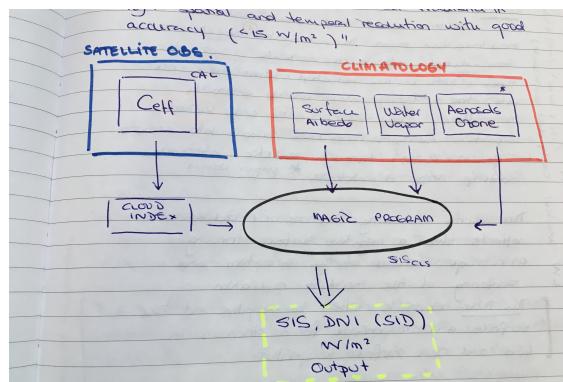
In order to evaluate the performance of a PV model it is necessary to use the data from real PV plants. However, these data is not easy to obtain, due to the fact that it belongs to private companies and it could compromise some of their strategies. This makes necessary to get agreements and confidential contracts between these companies and researchers that are not always easy to obtain if an immediate benefit does

---

<sup>3</sup>Jan Kleiss, Solar Energy Forecasting and Resource assessment, pp.22.23

### 3. DATA

---



---

FIGURE 3.3: Algorithm scheme: Retrieval of shortwave solar radiation from satellite images.

not arise from those relationships.

Some alternatives are the databases of aggregated data that are available for the European countries through the ETNSO-e data portal website. However, these data will be only useful for certain modeling exercises.

- ¿Tiene sentido que comente aquí los datos de ETNSO?
- No tengo claro si describir aquí las dos plantas que tenemos y que se utilizan en el segundo artículo

### 3.5 Temperature data

In addition to the direct relationship between solar irradiation and photovoltaic energy conversion, there are other atmospheric variables that influence the performance of solar cells, like temperature or wind speed. The model used in this work considers only temperature as a second order factor that reduces cell efficiency as it is explained in chapter 4.

Temperature data is easier to obtain than solar radiation data. Numerous weather stations provide temperature data at 2m over the surface at a local scale. For an overview of a large area gridded products are useful and easier to combine with the

satellite products.

E-OBS is a gridded dataset derived through interpolation from the ECA&D data based on stations that provides daily temperature data over Europe and the Mediterranean area <sup>4</sup>. This product has been validated in several research papers (Begert et al. 2008, Hofstra et al. 2009a,b) finding that some inaccuracies related to an oversmooth exists in areas where there are few stations, which affects mostly to the extreme analyses.

*¿Quizás incluir aquí un mapa con las estaciones a partir de las que se hace la interpolación*

---

<sup>4</sup><https://www.ecad.eu//download/ensembles/download.phpmaps>



---

## Methods

In order to fulfill the objectives of the present work it is necessary to apply different methodologies in each of the results chapters, all of them has been explained in detail at the corresponding section.

However, in the present chapter the different tools needed for the development of each study are described in a general manner. The three results chapters are based in a **modeling chain** that includes a photovoltaic model that is composed of two basic steps. First, the transposition to the plane-of-the-array, POA, of the components of the solar radiation, which implies also to assume some simple models of the atmosphere sphere seen from the generator plane and a model for the electrical performance of the system. Besides, the input of this PV model can be solar radiation measurements, satellite radition estimations o, as we saw in the previous chapter, the output of atmospheric/climate models.

In chapter 5 it is analysed the interannual variability and complementarity of solar resource over the Iberian Peninsula and for that purpose we apply **clustering techniques**. The regionalization allow us to simplify the spatio-temporal analysis of solar resource. The evaluation of the productivity, defined as the amount of energy produced by a PV system normalized by the power capacity of the system, means the **modelization of the photovoltaic system**.

The chapter 6 analyses the impact of **aerosols** in photovoltaic productivity over the Euro-Mediterranean area. Some climate simulations are used as the input of the photovoltaic model. The modelling chain allows to make a sensitivity test to quantify the role of aerosols in the area.

## 4. METHODS

---

Finally, the photovoltaic energy potential is analysed in the future, using **climate projections** from different climate models. Trends and anomalies with respect to a reference period are evaluated. In this case, a **multi-model analysis** through different RCMs simulations allow us to evaluate the solar resource under climate change scenarios. The representation of aerosols in the climate projections is considered as a fundamental variable for the shortwave downward radiation projections, SSR, and PV productivity.

### 4.1 Clustering algorithm applied to climate data

In a data-driven world, **pattern recognition** is being applied to many disciplines from biology to finance through social science, in order to obtain relevant information from different datasets. However, if the similarity measurements can not be defined first, due to unknown previous labeled data, another approach will be implemented. In that case, the task will consist in identifying the similarities afterwards, from a dataset of features, applying clustering algorithms for that purpose.

#### 4.1.1 Clustering algorithms

It is not under the scope of this work to analyse in detail all the clustering methods available, due to the vast number of them and its increasing complexity [ref: data clustering a review]. However, it is worth it to give an overview of its classification and application in order to better understand the choice made in the study of our problem.

There are different types of clustering algorithms based on its application and the criteria applied to construct the clusters. The two main categories are divided in **hierarchical** and **non-hierarchical algorithms**, but there are other taxonomies based on the algorithm construction that can be used to classify different methods [ref].

The **hierarchical clustering algorithms** create a number of nested clusters and they can be also divided themselves in agglomerative or divisive. The first one, obtain a smaller number of cluster in each step, whereas the divisive algorithms work on the other direction.

Most of these hierarchical algorithms are variants of the single-link or complete-link algorithms. They have a different way to measure similarity. In the first case, the distance between two clusters is the minimum of all the pairwise distances measured between the objects of the two clusters, in contrast, for the complete-link algorithm the distance is considered the maximum of all the pairwise distances.

On the other hand, the non-hierarchical or **partitional clustering algorithm**, create a number of non-overlapping clusters. These methods could be useful when the amount of data involved is large and the construction of a dendrogram, produced by the hierarchical algorithm could be computational expensive.

The partitional methods create the group of clusters using an **optimization function** which defines the similarity. The main inconvenient of these algorithms is the in-advanced definition of the number of clusters. Usually, the algorithm is implemented several times for a wide number of partitions and the optimum number of clusters is defined afterwards using validity index criteria.

#### 4.1.2 Clustering climate data

As previously commented, the use of pattern recognition and in particular, the clustering analysis, has been widely used across many different disciplines. The use of these techniques to group together atmospheric variables that could help in environmental classifications is a more recent research field.[ref]

Historically, the climatic divisions were based mostly on the differences in vegetation types around the globe. The Koppen-Terawata classification [ref] is the most widely known climate classification and bases its divisions in obtaining the vegetation thresholds from precipitation and temperature data. Some authors had already pointed the limitations of this classification method mostly due to two aspects: the first one is that vegetation thresholds are not well defined and that could be an issue when higher spatial resolution scales want to be defined. On the other hand is the fact that not only precipitation and temperature are influencing the vegetation species but also other atmospheric variables like solar irradiation, as well as other environmental factors that could be related to anthropogenic emissions or waste.

As these classical division are not adapted to the necessities of different fields and could be even biased, another way to geo-spatial classification is needed. In this sense it has become frequent to successfully use data-driven classification of different variables if it is necessary to give an spatially resolved answer. [ref]

#### 4.1.3 Applied clustering method

In this work a clustering method is applied to classify solar irradiation due to the fact that classical climate divisions are only based on temperature and precipitation. The spatial pattern that could be extract from the Koppen-Terawatta classification, can not be used for our purpose.

## 4. METHODS

---

A partitional commonly used clustering method has been selected to classify solar irradiation of the area. The '**k-means**' algorithm is easy to implement, due to the fact that it has been largely used over the literature. Moreover, the combination of a **principal component analysis** of the dataset previous to the k-means algorithm application has been proved to be an useful tool to reduce the data-dimensionality and apply the algorithm in a more efficient manner.[ref]

### K-Means algorithm

The k-means algorithm is a partitional clustering method that provides a set of partitions from the application of an optimization function, usually the euclidean distance, 5.1. The algorithm initiates from a pre-defined number of clusters, 'k', and randomly selects a group of centroids equal to the number of clusters. The optimization function is applied to assign each object to one of the clusters, depending on the similarity (distance) to each of the centroids. This process is applied until the algorithm converges.

$$J = \sum_{i=1}^k \sum_{j=1}^n ||x_i - c_j||^2 \quad (4.1)$$

One of the limitations of these methods is its dependency on the first selection of the centroids, that could lead to a local optimum instead to a global optimum. To solve this difficulty, the algorithm can be run many times to test the sensitivity of the algorithm to a different initial conditions. Another option is to use an initialization technique to select the centroids.

As the number of clusters 'k' is not normally known in advance, the algorithm is applied from 2 to 'n' times, with 'n' high enough to get the whole variability of the dataset and the optimum number of clusters is defined afterwards. This optimum 'k' will define the number of clusters that explain most of the variability and will assure that the increase in the number of clusters does not improve the variance representation. [ref]

### Principal Component Analysis

The Principal Component Analysis, PCA [ref], consists in a decomposition of the dataset into a number of vectors whose linear combination represents the original data. The transformation of the original data into a lower dimensional space is a reduction of the dimensionality retaining the maximum of the data variance.

The new orthogonal system has in its first coordinate axis, the projected values of the original data that preserve most of the variance in the dataset, and it is called the principal component. The rest of principal components decrease the amount of variance explained consecutively.

The PCA has been applied in advance to K-means clustering algorithm among the literature [ref]. It was explained by [ref] that the reduction in dimensionality is directly linked to K-means due to the fact that the clustering membership indicators are actually the eigenvectors given by the PCA. For this reason, it has been applied before the K-means algorithm fastening the algorithm.

### Validity Index

Clustering validation through the validation or validity index measures the goodness of the clustering results. As well as the clustering techniques, there is a classification for the methods used to validate the clustering results.

External validation techniques use information outside the dataset involved, whereas the internal validation techniques only need the information that is present in the data. The external methods know the number of optimal clusters in advance, so they are applied to select the best clustering algorithm. On the other hand, the internal validation methods will give us the optimum number of clusters after the application of the selected algorithm and only use the information that rely on the data.

Two different index are used in this study, the **Calinski-Harabasz**<sup>1</sup> index (Eq.:4.2), the **Davies-Bouldin**<sup>2</sup> index (Eq.:4.3), and the L-method proposed by Salvador and Chan, 2005 [ref]. Similar results are obtained with them. In order to analyse the optimum partition, is important to consider the nature of the variables that we are analysing. Most of the atmospheric variables, are continuous variables that could be closely related to other variables such as latitude. For that reason, the regionalization procedure can not be applied like for other discrete or non-continuous data. That characteristics should be considered when the results have to be evaluated.

---

<sup>1</sup>In the equation 4.2: The 'BCSM' is the Between-Cluster-Scatter-Matrix, and its trace is the sum of squares distances between each cluster center  $c_i$  and the global centroid vector of all the objects of the dataset. The term 'WCSM' is the Within-Cluster-Scatter-Matrix and the trace of this matrix is the sum of squares of the distances between the objects inside each cluster and the centroid. 'k' is the number of clusters

<sup>2</sup>In the equation 4.3:  $d_i$  is the averaged distance between the data classified to class i and the cluster center  $c_i$  and  $d(c_i, c_j)$  is the distance between the different cluster centers. 'k' is the number of clusters.

$$CH = \frac{trace_{BCSM}}{trace_{WCSM}} \times \frac{n-k}{k-1} \quad (4.2)$$

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{i=1, \dots, i \neq j} \frac{d_i + d_j}{d(c_i, c_j)} \quad (4.3)$$

## 4.2 Simulating a photovoltaic system

The process of simulating a photovoltaic system has two main steps. First, it is necessary to estimate the amount of energy that reaches solar cells to be transformed into electricity. This energy will be defined as **global effective irradiation**,  $G_{eff}(\alpha, \beta)$ . Secondly, the amount of energy that the system is able to give depends on the electrical performance of its components, which is the second part of the modeling process.

Both stages involve some modeling assumptions and each one would be explained in detail in the following sections. In general, due to the fact that it is very unlikely to obtain solar radiation measurements in the plane-of-array ( $G(\alpha, \beta)$ ), global horizontal irradiation ( $G(0)$ ) (the most common variable measured or modeled that can be obtain from different sources), will be the starting point. In order to estimate the amount of energy that reaches the generator surface,  $G(\alpha, \beta)$ , it is necessary to do a transposition of the different radiation components from the horizontal plane  $G(0)$  to the plane of the array (POA).

Once the solar irradiation components at the generator surface are calculated, it will be assessed how much of that energy can be transformed into electricity. The relative position between the generator and the sun gives the optical losses according to the difference with the optimum incident angle. After considering the optical losses the plane of array irradiation becomes the global effective irradiation  $G_{eff}(\alpha, \beta)$ .

The second step that will calculate the energy output will indicate the performance of the electrical components from the global effective irradiation and other variables that can influence on cells, like ambiance temperature. The whole photovoltaic system includes the generator, the inverter and the transmission elements and wires.

All the processes involved in the stage of simulating a photovoltaic system in this thesis have been made using solaR [ref], an R package that implements all the necessary functions to estimate the energy provided by the system.

### 4.2.1 Global effective irradiation

Assessing the amount of energy reaching cells of the photovoltaic generator requires to compute trackers movements, as well as the relative position between sun and panels throughout the year, a detailed description of these methods can be found in Perpinan.Marcos.ea2013. Once these equations are computed, the procedure is described below:

First, global irradiation on the horizontal plane is decomposed in two components, **direct** and **diffuse** irradiation, Eq.4.4. The third component of global irradiation, the albedo, it is not consider because its contribution is very low. To estimate these quantities we will consider equations proposed by [Liu1960] to characterize solar irradiation. Definition of *clearness index*, Eq.4.6, is the ratio between global irradiation and extra-terrestrial irradiation at the horizontal plane. They also proposed to relate that index with the *diffuse fraction*: the ratio of diffuse to global irradiation in Eq.4.5. This relation varies depending on the time scale. For daily values, we estimate correlation between the clearness index and the diffuse fraction using equations in [Aguiar1992]. After that, the diffuse component is obtained with the definition of the *diffuse fraction*, Eq.4.5.

$$G_d(0) = B_d(0) + D_d(0) \quad (4.4)$$

$$F_{D,d} = \frac{D_d(0)}{G_d(0)} \quad (4.5)$$

$$K_{Td} = \frac{G_d(0)}{B_{0d}(0)} \quad (4.6)$$

Secondly, daily irradiance has to be estimated from irradiation values. Considering low variability of solar irradiance, it is assumed that average irradiance in a short interval coincides with irradiation in that interval. Regarding equations proposed by [Aguiar1992], the ratio of the diffuse irradiance to diffuse irradiation is assumed to be equivalent to the ratio of extraterrestrial irradiance to extraterrestrial irradiation Eq.4.7, and the ratio of global irradiance to daily global irradiation is followed from the same reference Eq.4.8.

$$r_D = \frac{D(0)}{D_d(0)} = \frac{B_0(0)}{B_{0d}(0)} \quad (4.7)$$

$$r_G = \frac{G(0)}{G_d(0)} = r_D \cdot (a + b \cdot \cos(\omega)) \quad (4.8)$$

The third step considers only geometrical criteria to compute the direct and diffuse irradiance components at the inclined plane. Diffuse irradiance is calculated

## 4. METHODS

---

with the anisotropic model proposed in [hay1985estimating].

The last step estimates the effective irradiance incident on a generator subtracting dust and angle of incidence losses from the incident irradiance with the model proposed in [Martin2001].

In figure 4.1 the steps to assess effective irradiation at the plane-of-array is summarized.

### 4.2.2 Photovoltaic energy yield

Once the effective irradiation that reaches solar cells is being assessed, the transformation into power output depends on some factors regarding the photovoltaic system. In order to estimate potential for photovoltaic production, the term “yield” is defined as the energy produced by the power installed [ $\frac{\text{kWh}}{\text{kWp}}$ ]. That energy, comes from the integration in each time step of the power output of the photovoltaic system.

Considering that a photovoltaic generator is composed of modules, the generator nominal power output is calculated by multiplying the power output of a single module by the number of them, assuming the same electric performance of all modules.

$$P_{out} = I_m \cdot V_m \quad (4.9)$$

Ambient temperature influence cells performance. The assumption used to this assessment consider a linear relationship between cell temperature and global effective irradiation, Eq.4.10. NOCT in equation 4.10 is considered constant being the temperature of a cell when it works in determined conditions: irradiance of 800 [ $\frac{\text{W}}{\text{m}^2}$ ] and ambient temperature of 20°C.

$$T_c = T_a + G_{ef} \cdot \frac{NOCT - 20}{800} \quad (4.10)$$

After considering all these factors, the power output of the whole system is assessed by the consideration of a common inverter for transforming DC current into AC, also arrangement losses of the generator are included. Other systems factors that influence the performance of the photovoltaic system are shadows over the generator

---

*FIGURE 4.1: Algorithm scheme: steps on the calculation of incident irradiation at the inclined plane. Orange color means input of the calculation and blue color output or results. If a result in a previous step is used in the next one, arrows linking steps are orange. Right side of the scheme represents the method and equations needed.*

due to the positions of the PV modules over the land. This factor is not consider in our calculations assuming that we look for an estimation of the potential yield of an area, not the product of a real PV plant. Detailed description of the software employed to the assessment, solaR, is in [Lamigueiro2012]. The process for the photovoltaic output is summarized in table 4.1 including all the steps and elements involved as explained in [Perpinan2009].

## 4.3 Using Regional Climate Models

The modeling chain to obtain an estimation of the photovoltaic production of a system starts with the input of the PV model. We can consider the output of an atmospheric or climate model as the input of the PV system instead of solar radiation measurements. The uncertainty of the PV power output will increase but it is worth it to consider this option because of the lack of reliable and well spread station measurements.

In our context, with the purpose of analysing solar radiation and the potential PV production form a climatological point of view, climate models are a good tool that allow us to evaluate the resource in present conditions as well as its future evolution.

The history of climate modeling has been linked to the computational development since its beginning. The mathematical representation of the climate system was a consequence of the advance in numerical weather prediction models that took place for the first time in Princeton in 1952 and that had a fast grow. [ref and **develop a bit more?**].

Climate modeling is based on a 3-D representation of the whole climatic system, where the Earth is divided in a 3-D spatial grid whose size or resolution has evolved in time with computational power. Equations of motion, momentum and conservation laws are solved for each gridbox. As some of the atmospheric processes occurs in a finer spatial scale than the resolution of the model, parametrizations are needed for some processes such as convection. These models were first named as General Circulation Models, GCM, as their aimed was to represent main circulation flows in the atmosphere, but the development of more complex models that would start to include ocean and land, would name later these models as **Global Climate Models, GCMs**.

Since its origins, there has been a diversification and more research groups has developed its own climate models. Also, many international framework initiatives has unify efforts to understand better the future evolution of climate with the intensification and promotion of research collaboration programs that include all the models

#### 4. METHODS

---

Element	Method
PV generator	Identical modules with $dV_{oc}/dT_c = 0,475 \frac{\%}{^{\circ}\text{C}}$ and $NOCT = 47^{\circ}\text{C}$ . The MPP point calculated as in [garcia2005caracterizacion]).
Inverter	Efficiency equation proposed in [jantsch1992results]: $\eta_{inv} = \frac{p_o}{p_o + k_0^o + k_1^o p_o + k_2^o p_o^2} \quad (4.11)$ where $p_o = P_{ac}/P_{inv}$ is the normalized output power of the inverter. The characteristic coefficients of the inverters are: $k_0^o = 0.01$ , $k_1^o = 0.025$ , $k_2^o = 0.05$ .
Other losses	<ul style="list-style-type: none"> <li>• Average tolerance of the set of modules, 3%.</li> <li>• Module parameter dispersion losses, 2%.</li> <li>• Joule losses due to the wiring, 1.5%.</li> <li>• Average error of the MPP algorithm of the inverter, 1%.</li> <li>• Losses due to the MV transformer, 1%.</li> <li>• Losses due to stops of the system, 0.5%.</li> </ul>

TABLE 4.1: *Calculation procedure for the estimation of energy produced by a PV system from daily global horizontal irradiation data. Left column represents the element of the PV system and the right column the equations and methods used in each case for the efficiency of the elements.*

---

(WCRP, IPCC, CLIVAR...)

Global climate models provides reliable information about the possible evolution of climate in large scales. Their horizontal resolution, however, covers areas with very different regional characteristics, which potentially miss some information that affect in a particular manner to smaller and vulnerable areas.

It was in the early 90s (Dickinson 1989 and Giorgi 1990) when it was proposed to use global models as the necessary boundary conditions to force a 'limited area model LAM'. Those LAMs had been used for numerical weather prediction forecast, but they were applied to predict the weather just few days in advance. The idea of run these simulations for longer periods will result in the evolution of **Regional Climate Modeling**, what would provide regional climate information of processes that occurs in a spatial scale not resolved by a global models.

In a similar way than the GCMs, RCMs community has expanded in last decades and the number of research groups has increased. Also, some international initiatives try to engage different groups and modellers to develop better models and simulations (PRUDENCE, ENSEMBLES, CORDEX, EURO-CORDEX, MED-CORDEX).

In the present work we make use of RCMs focused on the Mediterranean area, considering the models and simulations included in the EURO-CORDEX and MED-CORDEX projects.

These RCMs can provide the necessary atmospheric variables for the analysis of renewable energy resources. Due to its complexity and the need of parametrizations for some atmospheric processes, they can have systematic bias that make difficult to use them for resource assessment. However, due to the low frequency variability observed in some variables and to the importance of considering the future projections, they are a valuable tool for analysing renewable energy resources and its evolution. Besides, the use of climate models allows to study specific climatic events and to understand its mechanisms, linking the implications of those situations with renewable generation.

For chapter 6, "Impact of aerosols in photovoltaic energy production", only one RCMs is used. Three different simulations are used in this chapter in order to quantify the sensitivity of PV energy production to changes in atmospheric aerosols content. The simulations are centered in the Mediterranean area, with the domain described by the MED-CORDEX initiative. The simulations length and characteristics of the aerosols dataset included are explained in detailed in the corresponding chapter.

#### 4.4 Future projections and scenarios

Behind the comprehensive study of the climate system it is not only the interest of a deeper understanding of the physics processes but also there is a concern about the impact that a changing climatic system could have in the ecosystems and different species of the Earth.

The Intergovernmental Panel on Climate Change, IPCC, was launched in 1992 and it is an international organism that was founded by the World Meteorological Association, WMO, and the United Nations Environmental Program, UNEP. Its purpose was to gather together all the scientific information available about the anthropogenic climate change. This information will summarize our knowledge about climate change, its impacts as well as some mitigation strategies were elaborated.

In order to elaborate the climate scenarios, which means the possible evolution of the climate system based on the actual climate conditions and models, it is necessary also to generate different plausible scenarios of emissions. The development of these emission scenarios of greenhouse gases or aerosols will be based on the assumptions about technological changes or demographic criterias as well as the socioeconomic development of each region. Those scenarios, with different options for the future evolution, will be the driving force for the climate simulations in order to elaborate the climate projections.

The first emission scenarios were elaborated in the 90s and they have changed since then. For the 5th AR of the IPCC, that was released 2014 were defined the RCP scenarios. They are different 'representative concentrations pathways' that correspond to the one of the possible scenarios that leads to a specific radiative forcing, i.e, the RCP8.5 represents a radiative forcing of  $8.5 \text{ [W/m}^2\text{]}$

For chapter 7, "Future projections for PV technology", several RCMs and its corresponding GCMs are used. The simulations include the historical period and the future scenario RCP8.5 or RCP4.5. The analysis of the results will be done with respect to a reference historical period.

In table 4.2 there is a summary of the climate models and the simulations that are used in chapter 6 and 7.

Study	Climate Model					Simulation
	GCM	RCM	Domain	Resolution RCM		
Aerosols' impact	CNRM-CM5	CNRM-ALADIN53	MED-CORDEX	0.44°	AER NO-AER TREND	
Future projections	CNRM-CM5	<u>ALADIN53</u>	EURO-CORDEX	0.11°	HIST/RCP85	
		<u>RCA4</u>			HIST/RCP85	
		<u>CCLM4</u>			HIST/RCP85	
	ICHEC-EC-EARTH	<u>RACMO</u>	EURO-CORDEX	0.11°	HIST/RCP85	
		<u>RCA4</u>			HIST/RCP85	
		<u>CCLM4</u>			HIST/RCP85	

TABLE 4.2: *Summarize of the climate models (global, GCM, or regional, RCM) used in chapters 7 and 8. The name of the global model includes the institute that elaborate the simulation and afterwards the global models's name. The domain of the simulations is MED-CORDEX or EURO-CORDEX, described in the referenced paper.*



## **Part III**

# **Results**



## **Long-term characteristics of solar resource over the Iberian Peninsula**



---

## **Abstract**

The spatio-temporal characteristics of different renewable resources like solar radiation, wind and precipitation are very different among them and a detailed understanding of each is important for an adequate planning and management of the electricity system. In this chapter, we elaborate a comprehensive methodology to analyse variability and complementarity of PV production that can be applied for long-term energy questions and for climate scales, but it is not limited to those time-scales or to solar resource thanks to the flexibility of the method.

The work is focused on solar irradiation as source of energy for photovoltaic (PV) generation over the Iberian Peninsula, IP, and on photovoltaic productivity, which is defined as the amount of energy generated normalized by the installed capacity at each location. The IP is an interesting site due to the fact that it is a large area with limitations in the interconnectivity with the rest of the European continent. It is also a complex climatic area [] due to its geographical location and its orography, which makes it interesting for testing the methodology.

The main steps of the method are the application of an objective clustering algorithm for performing a regionalization of the whole domain, the analysis of the temporal variability of solar radiation and photovoltaic energy yield, and the inter-comparison of the obtained clusters for examining their complementarity.

**The approach applied in this chapter means a different perspective to analyse the long-term issues in solar resource as the interannual variability, the influence of the large-scale circulation modes, trends or the synoptic patterns related to its variability. The use of the multi-step scheme allows to simplify the spatial analysis and answer the questions from a more applied point of view.**

Data from a satellite dataset of 30 years, which is usually considered as the time for defining a climatology, is used in this chapter and a long-term overview of solar radiation variability is analyzed over the Iberian Peninsula (IP). The spatial distribution and variability of photovoltaic (PV) power yield is calculated for different tracking systems. The variability is analyzed on an interannual time scale, which is relevant for energy supply security and year-to-year price stability. It shows robustness and stability of solar radiation and PV production on average for the whole domain, but with significant differences among clusters that could allow spatial compensation of PV production. The relationship between the variability of solar irradiation and of PV yield is not uniform among the different clusters. Areas where solar irradiation is higher are more sensitive to tracking type.

The whole process described in this chapter provides the information of how solar resource and the PV energy yield perform in a limited area and provide the tools to analyze the relationships between sub-areas and their variability. In this sense,

---

this method can be applied for isolated or nearly isolated electric systems located in regions with a variety of climates, or for interconnected systems involving several countries.

## 5.1 Introduction

The natural variability of renewable energy resources like solar radiation and wind presents some challenges for the management of electricity systems, which were designed for conventional technologies like nuclear or thermal power plants. For that reason a thorough knowledge and understanding of space-time features of solar radiation is needed. In the case of solar PV energy, its variability Widen2015 can be studied from the perspective of the resource or from the perspective of the PV power output, that includes some aspects of the PV generators involved in variability, like inverters or tilted and tracking panels which increase the complexity of the assessment. There are many studies focused on the short-term variability Zamo.Mestre.ea2014 that analyzed PV production ramps due to changes in solar incident irradiation associated with cloud motion Cros2014, IEA-PVPS-T14-1.32015. Also, the smoothing effect that a well-spread site planning has on the PV production is being investigated Marcos2012, Perpinan.Marcos.ea2013.

Not only short-term scales are important to address renewable resources intermittency but also longer time scales are relevant in order to make the system more efficient and reliable Davy2012. Policymakers, transmission system operators and investment companies, need an accurate evaluation of resources availability in present and future climate conditions for their mid and long-term planning. The analysis of interannual variability has a particular importance in order to assess stability of the resource and the financial viability of renewable energy plants pryor2006inter, as well as the likelihood of strong electricity price oscillations like the ones associated for example with the large interannual variations of hydroelectric production.

Regarding that perspective of long-term variability of solar resources, there are studies focused on long series from stations Sanchez-Lorenzo2009, Sanchez-Lorenzo2013, vazquez2012interannual or reanalysis data that identify low frequency changes in solar radiation, as the “dimming” and “brightening” periods Wild2005, that show relationship between solar irradiation and anthropogenic aerosols Nabat2014a. Some studies examine the influence of large-scale circulation atmospheric modes like the NAO (North Atlantic Oscillation) on solar radiation Pozo-Vazquez2004, Jerez2013, while others study the spatial variability instead of the temporal variability Gueymard.Wilcox2011a.

Some authors make use of regionalization techniques for atmospheric variables in order to analyze them from a climatological point of view Argueso2011 or for solar energy purposes, mainly for operation and short-term assessment Zagouras2013, Zagouras2014, Zagouras.Pedro.2014.

In this chapter a multi-step scheme that systematizes the time-space comparison of solar irradiation or photovoltaic productivity among sub-regions of the target area

is described, taking into account a large set of factors involved in PV production that could affect the variability of photovoltaic energy yield. Spatial complementarity between clusters is analyzed through correlation coefficients of solar irradiation and PV energy yield time series, showing possibilities for compensating PV production shortages in certain clusters.

The analysis over the Iberian Peninsula show that most of the results obtain in previous studies can be also seen with this approach. In addition, the methodology allow to give a clearer spatial answer to the questions as well as to sistematize the procedure.

This chapter is organized as follows: in first place, a description of the methodlogy is presented. Each stage of the multi-step scheme is summarized in section ?? and explained in more detail in the Methodology chapter. After that, the clustering algorithm is applied to the irradiation over the Iberian Peninsula and main results are shown.

### 5.1.1 Methods

The whole comprehensive scheme followed is represented in figure 5.1, showing the 4 stages and procedures that will be explained in detail in the following section.

From data of solar irradiation on the horizontal plane as the starting point of the method, the most common variable obtained from different data sources, the scheme will provide different outputs that can be used to evaluate resource and PV production:

- Regionalization of the area to facilitate the spatial analysis.
- Resource and energy yield aggregated by areas.
- Interannual variability of the resource and the PV production.
- Evaluation of the complementarity of the resource and PV production among areas.

#### Regionalization by clustering

Regionalization procedures provide the ability of extracting general information of the areas that could be treated as a coherent unit, facilitating the analysis and not

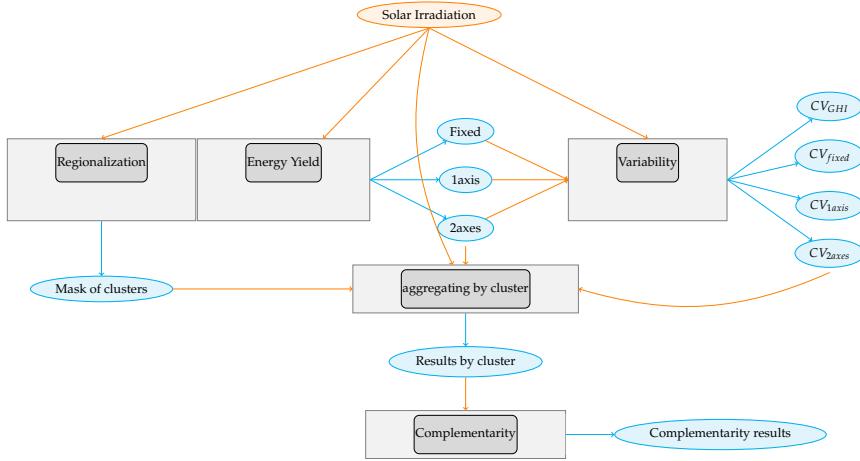


FIGURE 5.1: Scheme: Each gray block represents each of the operations needed to get the variability and complementarity results. Orange ellipses are the data employed and blue ellipses are the results of each stage. If the results of one of the blocks are used as input for another stage, connectors are represented in orange color.

considering those characteristics that are not under study. As it was explained in the methodological chapter, classical climatological classifications have some grade of subjectivity due to the fact that they rely on arbitrary assumptions Kottek2006 and their criteria are based on temperature and precipitation trewartha1980koppen. For our purpose, objective and data-derived criteria are more suitable due to the fact that a different variable is analyzed and its classification does not match classical climate divisions. Objective methods based on clustering techniques have been applied successfully over the literature for the analysis of renewable energy resources Polo2015, Zagouras2013, Zagouras2014, Zagouras.Pedro.2014, gomez2015characterization and atmospheric variables Argueso2011, garcia2012seasonal.

A commonly applied regionalization methodology includes the Kmeans algorithm after preprocessing the data through Principal Component Analysis Ding2004. This two-step method first reduces redundant information by a Principal Component Analysis that decreases dimensionality of the original dataset. After that, k-means algorithm is applied to the reduced data to find the optimal partition of clusters, which is based on similarity between each element or object inside the cluster and its centroid. This is considered as the most representative element of the cluster, and similarity is

measured by an objective function defined in the cluster algorithm.

This method presents some problems regarding the random selection of the cluster centroids in the first step. Different initial centroids can lead to different solution or a local optimum could be found. Also, there could be some computational problems if many iterations are needed to get the final partition.

The procedure used in [Argueso2011] and in [Zagouras2014] is adapted to get the optimal partition in our scheme from a combined clustering grouping and avoiding the above mentioned problems: the **k-means partitional algorithm is initialized with a hierarchical clustering solution of a the dimension-reduced data by a Principal Component Analysis**. For the particular case applied in this work, vectors of daily solar irradiation are used for the regionalization. The following steps are needed to get the optimal partition of clusters in the area:

- **To reduce data dimensionality.** Principal components are eigenvectors of an orthogonal matrix after applied a singular value decomposition (SVD) to the original data, daily solar irradiation vectors, whose initial dimension is reduced to the first eigenvectors that retain 95% of the variance. Considered that, a linear combination of these eigenvectors represents the initial data.
- **Hierarchical clustering to initialize k-means.** A hierarchical clustering method classifies data based on a hierarchy. If it is agglomerative, it will start with a cluster for each observation of the data and observations will group together recursively by similarity using the “complete linkage” method. Once the hierarchy is obtained, centroids can be calculated for each emerged partition with a number of clusters between 2 and  $n$ , where  $n$  is a high enough selected number of clusters. Centroids will be the initial seed for the Kmeans algorithm, avoiding the computational problems and favoring reproducibility.
- **Kmeans algorithm.** The k-means algorithm is a partitional clustering method that minimizes an objective function that defines similarity among the elements of each cluster. In our case we made use of the Euclidean distance, Eq. 5.1 between the objects or elements in the cluster and its centroid as the objective function. The number of clusters in which the data is divided into has to be known beforehand. In order to overcome the inconvenience, the algorithm is run from 2 to  $n$  clusters and the optimum number is determined by making use of a clustering validity index after that.

$$J = \sum_{i=1}^k \sum_{j=1}^n \|x_i - c_j\|^2 \quad (5.1)$$

- **Validity index.** In order to determine the optimal partition, validity clustering techniques are applied. There are two types of validation for the clustering methods. First, external clustering validation that make use of external information out of the data; and secondly, there are internal clustering validation methods that rely only on information from the data 5694060. The latter are used to preserve objectivity as much as possible and are based on two criteria: compactness and separation of the clusters emerged. We use one of the most applied validity index, the Calinski-Harabasz index CalinskiH, CH, that evaluates the average between and within cluster sums of squares.
- **L-method.** CH index is calculated for every partition from 2 to  $n$  clusters. The resulting CH graph in figure 5.2 for the Iberian Peninsula regionalization is shown for a number of clusters between 2 and 70 as an example. Theoretically, the partition with the maximum CH is the optimum, but the graph shows a decreasing trend which leads to imprecision in finding the optimum. The large number of data and the continuous variable analyzed are responsible for that. For that reason, the L-method is applied Salvador2004. This method selects the intersection of two best-fit lines in the graph CH vs.  $k$ , where  $k$  is the number of clusters of the partition Zagouras2013. All possible pairs of lines that fit linearly to the left and right sequence of data points are created. Each line has at least two points. The total root mean squared error is calculated as in Eq.:5.2:

$$RMSE_T = \frac{c-1}{k-1} RMSE_{left} + \frac{k-c}{k-1} RMSE_{right} \quad (5.2)$$

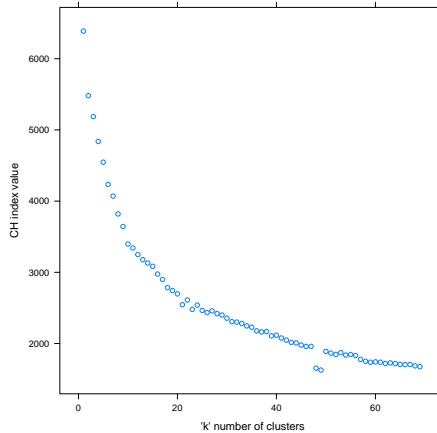


FIGURE 5.2: *Calinski-Harabasz index by 'k' number of clusters*

Where  $c$  is the number of clusters where the graph is split into the two fit-lines,  $k$  is the total number of clusters. The “total root mean square error” is a weighted error with two terms, one for each side of  $c$  in the graph. Each side has a heavier weight depending on the points involved in the fitting. The minimum of  $RMSE_T$  gives us the optimum number of clusters of the data Zagouras2013 which are used in the following steps.

### Photovoltaic energy yield

The simulation of a photovoltaic energy system is described in a previous chapter. Here it is summarized in order to do not miss the coherence of the text.

The assessment of the power output of a photovoltaic system is carried out in two main steps:

1. In first place, global irradiation on the horizontal plane  $G(0)$ , which is the most common variable obtained from data sources, has to be transformed into the plane-of-array irradiation,  $G(\alpha, \beta)$ , where  $\alpha$  is the azimuth angle and  $\beta$  the inclination angle of the generator plane. Due to optical losses (reflection, angle of incidence, and dust), the irradiation available is reduced for the photovoltaic cells inside the panels and the plane-of-array irradiation is then denoted as effective irradiation on the PV generator  $G_{eff}(\alpha, \beta)$ .

Three different types of tracking types are considered for the photovoltaic generator that influences on the tilt of the panels:

- **Fixed** panels with an optimum angle of inclination that depends on the latitude of the place.
  - **North-South** oriented panels that track the sun daily varying the azimuth angle, we will refer to them as “one axis”
  - **Two-axes** tracking system that allows variation of the azimuth and inclination angles, we will refer to them as “two axes”.
2. Once the effective irradiation that reach solar cells has been assessed, second step is the transformation into power output that depends on the photovoltaic system. The photovoltaic system is composed of a PV generator, consisting of several PV modules, and an inverter to transform the DC current output from the generator into AC current to be integrated into the network. In order to estimate potential for photovoltaic production, the term *yield* is defined as the system energy produced divided by the power installed [ $\frac{kWh}{kW_p}$ ].

### Variability and complementarity

The metric to analyze interannual variability is the coefficient of variation, CV Eq.5.3, which is defined as:

$$CV = \frac{\sigma}{\bar{X}} \quad (5.3)$$

In this equation,  $\sigma$  is the standard deviation of the variable analyzed and it is divided by the mean of the variable in the period of the study. Sometimes CV is represented in percentage. This measure is dimensionless and can be applied in different time scales, which is helpful for comparisons.

To assess complementarity of the solar resource in the area of study, the Pearson's correlation coefficient between the time series of pairs of clusters Eq.5.4 is calculated:

$$\rho_{i,j} = \frac{\sigma_{c_i, c_j}}{\sigma_{c_i} \sigma_{c_j}} \quad (5.4)$$

In this equation,  $c_i$  and  $c_j$  are the time series corresponding to the clusters  $i$  and  $j$ . The concept of complementarity is associated with negative correlations between sub-regions of a wider area. If there is such a complementarity, a positive change in the variable for one of the clusters will be associated to a negative change for the other one. This is relevant for identifying spatial compensation possibilities and reducing overall variability in a network with high penetration of solar PV energy.

Complementarity can occur in different scales, either spatial or temporal and to understand it sometimes is a matter of balance. Spatial resolution for the complementarity assessment must be high enough to make sense of the comparison between zones, due to the fact that it is clear that geographically dispersed areas, far from each other, will have very different evolution of atmospheric variables, but may not be interesting from the electricity generation point of view. On the other hand, if the area of study is too small, atmospheric variables and therefore renewable resources will evolve in a very similar way. For such small areas, local complementarity between different resources can be analyzed, but not spatial complementarity of one resource. The methodology presented here addresses this issue by applying the inter-cluster comparison, that ensures homogeneity within each cluster and differences between them.

Correlation coefficient for a long time series may hide changes in complementarity for shorter sub-periods with higher frequency correlations. For that reason a moving window is applied to the time series calculating the correlation coefficient and providing an indication of how complementarity between clusters varies during the

whole period. Width of the window depends on the length of the time-series analysed.

In order to obtain the more important cluster pairs regarding complementarity, the median of the correlation coefficient series is calculated for each pair. After that, the cluster pairs are reordered from lower to higher values of this median correlation.

### 5.1.2 Results

The optimal partition after having applied the clustering method is represented in figure 5.3. The CH validation procedure gives an optimum number of 19 clusters for the area, where each of the clusters has an homogeneous time evolution of solar irradiation. Due to the nature of clustering techniques, there is not an unique/best method to select the optimum partition. Another index (Davies-Bouldin, [davies1979cluster]) has been applied for comparison, and the obtained optimum number of clusters was of the same order than for CH index.

#### Variability and complementarity results

After regionalization, it is performed an analysis of solar irradiation on the horizontal plane and PV yield by tracking system, including their temporal variability.

Regarding interannual variability, we have calculated the CV of two time-aggregated means of solar irradiation and PV yield:

- On one hand it is applied for the yearly mean of daily irradiation  $G_{d,y}(0)$  and yearly PV yield. This metric gives the variation of the energy from one year to another and if it is low, general stability of the solar resource and PV production is guaranteed.
- On the other hand the interannual variability of the monthly time series  $G_{d,m}(0)$  and monthly energy yield is also investigated in order to quantify differences in the annual cycle.

The CV is also aggregated by cluster, in order to facilitate the intercomparison among areas.

Power from the PV generator depends quasi-linearly with solar irradiation at the plane-of-array ( $G_{eff}(\alpha, \beta)$ ), besides second order effects (spectrum, wind, etc) Perpinan2007 . Due to that the fixed typology is the one with lower yield because the amount of irradiation reaching cells is lower than the amount of energy reaching panels when trackers have one or two axes movements.

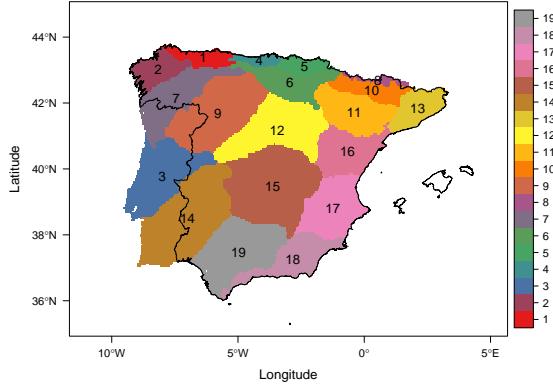


FIGURE 5.3: *Optimal partition of 19 Clusters after applied the algorithm and the validity index*

For areas where solar irradiation is higher, yield differences between trackers are higher. This can be seen in figure 5.4 where yearly mean yield for the 30-years period is aggregated by cluster and tracking system, and clusters are sorted vertically from less to more energy yield. A noteworthy result is that yield increase from fixed panels to one-axis panels is non-linear. This increase ranges between 17% for the clusters with less solar irradiation, located at the northern coast (clusters 4, 5), and 30% for the southern clusters with more solar irradiation (clusters 18, 19). In contrast, energy yield increase from one-axis to two-axes panels is almost constant, around 12% for all clusters. A consequence of the non-linear PV yield increase from fixed to one-axis panels is that the energy yield differences between clusters are much higher for tracking than for fixed systems. While for fixed panels PV energy yield varies between 1000 and 1450  $\frac{\text{kWh}}{\text{kWp}}$ , for two-axes systems it varies between 1350 and 2100  $\frac{\text{kWh}}{\text{kWp}}$ . These average values are coherent with results obtained in [Antonanzas-Torres2013] when considering a value of 0.75 for the system performance ratio.

Electricity price variations significantly depend on the variations of the monthly renewable electricity production from year to year. This time-scale is also the most influenced by the large scale circulation modes for solar potential in the Iberian Peninsula Jerez2013a. The winter half of the year, from October to March, is especially variable.

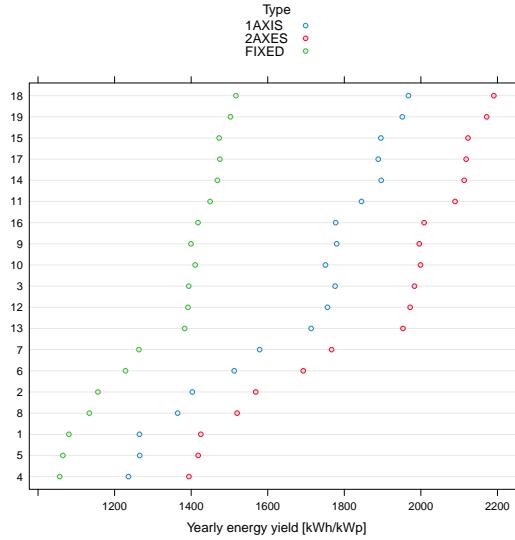
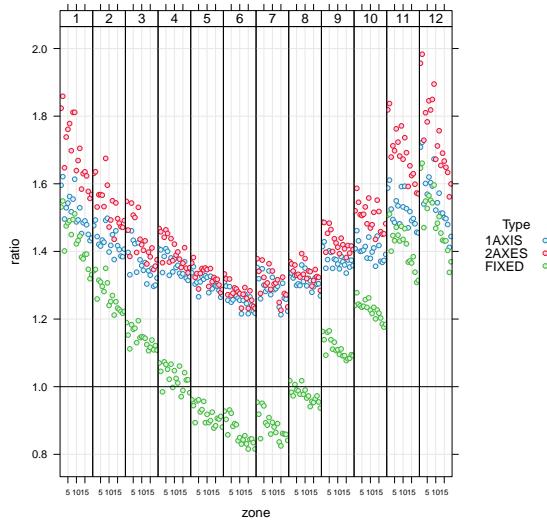


FIGURE 5.4: *Yearly mean of PV yield by cluster and for each tracking system [ $\frac{\text{kWh}}{\text{kWp}}$ ]. Values are sort from lower to higher yield values.*

The interannual variability for monthly yield is higher than for the irradiation at the horizontal plane, as it occurred for yearly values which results are in the appendix. In winter months, these differences in CV are much higher than in summer. This behavior is more pronounced in northern areas.

In order to quantify these differences in variability between solar irradiation and solar power output, the ratio between variability of yield by tracking system and solar irradiation is represented in figure 5.5 for each month and cluster. If CV of energy yield is higher than CV of solar irradiation, values are above one. On the other side values will be below one if CV of solar irradiation is higher than CV of energy yield.

The highest ratios are obtained between  $CV_{2\text{axes}}$  and  $CV_{G0}$ . The ratio of  $CV_{1\text{axis}}$  is clearly lower in winter months, but is very similar in summer to the ratio of  $CV_{2\text{axes}}$ . Yield with an 'horizontal' axis tracker and 'two-axes' trackers increase the variability between 20% in summer and more than 80% in some areas in winter. The fixed typology ratio,  $CV_{fixed}/CV_{G0}$  has a much wider range in the whole year. In winter months, it has values between 1.2 and 1.6, depending on the cluster, and is not far from the other two typologies. In contrast, this ratio decreases rapidly in summer months, reaching values below one between May and August. This means that for that period, variability of the 'fixed yield' is smaller than variability of solar irradiation at the horizontal plane.



**FIGURE 5.5:** *CV ratios between each type of tracking system and solar irradiation at the horizontal plane, grouped together by month in the graph. Ratios are calculated for each cluster, represented in the x axis. “Fixed” represents  $\frac{CV_{fixed}}{CV_{G0}}$ , “1 axis” is  $\frac{CV_{1axis}}{CV_{G0}}$  and “2 axis” is  $\frac{CV_{2axes}}{CV_{G0}}$*

The results of CV show that variability of PV energy yield at tilt panels is higher than variability of solar resource at horizontal plane in most cases, explained by the nature of solar irradiation at tilt panels and its dependency of solar irradiation at horizontal plane, [Perpinan2009].

The monthly time series are also selected for the complementarity analysis.

Regarding solar power complementarity, opposite-evolving time-series for different areas would strongly increase the reliability of the whole electric system, as shortfalls of solar irradiation in certain areas could be compensated by above-normal irradiation in others. However, this ideal situation is difficult to find in a rather limited area like the IP, at least for monthly timescales over a long time period of 30 years. In this case, the absence of correlation becomes also important, as it avoids simultaneous shortfalls or simultaneous above-normal values, and therefore softens the overall power production. The correlation matrix for the 30-year period and each month is in the appendix, showing the results for the whole period. In most cases, the correlation coefficient is highly positive, specially between pairs in the northern part of the Iberian Peninsula. For the southern part, the correlation coefficient is also positive but it decreases in July and August. There are some exceptions between the north-

ern clusters 4 and 5 and the southern clusters 14 to 19, these pairs are only slightly correlated, not correlated or even slightly anticorrelated for some months.

Overall, southern and eastern clusters are uncorrelated at least during part of the year with northern and northwestern clusters. In some cases, the absence of correlation is found between nearby clusters: in winter months, the north-eastern cluster 11 (central Ebro valley) is uncorrelated to the closely-lying clusters 4, 5 and 8 (in the northern coast and Pyrenees). This is probably related to persistent atmospheric situations with north to northwesterly winds, that cause cloudiness in the windward clusters and clear skies in the leeward Ebro cluster, due to a foehn effect. This result points out the selective character of the clustering method.

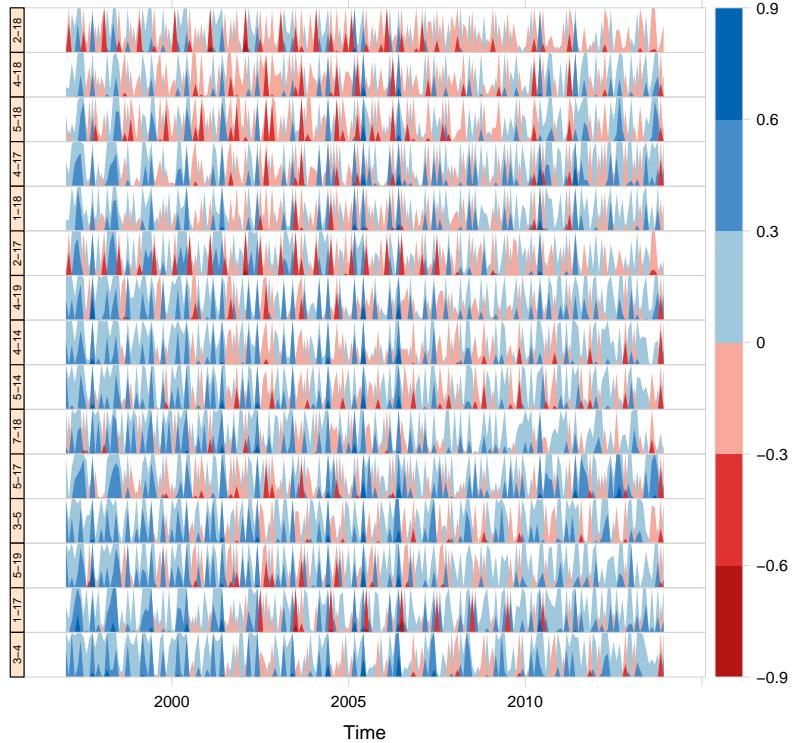
It could be that the obtained clusters present higher complementarity in shorter sub-periods. We have divided the whole 30-year period in sub-periods of consecutive 15 years. The correlation coefficients have been calculated again for the resulting 15-year moving window, for each pair of clusters and for each month. In this way, we obtain how each correlation coefficient evolves during the 30 year period. The analysis has been applied for the four variables in the study: solar irradiation at the horizontal plane and PV energy yield for each tracking system.

The median value of the correlation coefficient series is calculated for each pair. Each serie comprises 12 monthly values for each of the 15-year moving windows. After that, the cluster pairs are reordered from lower to higher values of this median correlation, and the first 15 pairs (with the lowest median correlation) are selected as the most representative of complementarity. These cluster pairs are represented in figure 5.6, showing the time-evolution of its correlation coefficient. These results are for solar irradiation on the horizontal plane. The corresponding figures for PV yield by tracking system are not shown due to the similar results obtained for this analysis.

The most relevant pairs in terms of complementarity include a northern (1, 2, 4 or 5) and a south-eastern cluster (17 or 18), as can be seen in figure 5.6. The negative correlations for these pairs reach values below -0.6, at least in some 15-year sub-periods. These clusters are negatively correlated in most cases. Clusters 19 and 14 (southern and south-western IP) are also negative correlated with clusters in the north, although with lower values than the south-eastern clusters.

It is important to notice the appearance of cluster pairs 3-4 and 3-5 in this figure. These clusters are closer than the previously commented cases, which highlights the adequacy of the clustering method. All three are Atlantic coast clusters, but while cluster 3 includes part of the western coast, clusters 4 and 5 are northern coast areas. This fact, together with the position of the main mountain ranges, can explain their partially complementary behaviour.

In order to highlight the months with maximum anti-correlation, figure 5.7 presents, for the same 15 cluster pairs as above, the minimum values of the monthly correlation coefficient (where the minimum for each month is calculated over all 15-year sub-periods). Differences between months are clearly observed in this graph. Only two months (March and June) show consistently positive values of this parameter, and



**FIGURE 5.6:** Correlation coefficient of solar irradiation at the horizontal plane: evolution of a 15-year moving window of monthly values, for the 15 cluster pairs showing the smallest median correlation values. Monthly negative correlations (in red) and monthly positive correlations (in blue) are represented in the same axis to facilitate the comparison of the multiple time-series. Also, higher correlation values overlap with lower ones, enabling a compact presentation of all the information in a narrower plot. The cluster pairs are indicated on the left, while the year in the x-axis indicates the end of each 15-year moving window.

therefore a low complementarity. In the other months, this parameter has predominantly negative values, revealing a certain degree of complementarity. July, August and September show rather low values and relatively high complementarity, which is important as these are months with a high productivity and also include the summer demand peak. In general, for the second half of the year, the values of this minimum correlation value are lower than for the first half of the year.

### 5.1.3 Conclusion

### 5.1.4 Summary

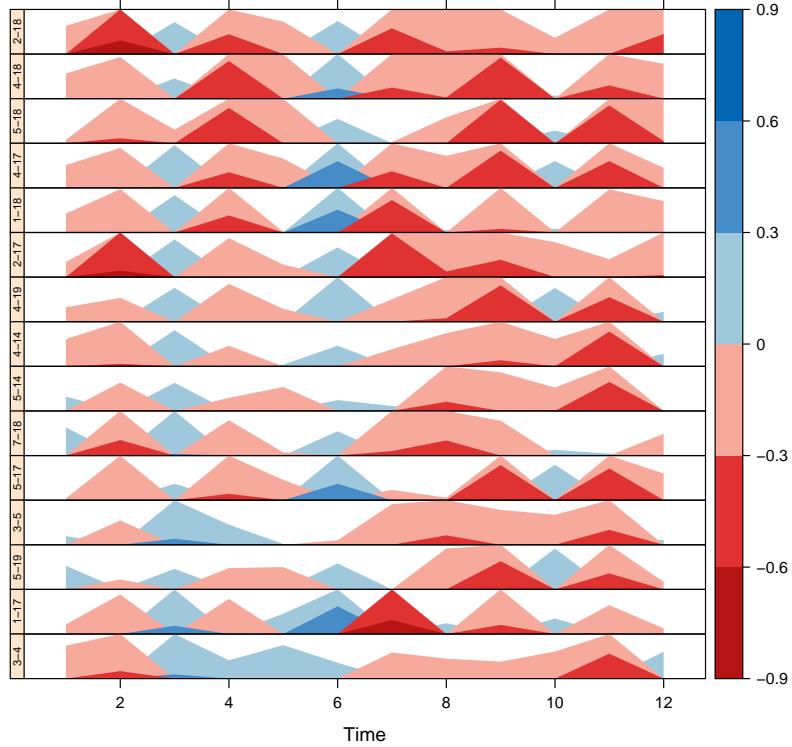


FIGURE 5.7: *Correlation coefficient of solar irradiation at the horizontal plane: minimum values of the monthly correlation coefficient, for the same cluster pairs as figure 5.6. The minimum is calculated over all 15-year sub-periods. The type of representation of correlation values is the same as in figure 5.6. The cluster pairs are indicated on the left, while the numbers in the x-axis indicate the months.*

---

## **Impact of aerosols in photovoltaic energy production over the Euro-Mediterranean area**



---

## Abstract

The increase in the photovoltaic energy installed capacity over the world leads to the need of a better understanding of solar resource and its variability. The aim of this work is to assess the influence of aerosols on photovoltaic energy production from seasonal to multi-decadal time scales. For this purpose we use various coupled aerosol-climate simulations that take into account the complex spatial and temporal patterns of natural and anthropogenic aerosols over the Euro-Mediterranean domain.

The results show that aerosols strongly influence the spatial pattern, seasonal cycle and long-term trend of PV production. The most affected area is Central Europe where sensitivity of PV production to aerosols is higher. The annual production loss due to aerosols ranges from no impact to  $-16\%$  in The Netherlands, with variation depending on the area and on the typology of the tracking system. The summer production loss can even reach  $-20\%$  over regions of Africa and Syria-Iraq.

We conclude that aerosols can not be neglected in the assessment of PV production at large time scales over the Euro-Mediterranean area. Besides, the potential increase in energy due to reduction in the anthropogenic aerosols is shown in the simulation of the brightening period over Europe, with an increase of  $2000 \frac{\text{kWh}}{\text{kWp}}$  in a PV lifetime for the most affected areas. It illustrates the evolution that PV potential could follow in highly polluted areas through the effective implementation of pollution control measures.

## 6.1 Introduction

In the last decades there has been an overall increase in the deployment of photovoltaic (PV) power plants over the world, demanding research in the fields that contribute to a better integration of this technology into the electricity system. Due to the variability of solar resource, a comprehensive study of its spatiotemporal behavior is needed.

Aerosols particles influence the climate system directly, affecting the Earth's radiation budget by scattering or absorption of solar radiation, or indirectly, changing cloud properties. Due to their importance in the amount of solar radiation that reaches the Earth's surface and their effect in the changing climate, the evaluation of their impact on the solar resource for energy purposes is of special interest.

A general lack of well-spread surface stations, which provide solar radiation measurements, has made the satellite information the main and most reliable source of data up to now, due to its spatial and time resolution Posselt2012, Ineichen2014. However, satellite retrievals were not available several decades ago and they do not allow to quantify the effect of specific factors on solar irradiance. Thus for investigating long-term statistics and to disentangle the various factors influencing solar resource, a different approach is needed. Models are the best tool to understand processes that occur in the atmosphere and their link with the resource variability, as individual factors can be included or removed in them, allowing the isolation of their effects.

Due to the increasing concern about the availability of renewable energy resources under climate change scenarios, climate modeling has revealed itself as a valuable tool for evaluating future energy potential Crook2011, Gaetani2014, Gaetani2015, Jerez2015, Jerez2015climix, Tobin2016. However, representation of the clouds is still one of the main challenges for these models, and the spatio-temporal variability of aerosols is rarely taken into account in some of regional climate models Bartok2017, which could lead to significant errors in PV power forecasting or future energy estimations Rieger2017. Such climate simulations have to be combined with an accurate PV model capable of reproducing the system performance. Existing studies analysing the influence of aerosols on solar irradiation lack spatial detail (because of the use of relatively coarse global climate models) and/or do not apply a detailed PV production model Bergin2017.

The Mediterranean region is considered as highly influenced by aerosols coming from different sources Lelieveld. These aerosols have a deep impact on the climate of the region Nabat2014, Nabat2014a, thus on the shortwave solar radiation reaching the surface Mallet2016. Regarding possible changes in anthropogenic aerosols in the future Gaetani2014, Jimenez-Guerrero2011, the relevance of the near-term climate change scenarios and the expected PV deployment, the study of the Euro-Mediterranean area is important for solar energy.

In this work we use a regional climate model Nabat2014 with a realistic aerosol representation combined with an accurate PV model Lamigueiro2012. The influence

of these aerosols in the spatiotemporal variability of PV production over the region is quantified. The analysis is made in present climate conditions for simulations between 2003 and 2009 and different tracking types are considered in the study, due to the different sensitivity of each typology to changes in solar radiation Gutierrez2017. On the other hand, the impact that trends in anthropogenic aerosols have in PV energy production is also investigated using longer simulations for the “brightening” Wild2005 period, 1980-2012, reflecting how pollution control policies could benefit the PV energy production in highly polluted areas.

This paper is organized as follows: in section 2 the climate and the photovoltaic model are described. In addition, there is a description of the aerosols and the datasets used for evaluation. Section 3 presents the results and shows the impact of aerosols on photovoltaic energy production. It is organized depending on the space-time scale analysed and there is a subsection for the tracking system sensitivity. Finally, section 5 is a discussion section for limitations and future perspectives and section 6 shows the main conclusions.

## 6.2 Data and Methods

Different climate simulations are used as an input of a PV power model. These climate simulations provide the daily-mean shortwave solar radiation, SSR, at the surface. The energy production model simulates the performance of a general photovoltaic system and includes different tracking types, considering the tilt of photovoltaic panels as a relevant component of the whole assessment. Computation of the photovoltaic energy model is made using the R open-source package named solaR Lamigueiro2012.

### 6.2.1 Climate Data

The climate model used in this study (CNRM-RCSM4, Sevault2014) is a coupled Regional Climate System Model (RCSM) dedicated to the study of the Mediterranean climate. CNRM-RCSM4 is one of the RCSMs contributing to the multi-model Med-CORDEX initiative Ruti2016. It has the specificity to represent various components (atmosphere, land surface, river, ocean) of the Mediterranean regional climate system at high-resolution as well as their high-frequency coupling. The horizontal resolution is 50 km for the atmosphere, the land surface and the river network, and about 10 km for the Mediterranean Sea. In addition, the atmosphere part of the model, the so-called ALADIN-Climate version 5.2 Colin2010 is one of the few available Regional Climate Models which can take into account a realistic representation of the spatiotemporal variability of the aerosols Nabat2014. The model has been extensively described, evaluated and intercompared with other Med-CORDEX models in previous studies, [Sevault2014; Nabat2014; Nabat2014a; Flaounas2016; Gaertner2016; DellAquila2016; Harzallah2016; Cavicchia2016]

## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

---

The detailed interannual aerosol dataset used in the climate simulations Nabat2013, NAB13, is able to reproduce the spatiotemporal variability of AOD (aerosol optical depth) over the Mediterranean region. It improves the representation of aerosols against older climatologies commonly applied in regional climate studies like [Tegen1997] or [tanre1984first].

The NAB13 dataset includes five different aerosol species: Sea Salt, Black Carbon, Sulfate, Organic Carbon and Desert Dust (ss, bc, su, or, sd); with spatial and temporal variability. It is based on a blending of a satellite-derived AOD product and a high-resolution regional climate model using up-to-date interactive aerosols module. This dataset has also been evaluated against ground stations Nabat2013.

During the eighties, some policies against the emissions of certain types of anthropogenic aerosols were implemented in Europe, which has been linked with the observed increase in the shortwave solar radiation in the area Wild2005. For simulations over this commonly named “brightening period” (1980-2012), a trend for sulfate aerosols is included in NAB13, being able to reproduce the shortwave solar radiation trend observed over Europe since 1980 Nabat2014a.

There is a large spatial and seasonal variability of the AOD at 550nm over the Euro-Mediterranean. Spring and summer months are highly influenced by dust aerosols in the south of the domain. In winter, anthropogenic aerosols dominate in central Europe and during autumn, there are few areas with high values in opposition to the rest of the domain.

The domain considered in the simulations covers the Mediterranean area in addition to a large part of Europe (see figure 6.1).

For the first period, 2003-2009, a pair of runs is analysed. We refer to them as **AER** and **NO-AER**. The AER simulation includes the NAB13 dataset, whereas no aerosols are included in NO-AER. This pair allows to easily attribute the obtain differences to the aerosols effect, therefore to quantify the impact of aerosols on the Euro-Mediterranean SSR and PV productivity. It is an important point considering the fact that some of the state-of-art RCM do not include aerosols in their simulation, so it gives an idea of that missing forcing.

Secondly, a longer simulation between 1980 and 2012, **TREND**, covering the “brightening” period observed in Europe is also analysed. It will show the effect of a decreasing trend in sulfur aerosols on the shortwave solar radiation and on the PV productivity.

A summary of the different simulations is reported in table 6.1.

### 6.2.2 PV model description

In this section we make use of the commonly used terms: shortwave solar radiation, SSR, is equivalent to the global irradiation at the horizontal plane  $G(0)$ , which is composed of the beam component,  $B(0)$ , the diffuse irradiation,  $D(0)$ , and the albedo  $R(0)$ .

Simulation	Aerosols	Period
AER	NAB13	2003-2009
NO-AER	Not included	2003-2009
TREND	NAB13 + sulfates trend	1980-2012

TABLE 6.1: *Simulations of the CNRM-RCSM4 regional climate model to obtain SSR and temperature as input of the photovoltaic model, period and representation of aerosols in each simulation.*

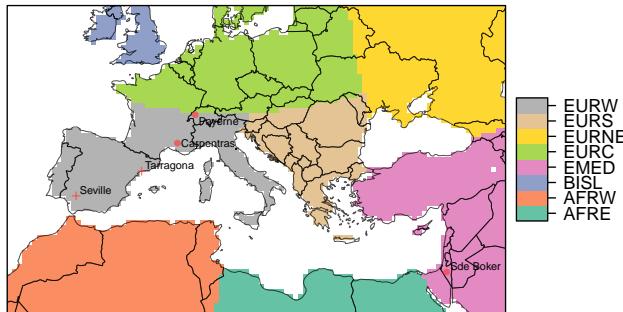


FIGURE 6.1: *Areas defined to evaluate the difference in solar radiation between the satellite and the climate model simulations: Western Europe, EURW; Southern Europe, EURS; North-Eastern Europe, EURNE; Central Europe, EURC; Eastern Mediterranean, EMED; British Island, BISL; Western Africa, AFRW; Eastern Africa, AFRE. The location of the BSRN stations is represented with points “.” and the two PV plants are plotted with “+” in orange.*

## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

---

A photovoltaic energy system is mainly composed of the generator, consisting of several modules that include the cells to transform solar radiation into electricity, and the inverter, which transforms direct current into alternate current to be integrated into the electricity network. The assessment or estimation of the power output of a PV system implies modeling its components. For this purpose, two main steps have to be considered. The first is related to the amount of energy that reaches the generator surface and the other is based on the performance of the electrical components. The procedure is described in [Perpinan2009] and the same methodology was applied in [Gutierrez2017].

1) Global irradiation at the horizontal plane  $G(0)$ , as an output of the climate model, has to be transformed into the global effective irradiation,  $G_{eff}(\alpha, \beta)$ , which is the amount of energy reaching the tilted surface of the generator (where  $\alpha$  is the azimuth angle and  $\beta$  the inclination angle) after considering reflection losses, angle of incidence and accumulated dust.

First, for the decomposition of  $G(0)$ , empirical equations that correlates the *clearness index* and the *diffuse fraction* Page1961 are used. The clearness index is a measure for the energy lost when solar irradiation goes through the atmosphere, and the diffuse fraction is the relationship between the diffuse component,  $D(0)$ , and the global irradiation at the horizontal plane  $G(0)$ . Its value is the amount of the diffuse component in global irradiation.

The equations that relate the clearness index to the diffuse fraction depend on the place and time scale involved. There are many relationships proposed depending on the area of the study deMiguel2001, Gopinathan1995. For daily time-scales, the equations that relate the clearness index and the diffuse fraction proposed by [Aguiar1992] are used in our case.

Three different tracking types are considered for the photovoltaic generator and the equations describing movements and relative position between the system and the sun are described in [Perpinan2009]

1. **Fixed panels** with an optimum angle of inclination that depends on the latitude.
2. **One axis trackers**, with a generator rotating on an axis oriented North-South. We will refer to them as “**one**”.
3. **Two-axes** tracking system that allows variation of the azimuth and inclination angles, we will refer to them as “**two**”.

To obtain irradiation components in the tilted surface at daily time scale, it is first necessary to estimate the irradiance profile using empirical relationships Collares-Pereira1979. Irradiance profile is then transformed into its components at the tilted surface and integrated in time to obtain the energy reaching the generator surface.

Direct irradiance can be transformed to the tilted surface using only geometrical criteria whereas the diffuse fraction is obtained with the model proposed by [hay1985estimating]. This model considers an approximation where the sky sphere is seen by the generator as isotropic except for the circumsolar region, which is considered to emit direct irradiance. The albedo component is considered as isotropic, due to the fact that its contribution to global irradiance is low.

Finally, we apply equations from [Martin2001] to obtain  $G_{eff}(\alpha, \beta)$ . It includes optical losses due to the fact that, except for the two-axis tracking system, the incident irradiation deviates from the normal of the generator. Also, transmittance losses are included for accumulated dust over the surface, considering a “moderate dust degree” in the terms used in the referenced paper [Martin2001]. In this case, any spatial distinction is considered and the same coefficient values are used to calculate angular and transmittance losses.

2) The second step is the transformation into power output, depending on the electrical characteristics of the components in the photovoltaic system and second order effects like temperature. A summary of the modules and the inverter model can be found in table 6.2. Detailed information about this step can be found in [Perpinan2009].

Ambient temperature is needed since the efficiency of the cells decreases with the increase of temperature, assuming a linear relationship. In this work we use the equation of [Crook2011], to calculate the daytime temperature from the daily maximum and minimum temperature, obtaining a daily profile.

Computation of the photovoltaic energy model is carried out with the `solaR` package Lamigueiro2012. This package implements in R a set of functions that include the sun apparent movement equations, the solar irradiation and irradiance decomposition, transposition models, and the PV generator and inverter models.

### 6.2.3 Datasets for evaluation

#### Satellite product: CM-SAF

CM-SAF (Climate Monitoring Satellite Application Facility) Schulz2009 was created as part of EUMETSAT Satellite Application Facility (SAF) when the importance of “contributing to the operational monitoring of the climate and the detection of global climatic changes” was recognized. Besides its operational products, they provide Climate Data Records generating long-term data, which is a valuable product for climate variability studies. In this work, the SARAH Muller2015 dataset for daily shortwave solar radiation, SSR, is used with an horizontal resolution of 0.44° to be consistent with the model simulations and for the period 2003 and 2009. For this product, the 85% of absolute differences with shortwave solar radiation measurements is below 10 W/m<sup>2</sup> for monthly values and 13 W/m<sup>2</sup> for daily means.

Concerning aerosols representation, the satellite dataset includes information from MACC Benedetti2009, Morcrette2009, provided by the European Centre for Medium-

## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

---

Element	Method
PV generator	Identical modules with $dV_{oc}/dT_c = 0,475 \frac{\%}{\text{C}}$ and $NOCT = 47^\circ\text{C}$ . The MPP point calculated as in [garcia2005caracterizacion]).
Inverter	Efficiency equation proposed in [jantsch1992results]: $\eta_{inv} = \frac{p_o}{p_o + k_0^o + k_1^o p_o + k_2^o p_o^2} \quad (6.1)$ where $p_o = P_{ac}/P_{inv}$ is the normalized output power of the inverter. The characteristic coefficients of the inverters are: $k_0^o = 0.01$ , $k_1^o = 0.025$ , $k_2^o = 0.05$ .
Other losses	<ul style="list-style-type: none"> <li>Average tolerance of the set of modules, 3%.</li> <li>Module parameter dispersion losses, 2%.</li> <li>Joule losses due to the wiring, 1.5%.</li> <li>Average error of the MPP algorithm of the inverter, 1%.</li> <li>Losses due to the MV transformer, 1%.</li> <li>Losses due to stops of the system, 0.5%.</li> </ul>

TABLE 6.2: *Calculation procedure for the estimation of energy produced by a PV system from daily global horizontal irradiation data. Left column represents the element of the PV system and the right column the equations and methods used in each case for the efficiency of the elements.*

---

Range Weather Forecasts (ECMWF). Monthly long-term means of a 0.5x0.5 degrees grid are spatially interpolated to assign the values of each pixel.

### **BSRN stations**

The BSRN (Baseline surface radiation network) Ohmura1998 is a set of ground-based measurements, from the WRCP Radiative Fluxes Working Group. It was created for supporting research in the radiation budget of the Earth-atmosphere and the radiation distribution, considering its main role in climate processes. The main objective is to provide high-quality measurements that are able to deal with the scarcity of the existing radiometric network. Although there are not many stations available for the period of the study in the Euro-Mediterranean domain, the high quality data provided by BSRN is necessary for a first evaluation of the climate simulations.

The three stations of the BSRN network used in this study (Payerne, Carpentras and Sede Boker) covered the period 2003-2009 and provide SSR monthly data. Their location is represented in figure 6.1 together with the PV plants considered in the study.

### **Temperature data: ECAD**

The PV production assessment calculated using the SSR from satellite data needs also the temperature for the performance of cells inside the module. The gridded E-OBS data set from the EU-FP6 project ENSEMBLES Haylock2008 is used in the energy production model at daily resolution. Mean, maximum and minimum temperature from the dataset, in a spatial resolution of 0.25°, are interpolated to the same grid of the climate model.

### **PV production data**

Data from two different power plants are used for the evaluation of the simulated PV power and the assessment of the added value of the aerosol inclusion in the climate simulations.

The two power plants are in the Iberian Peninsula and their location represented in figure 6.1. The first one, located in Tarragona in the North-East area, is a PV system with fixed structure. The second one is a two-axis tracking PV plant located in Seville, in the South of Spain. Details of both PV power plants are in table 6.3, including the electrical characteristics of their components.

Data of PV production are difficult to obtain due to confidentiality contracts. Moreover, when data are available time series are not always complete. Maintenance, modules substitutions, inverter problems and other stops in the production may lead to common time-gaps in the datasets. These limitations must be taken into account when establishing statistical comparisons between models and real data. In this case,

## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

---

	Seville	Tarragona
Type	two-axes	fixed
Generator	$P_g = 27.31 \text{ kWp}$	$P_g = 100.18 \text{ kWp}$
	$N_{mp} = 12$	$N_{mp} = 27$
	$N_{ms} = 11$	$N_{ms} = 35$
Inverter	$P_{inv} = 25 \text{ kW}$	$P_{inv} = 100 \text{ kWp}$
	$V_{min} = 405 \text{ V}$	$V_{min} = 450 \text{ V}$

TABLE 6.3: *Summary of the electrical components of the two photovoltaic plants, including generator characteristics (generator power  $P_g$ , and modules in parallel and serie,  $N_{mp}$  and  $N_{ms}$ ) and the inverter characteristics (power of the inverter  $P_{inv}$  and the voltage  $V_{min}$ ).*

---

the two PV power plants provide daily data within the period 2003-2009, the details are in table 6.4. Monthly means of these data are compared against the monthly means of simulated daily PV productivity, energy produced by the power installed  $\frac{\text{kWh}}{\text{kWp}}$ , with the models and the satellite. Only months with more than 15 days of data available are considered for the monthly mean and compared against simulated data.

## 6.3 Results

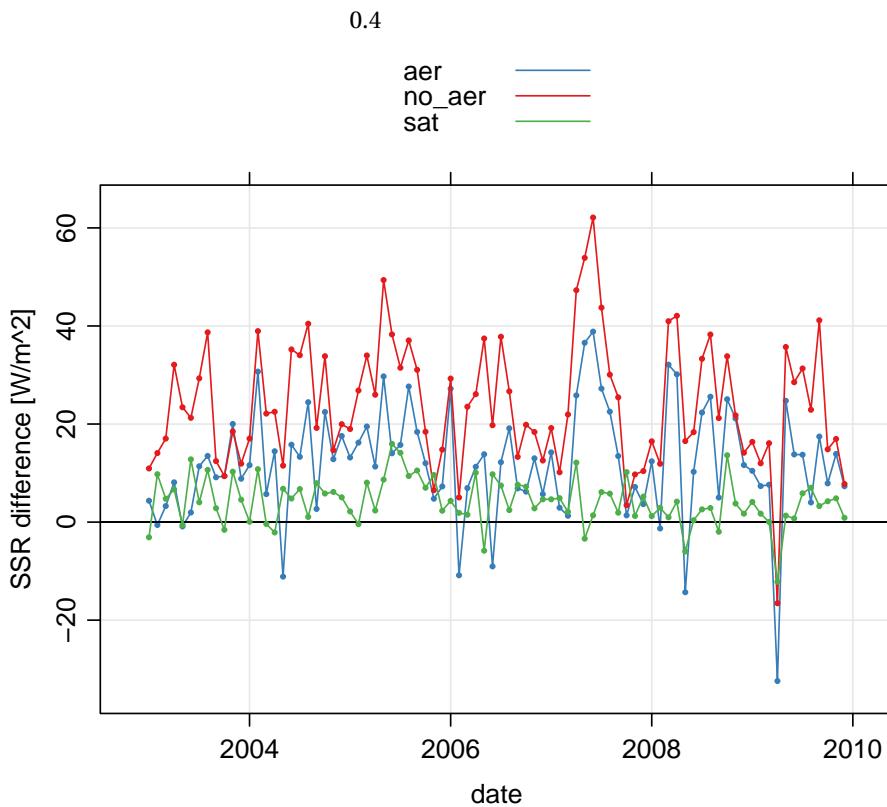
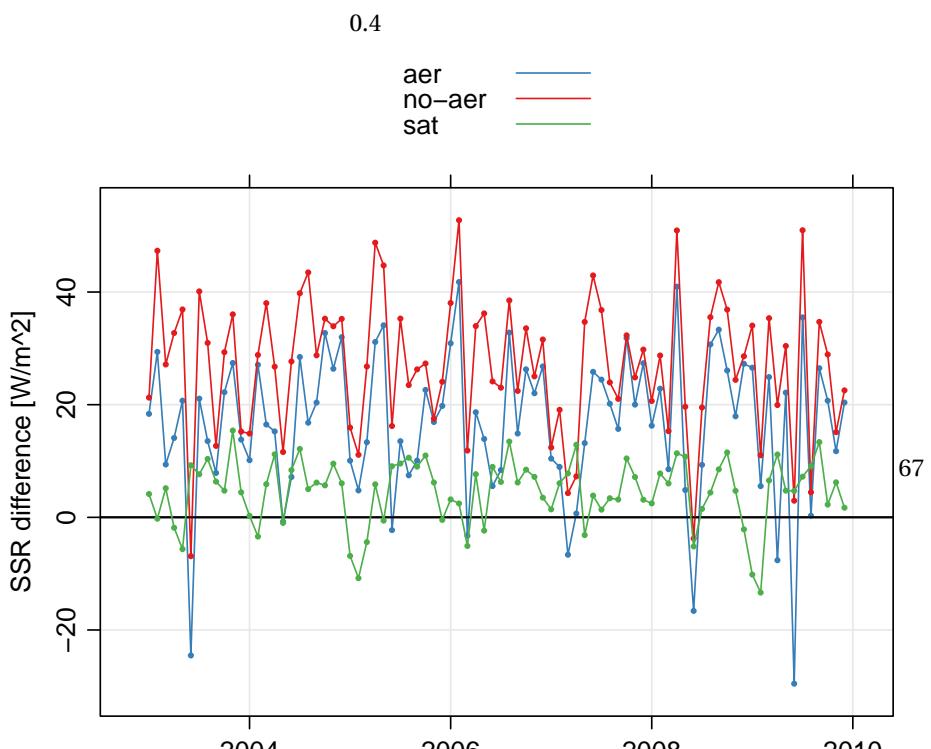
Although the AOD dataset and the climate simulations used in this study have already been evaluated against observations and satellite datasets in [Nabat2013; Nabat2014; Nabat2014a], an additional assessment of the SSR form the climate simulations is also made in this work against the CM-SAF SARAH dataset Muller2015 for the period 2003-2009 and against some BSRN Ohmura1998 stations at a local scale.

### 6.3.1 Local scale

Three different stations from BSRN cover the period between 2003-2009 with SSR monthly data. Stations are represented in figure 6.1 with points. Also, two different power plants in Spain, represented in figure 6.1 with a cross, are used to evaluate PV power simulations at a local scale. A summary of the data used in this section, the periods and the resolution can be found in table 6.4.

In figure 6.5 differences in monthly time series of SSR from simulations and satellite data and the stations measurements are represented.

For the three stations, the satellite has better results than the simulations, although the error magnitude varies depending on the station. There is also a general improvement of the AER simulation against the NO-AER. Carpentras station is the

FIGURE 6.2: *Carpentras*

## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

Data from	Seville	Tarragona	Payerne	Sede Boker	Carpentras
Variable	PV productivity	PV productivity	SSR	SSR	SSR
Time res.	day	day	month	month	month
	518 daily values	300 daily values			
Period	between: 02-07-2007 30-11-2008	between: 01-01-2003 19-03-2005	2003-2009	2003-2009	2003-2009
% no data	0	8 %	0	10.71 %	0

TABLE 6.4: *Summary of the local data of PV power plants and SSR from BSRN stations used for the evaluation of the simulations. SSR is evaluated as input of the PV model and the PV output as the result of the whole modeling process.*

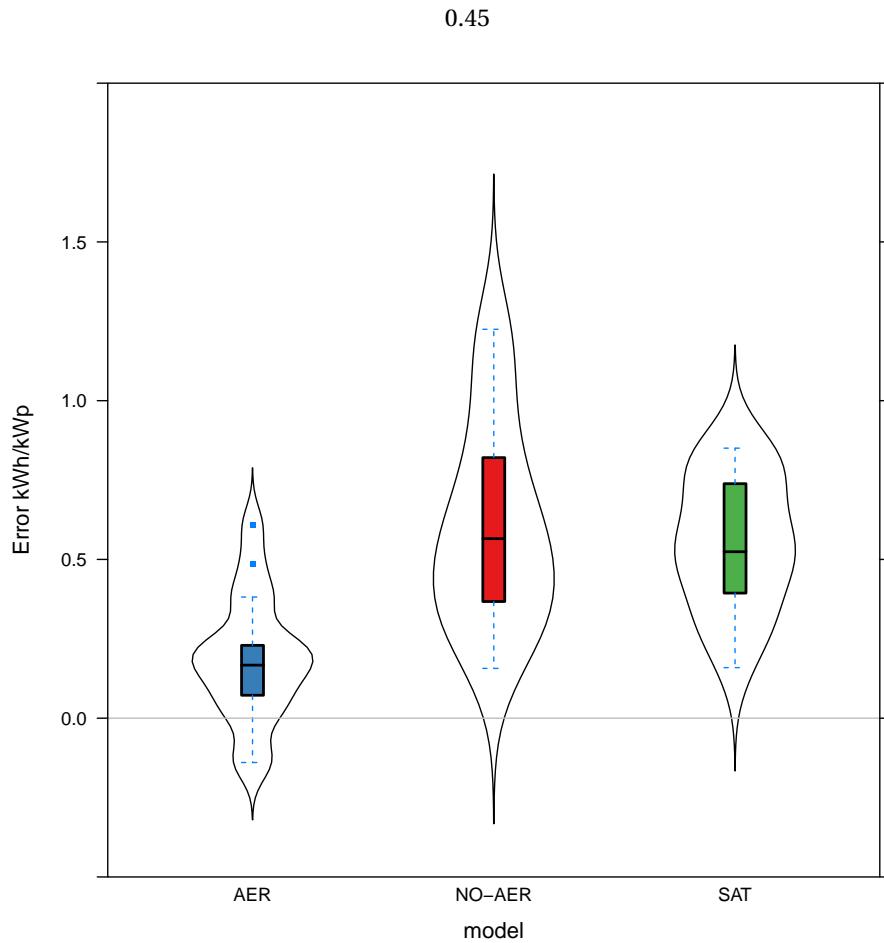
one with lower bias with respect to the observed data 6.2. In the case of the Sede Boker station 6.4, there is a seasonal bias in the AER simulation. High negative differences appear from May to August, showing underestimation of the SSR in these months.

Several statistical measurements summarize the general performance of the SSR from the climate model and the PV simulations at each location and are reported in table 6.5.

For the simulation of PV power output, the daily mean of PV production data averaged for each month is compared with simulated PV energy production at each power plant, using the three possibilities of SSR data as input: AER, NO-AER and SAT. For these simulations, in order to help with the visualization of the results a *violin plot* is used to visualize the absolute error and its distribution. For the Seville PV power plant, AER performs better than NO-AER and SAT simulations showing lower errors 6.6. Besides, only AER has some negative errors, which means underestimation for some months, whereas the satellite and the NO-AER have only positive error values.

The median error for the AER simulation is less than  $0.25 \frac{\text{kWh}}{\text{kWp}}$ . Differences are concentrated around this value, which makes the distribution peaks around it in a narrow shape, although the range is wider due to higher values above  $0.5 \frac{\text{kWh}}{\text{kWp}}$  that spread the distribution.

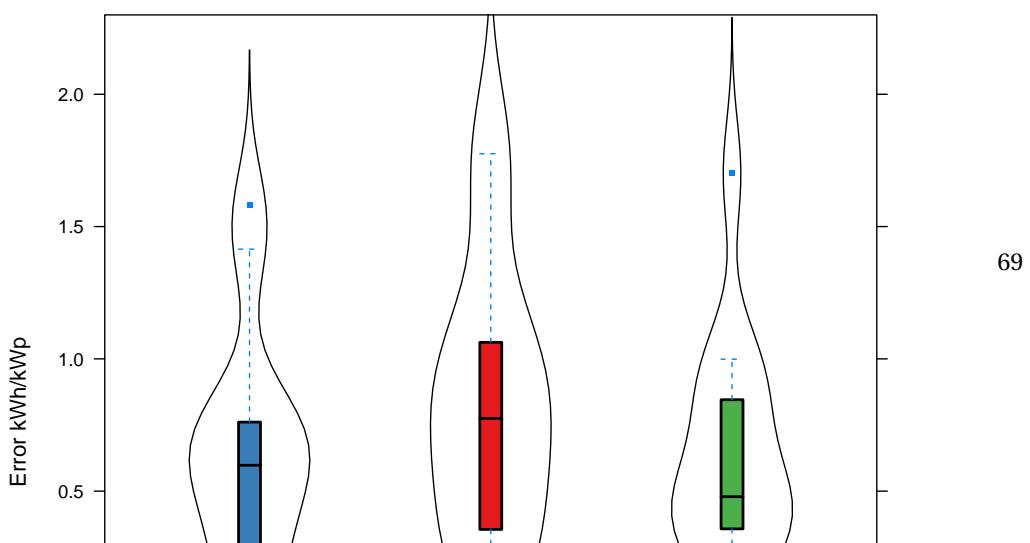
NO-AER simulation has a wider range of errors than AER and SAT, and a wider distribution. The satellite presents a median error close to  $0.5 \frac{\text{kWh}}{\text{kWp}}$ , as it could be expected from the evaluation and report of the CM-SAF dataset.



---

FIGURE 6.6: *Seville*

0.45



## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

---

Location	Simulation	RMSE	MBE	cor	sd
Seville	AER	0.27	0.18	0.98	1.34
	NO-AER	0.67	0.60	0.95	1.29
	SAT	0.59	0.55	0.98	1.45
Tarragona	AER	0.77	0.61	0.87	1.21
	NO-AER	0.96	0.82	0.9	1.29
	SAT	0.76	0.64	0.88	1.13
Payerne	AER	21.21	16.62	0.97	77.4
	NO-AER	29.70	27.07	0.98	81.88
	SAT	7.36	4.6	0.99	83.29
Carpentras	AER	16.59	12.05	0.98	90.05
	NO-AER	27.26	24.10	0.99	90.57
	SAT	6.38	4.26	1.00	88.71
Sede Boker	AER	18.89	8.17	0.98	62.83
	NO-AER	37.42	35.63	0.98	76.06
	SAT	12.00	10.27	0.99	77.62

---

TABLE 6.5: *Values for the root mean squared error (RMSE), mean bias error (MBE), temporal correlation (cor) and standard deviation (sd) for the simulated PV production in Seville power plant and Tarragona power plant compared with the final productivity data measured  $\frac{\text{kWh}}{\text{kWp}}$  and the SSR from the climate model simulations and the satellite in comparison with the SSR from BSRN stations  $\frac{\text{W}}{\text{m}^2}$*

---

The Tarragona PV power plant presents larger errors. The SAT has slightly better performance than AER and the improvement with respect to NO-AER can be observed in both cases.

AER simulation's median error is  $0.60 \frac{\text{kWh}}{\text{kWp}}$  and NO-AER simulation error has larger values, the median is  $0.77 \frac{\text{kWh}}{\text{kWp}}$ . The SAT error has lower median value:  $0.48 \frac{\text{kWh}}{\text{kWp}}$ .

It can be also appreciated in figure 6.7 that there are two months in the simulations where errors are specially large. As these values are clear outliers, it might be that some maintenance activities in the plant are the cause of the lower energy output than expected, which would explain the larger errors in those months.

The added value of including aerosols in climate simulations for the estimation of PV production is clearly illustrated by our results. The model using climate simulations with a good aerosols representation could be locally as good as the satellite, at least at the location of the PV power plants used (despite of its biases in some areas).

The satellite product has revealed itself also as a good dataset, either for SSR (seen in the evaluation with BSRN stations) and for the PV production estimation.

### 6.3.2 Regional scale

The regions defined in figure 6.1 are the same areas selected in [Nabat2014], and the acronyms used in this section are also in the figure. The difference between simulations and satellite data in annual terms is represented in figure 6.9. The improvement of the simulation including aerosols is noticeable for every region. It can be appreciated also that there are some weak changes in the interannual variability of the bias due to aerosols inclusion, like in the BISL area, AFRE region in 2009 or AFRW.

Western (EURW) and Southern Europe (EURS) have the lowest biases in SSR, as well as eastern Africa (AFRE) and the South of British Islands (BISL). Negative biases only appear for Africa and they could be due to the fact that in some cases the annual amount of aerosols is overestimated, or due to the optical properties of the aerosols in the model. The rest of the areas have positive biases, probably due to an underestimation of the cloud cover in the model Nabat2014. The east side of the domain shows a higher bias (NEEUR), although it is clear that the addition of aerosols improves the representation of SSR.

In order to calculate the impact on photovoltaic production, shortwave solar irradiation, SSR, from the climate model simulations and the satellite dataset is used as an input of the photovoltaic model. For a fixed typology of the panels, the difference between simulations with respect to the satellite dataset as input, in the mean of daily productivity for each month, is represented in figure 6.10. The productivity is defined as the energy production per unit of power installed kW<sub>p</sub>.

In general, The NO-AER simulation gives more production than the AER and the satellite, overestimating the PV power potential. Climate model simulations, AER and NO-AER, differ more from satellite PV production output in winter months, whereas in summer months, AER and satellite are close to each other, except for the EURNE region. This area is the most biased area of the model in comparison with the satellite, perhaps due to a bias in the cloud cover.

The regions from North Africa, AFRE and AFRW, are slightly different from the rest with a roughly constant bias curve of the the annual cycle, therefore the difference in PV production between summer and winter months is lower due to the mean latitude of the area. Besides, for these areas, the AER simulations give in summer month lower PV production values than the SAT and the NO-AER, which is not the case in the rest of the regions.

The EURS area is the one with largest amplitude between winter and summer months. For April and May, the AER simulations has small difference with the satellite simulation.

## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

---

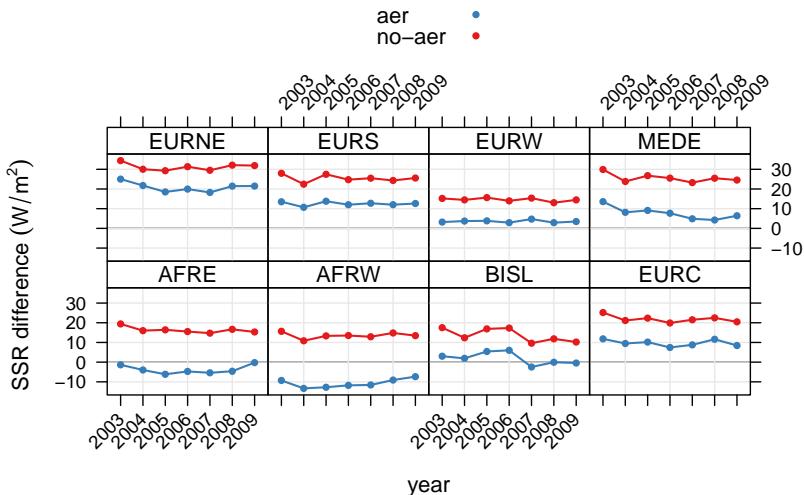


FIGURE 6.9: *Difference in shortwave solar irradiation,  $\text{SSR} \frac{\text{W}}{\text{m}^2}$ , between simulations and the satellite aggregated by areas defined in figure 6.1 in the period 2003-2009.*

### 6.3.3 Impact of aerosols by tracking type

#### Mean behavior. Period 2003-2009

In addition to the absolute difference in PV productivity between AER and NO-AER, the relative difference between simulations is presented in this section, showing the relative impact in each case. It allows us to contextualize the loss with respect of the potential of the place.

The difference in yearly PV production between both simulations, for every type of tracking panel, is represented in figure 7.4. The spatial pattern shows areas where the aerosols affect more the shortwave solar radiation. Central Europe, Po Valley and the South of the domain, along the African continent coast, are the areas where the differences are more noticeable.

For the fixed panels, the differences are around  $-150 \frac{\text{kWh}}{\text{kWp}}$  for some of these areas in annual terms and reaching  $-200 \frac{\text{kWh}}{\text{kWp}}$  in the Po Valley, Syria, Iraq and between Algeria and Tunisia.

When the two other tracking systems are considered, the difference between both simulations increases, with values of  $-300 \frac{\text{kWh}}{\text{kWp}}$  for the two-axes type over the most affected areas and even higher values like in the south of Turkey. These results are

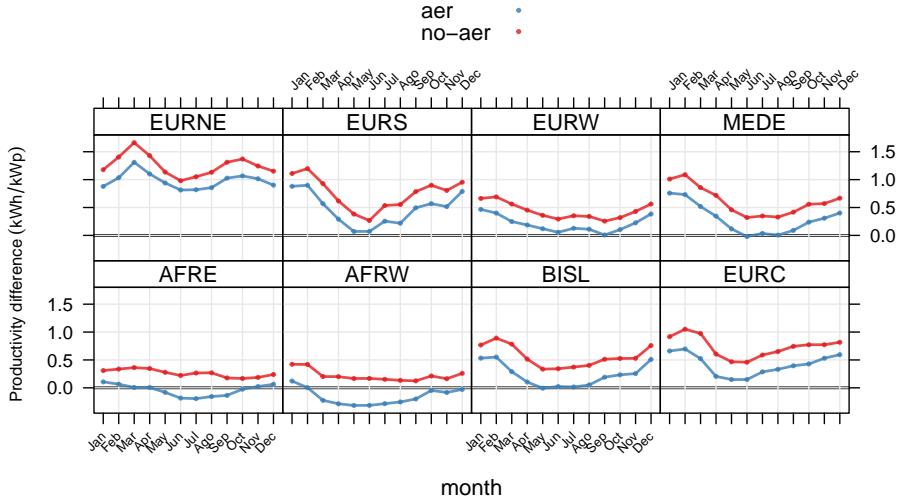


FIGURE 6.10: Annual cycle of daily energy productivity  $\frac{\text{kWh}}{\text{kWp}}$  differences by area for the AER simulation and NO-AER simulation with respect to the satellite as inputs of the PV model, considering fixed panels.

consistent due to the fact that the two-axes and the one-axis tracking systems are more efficient systems to give energy to the generator.

The results of the country averages for the annual productivity are shown in figure 6.12. It is shown that the aerosols impact range from -4% to the PV production to -16%. It can be seen that from the fixed panels to the one-axis, and then to the two-axes tracking type, there is an increase in the productivity losses that is more noticeable in Central Europe, like in Belgium-The Netherlands.

Germany, that have installed a high amount of PV capacity, is affected with values around -10% of loss for fixed system and more than -12% for the one-axis and two-axes panels. Thus, some countries with high PV production have moderate losses due to aerosols in relative terms. For countries in the west and south of the domain like Portugal, Spain, Morocco, Jordanian or the extension of Saudi Arabia included in the domain, the relative amount of energy loss is smaller due to its high potential, with differences between -5% and -6% for fixed panels and reaching -8% in some cases for the other types of tracking systems.

## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

---

1

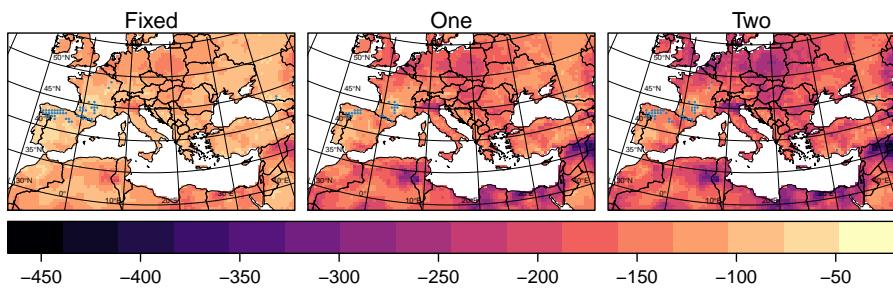


FIGURE 6.11:

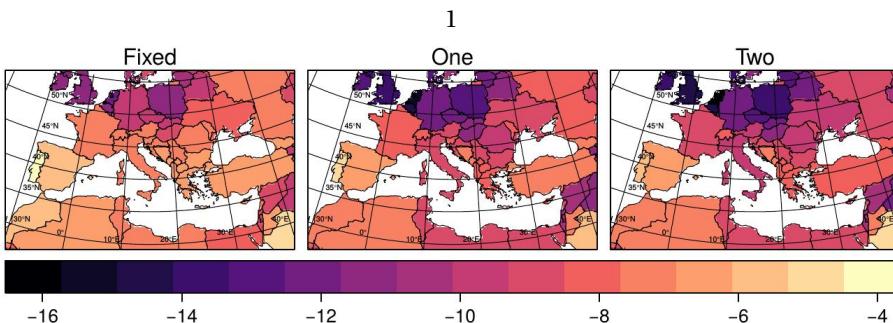


FIGURE 6.12:

FIGURE 6.13: Differences in PV yearly productivity: (a) absolute  $\frac{\text{kWh}}{\text{kWp}}$ , (b) relative averaged by country (%); for the period 2003-2009 between AER and NO-AER and for the three different types of tracking. For the non-significant differences, calculated with a 't-test', a point is overplotted for (a)

### Seasonal cycle

In seasonal terms spring (MAM) and summer months (JJA) show statistically significant values of the PV productivity difference. For winter (DJF) and Autumn (SON), there are few significant areas, mostly located at the south of the domain, in the African continent. An exception can be found for fixed panels in winter, with significant zones in Europe with values above  $-10\%$ , like the Po Valley or British Islands.

The spatial pattern for the spring season shows higher values in Central-Europe, in specific countries like Poland, Belgium and The Netherlands, or British Islands, and in northern parts of Syria and Iraq. Maximum values in these areas, range from an impact higher than  $-10\%$  for fixed panels to around  $-15\%$  for the two axes tracking type. Almost the whole African part of the domain shows significant differences, with impact values above  $-5\%$  and local values reaching more than  $-10\%$  between Algeria and Tunisia and in areas of Libya or Egypt.

Summer months have higher loads of aerosols over the Mediterranean. For fixed panels, only a few areas in north Africa, Middle East and north-western Europe show values with a loss higher than  $-10\%$ . One-axis tracking enlarge the extension of areas with a decrease higher than  $10\%$  and some areas with more than  $-15\%$  appear in Belgium and The Netherlands, Algeria and Syria-Iraq border. Also some smaller areas appear with PV production losses above  $-15\%$  within the above mentioned areas. Between the one-axis and the two-axes tracking type, there are no substantial differences in the spatial pattern but maximum values reach in these cases  $-17\%$  to  $-20\%$  in the same areas.

Significance is not clearly linked to the magnitude of the relative difference in productivity. High relative differences in DJF and SON in central and northern Europe are not significant, due to the high climate variability in these areas and seasons, whereas lower relative differences in spring in Africa and in summer in southern and eastern Europe are mostly significant. The fact that JJA changes are significant over most of the domain explains the statistical significance of most yearly differences (figure 7.4), due to the high contribution of summer to the yearly production and yearly interannual variability (see e.g. [Gil2015]).

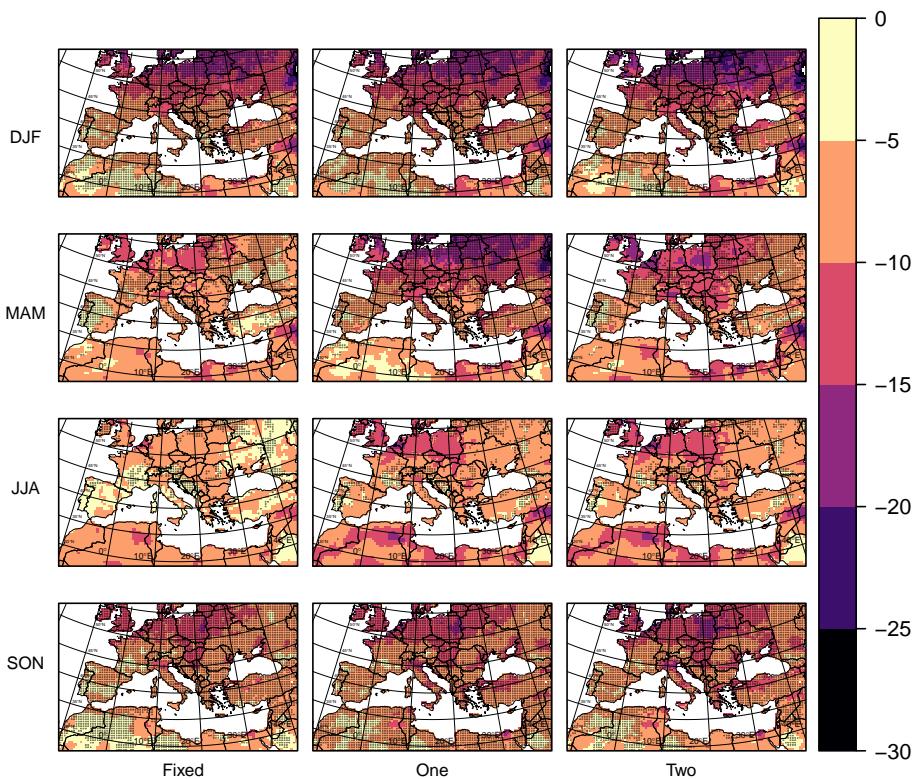
#### 6.3.4 Long-term trends. Period 1980-2012

The impact of long term trends of solar shortwave on PV production can be illustrated with the results of the simulations of the brightening period over Europe. The simulation including the aerosols dataset, NAB13, and the decreasing trend in sulfur aerosols, TREND, is able to reproduce the observed positive trend in SSR Nabat2014.

As a compromise between the lifetime of a PV plant and the length of the simulation, two 15-year periods (at the beginning and end of the simulated period) are evaluated as a proxy for a PV project and the potential amount of energy that can be obtained during such a project. A fixed system is selected for the panels.

## 6. IMPACT OF AEROSOLS IN PHOTOVOLTAIC ENERGY PRODUCTION OVER THE EURO-MEDITERRANEAN AREA

---



---

FIGURE 6.14: *Relative difference (%) in seasonal PV productivity between both simulation for all type of panels. For the non-significant absolute differences, calculated with a 't-test', a dot is overplotted.*

The relative difference between the accumulated energy obtained by a 15-year project between 1997-2012 and energy obtained by a 15-year project between 1980-1995 can be seen in figure 6.15.

In Central-Europe the differences in the energy obtained are higher, due to the fact that it is the region where anthropogenic aerosols decrease more. Accumulated over 15 years, an increase higher than 2000  $\frac{\text{kWh}}{\text{kWp}}$  is found for this area. It means that about 10-14 % more energy can be produced in the lifetime of a PV plant at the end of the period.

These results highlight the impact that environmental policies could have on the PV energy production, showing that anthropogenic aerosols are able to reduce potential PV power of a project significantly. It also illustrates that future evolution of regional anthropogenic aerosols load is likely to influence expected PV production on a given site on the lifetime of a PV plant.

## 6.4 Discussion

### 6.4.1 Limitations

The coarse resolution of climate models and their bias in some variables, especially in cloud cover, makes difficult their application for solar resource assessment at local scale in most areas. However, the RCM will evolve to finer resolution and besides, the use of climate models is mandatory for taking into account future climate evolution. They are also relevant tools to perform sensitivity tests allowing to disentangle the driving factors of the resource variability such as aerosols.

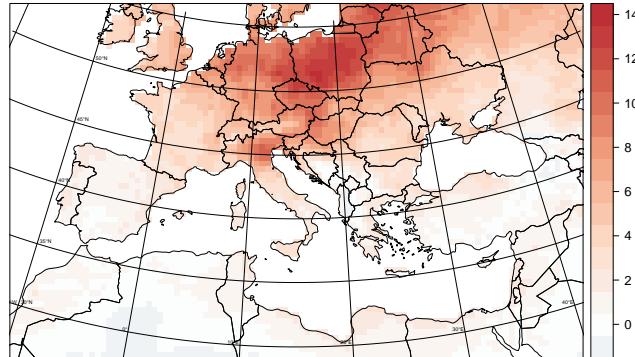
A multi-model approach would be necessary to obtain a more robust answer applying the same methodology to more models but, up to now, aerosols are poorly considered in many RCM, which does not allow that type of study considering actual simulations.

For the limitations in the PV model, it is important to notice that in the decomposition of SSR and the transposition to the tilted panels, some empirical relationships are used. If components of solar radiation were an output of the RCM, the additional steps for decomposition could be avoided.

It could also be pointed that the spectral decomposition of the SSR reaching the panels would give a better input in order to calculate spectral losses of the PV modules, although for periods longer than a day the spectral effects become less significant Martin1999.

In the optical losses, deposition of dust over the generator surface is approximated in the assessment but could be underestimated in desert areas because it is not spatially modeled. That could mean a higher drop in transmittance and final energy production.

The time-scale is also important for the AOD representation. Several processes involving aerosols in the atmosphere are in day to weeks time-scales. We have used



---

FIGURE 6.15: *Relative difference in PV yearly productivity [%] accumulated for a 15-year period at the end and the beginning of the period: 1980-1994/1998-2012*

a realistic interannual dataset of the AOD that improves the state-of the art used in climate modeling but next studies should go further with an improvement in the aerosols representation. This will include a prognostic scheme of aerosols that allows to study finer time-scales and future scenarios, through a fully-coupled and fully-interactive aerosol-climate model.

#### 6.4.2 Implications for the climate services dedicated to the energy sector

The results show the necessity of considering aerosols in climate simulations used to deliver energy-related climate services. The inclusion of spatio-temporal variability of aerosols in RCMs may change the current estimates of future PV production over Europe Jerez2015. An accurate ensemble of models is essential for bridging the gap between services providers and potential users considering that some of the discrepancies between model simulations could come from the aerosols inclusion.

#### 6.4.3 PV production data issue

The scarcity of real power data at PV plant sites is an important issue that has to be overcome in order to improve research in the fields of PV forecasting, climate ser-

vices or energy modeling. The potential synergies between research institutions and different stakeholders of the energy sector will enhance the PV integration, the management and planning operations as well as the efficiency of the overall system and its development. Not only production data is needed, but also, accurate metadata will be extremely important in order to integrate into the modeling chain important factors such as: days of maintenance activities, cleaning-panel days, installed capacity, electrical characteristics of the PV power plant, among others.

## 6.5 Conclusions

Two main questions are addressed in this paper: first, the evaluation of the capacity of an RCM to produce reliable estimations of PV production over the Mediterranean area. Secondly, the role of aerosols in PV production using sensitivity tests performed with the RCM.

It is demonstrated that the use of a RCM as input of a photovoltaic production model is able to reproduce real PV data accurately at monthly time scales for two locations at the domain. Besides, the added value of including aerosols in the simulations is observed over the whole area as the simulation with aerosols shows less bias in SSR than the simulations without aerosols and it is close to the simulated PV using the satellite dataset across the whole area.

The results show that the most impacted areas are (with some exceptions) mostly in central Europe, where the lower resource amount in combination with the influence of aerosols gives a significant reduction in potential electricity production. For the annual averaged by country productivity, relative differences are around -12% for fixed typologies and are seen over Central Europe (Poland and The Netherlands). Differences increase from one axis typology to the two-axes, reaching around -16% between both simulations in Belgium and around -13% and -14% in many countries. In seasonal terms, the loss can reach values of -20% in some areas for summer months.

In the multi-decadal simulation 1980-2012 a noticeable increase in productivity has been obtained in central Europe as a result of the decreasing trend of anthropogenic aerosols observed from the end of the eighties. This trend has been associated with pollution control measures as well as economic crisis in Western Europe. This result has implications beyond the domain of this study: highly polluted countries like India and China could obtain an increase in PV productivity if pollution control policies are effectively implemented.

The non-negligible impact of aerosols on PV production in the area suggests that the inclusion of aerosols in future scenarios is necessary for solar energy assessment.



## **Future projections of solar resource for photovoltaic applications**



---

## Abstract

With the ongoing energy transition, the evolution of renewable energy resources under different climate change scenarios is key for the investors and stakeholders of the energy industry. The possibility of changes in the actual conditions for the operating plants and the projected resources can vary the financial frame of the projects.

Although climate models give a robust answer in terms of global warming and other important climate variables, they disagree in the projected changes of solar resource over Europe. Whereas global climate models, GCM, present a clear positive signal around the mid of XXI century, regional climate models, some studies show a negative anomaly for the same period in RCMs.

In this chapter we try to explain the reasons behind the different behaviour focusing on the representation of aerosols in the scenario simulations. We use regional climate models from the EURO-CORDEX project and we focus on the RCP8.5.

A fairly pairwise comparison show that regional climate models that include evolving aerosols for the scenarios reverse the sign of the signal in the shortwave solar radiation over Europe.

We analyse total cloud cover in the regional models simulations and it is not find a clear relationship between the models with aerosols and the anomaly in cloud cover.

Due to the clear relationship between shortwave solar radiation anomaly and aerosols evolution, it is necessary to prepare an experiment that is able to give a robust answer in the role of aerosols in terms of shortwave solar radiation projections, which means, that is able to narrow down the uncertainties.

In terms of photovoltaic (PV) potential it has been shown that for RCM with evolving aerosols countries in Central-Europe has a positive anomaly, in contrast with the rest of the RCMs projections. Those RCMs projections including evolving aerosols agree with GCM projections. This result reveals the impact that the representation of aerosols has in the projections of PV potential for the climate services.

## 7.1 Introduction

The generalized increase in the photovoltaic installed capacity in last decades demands the detailed study of spatio-temporal features of solar resource.

Due to the link between solar energy production and atmospheric variables, there is an increasing concern motivated by the availability of resources under climate change scenarios. Due to that, climate modelling is a key tool to evaluate future energy potential despite of some constraints like its low spatial resolution or cloudiness representation.

Although variability of solar radiation are mainly due to changes in cloudiness, in clear days other constituents of the atmosphere, mostly aerosols, decrease the amount of energy reaching the generator surface[ref]. Because of its geographical situation, the Euro-Mediterranean area is one of the most influenced areas by natural and anthropogenic aerosols coming from different sources, affecting the spatio-temporal distribution of solar resource.[ref]

Different CMIP5 simulations with different GCMs have been evaluated to assess the photovoltaic potential under climate change conditions (wild et al.) and projecting an increase over Europe. However, later studies using regional climate models (Jerez et al. Bartok) have shown an opposite behaviour, an overall decrease in short-wave solar radiation and photovoltaic potential in the same area.

The added value of regional climate modeling against global simulations lies on the better representation of local features that can not be solved with coarser models. However, the increase in resolution has led to a simplification of other processes in order to not compromise the computational time. Many regional climate simulations have been done using a very simplistic representation of aerosols content (nabat) and they do not include an evolution of aerosols in future projections.

In this work we analyse the projections of photovoltaic potential over Europe for the RCP8.5 scenario using regional climate models simulations from EURO-CORDEX initiative. We classify different simulations attending to the aerosol representation in the model and its driving global climate model.

Section ?? describes the regional models and simulations used in the study and in section ?? there is a description of the datasets used by them. In section 3 the main results are explained and a discussion section followed.

## 7.2 Climate data

### 7.2.1 EURO-CORDEX

The EURO-CORDEX (Jacob et. al 2014) initiative started as a 'branch' from the Coordinated Regional Climate Downscaling, CORDEX, whose aim is to provide regional climate simulations in different domains. EURO-CORDEX develops climate projections focused on the European continent at different horizontal resolutions ( $0.44^\circ$ ,  $0.11^\circ$ ). These simulations are driven by different CMIP5 GCMs (Taylor et. al 2012).

In table 1 there is the information about the aerosols datasets included in the EURO-CORDEX RCMs. It can be observed that only two RCMs include time-evolution of aerosols: RACMO22e and ALADIN5.3. Each of these RCMs have a different dataset of aerosols for projections. Whereas ALADIN5.3 select the Szopa dataset, RACMO22E uses the CAM inventory (Lamarqué et al.).

The choice for the different simulations in this chapter has followed the next 'pairwise' principle: firstly, RCMs simulations from the EURO-CORDEX database including time evolving scenarios of aerosols are selected. Then, a family of RCMs simulations driven by the same GCMs is constructed around it. Following these steps, we will obtain several groups of RCMs simulations where each one will be called 'family' with the same driven GCM and with only one simulation including aerosols scenarios. Besides, we consider only the RCP8.5 because changes in SSR will be more noticeable and differences between simulations easily detected.

### 7.2.2 Aerosols datasets

Different datasets are used for the RCMs simulations. Only ALADIN5.3 and RACMO22E include in their scenario runs evolving aerosols. For ALADIN5.3 the Szopa ?? No tengo

TABLE 7.1: *RCMs from EURO-CORDEX grouped by the CMIP5 GCMs drivers generating the 3 families of simulations studied.*

CMIP5 GCM	Institution	RCM
CNRM-CM5	CNRM	ALADIN5.3
	CLMcom	CCLM4-8-17
	SMHI	RCA4
ICHEC-EC-EARTH	KNMI	RACMO
	CLMcom	CCLM4-8-17
	SMHI	RCA4

claro si incluir esta sección. De hacerlo, ¿describo todos los dataset? ¿únicamente los que utilizan los modelos que tienen evolución temporal?

## 7.3 Methods

### 7.3.1 Spatial analysis

In order to analyse the spatial behaviour of shortwave solar radiation, SSR, it is calculated the mean yearly anomaly for the period 2021-2050 with respect to the reference period 1971-2000. The yearly anomaly is obtained only for summer months, June, July and August, JJA, which correspond to the season when the AOD (aerosols optical depth) is higher over Europe and the Mediterranean area.

The 2021-2050 period has been selected because of two reasons: in the mid-century the effects of global warming are not still very remarkable among the simulations, which could help to better detect the impact of including evolving aerosols. On the other hand, the PV power plants operating during 2021-2050 are the ones that will be planned shortly, so the results are important for the industry.

As an important driver of the variability of SSR the same procedure is applied to the total cloud cover variable, CLT. This allows us to know if there is correlation between both anomalies and how this correlation behaves for models including aerosols.

TABLE 7.2: RCMs from EURO-CORDEX and aerosols description.

Institution	RCM	description	classes	scenarios
CNRM	ALADIN5.3	Climatology. Szopa 2013	5 classes: sea salt, sulfate, black carbon, desert dust, organic carbon.	evolve with RCP
CLMcom	CCLM4-8- 17	Climatology. Tanré 1984	4 classes: sea, land, desert, urban.	no evolution
SMHI	RCA4	Parametrization in radiation fluxes	Single integrated class.	no evolution
KNMI	RACMO	CAM inventory, Lamarque 2010	6 classes: sulfate, organic matter, desert dust, sea salt, stratospheric aerosols, volcanic	evolve with RCP

### 7.3.2 PV potential

In order to obtain a projection of the photovoltaic productivity over Europe under climate change scenarios, a modeling chain approach is considered. Shortwave solar radiation from different climate simulations is used as an input of a photovoltaic model that will give an estimation of the power output. The complete process is explained in chapter 4.

The two steps followed in the modeling process of the photovoltaic output are: first, it is necessary to get solar irradiation that reach solar cells, after that the electrical performance of the system is modeled.

The process is described in 4. In this case, monthly means of daily solar irradiation are used for the decomposition and transformation to the plane of array. Correlation equations between de diffuse fraction and the clearness index described by Page are applied. After the decomposition, the components are transposed to the plane-of-array using equations 4.7 and 4.8.

## 7.4 Results

### 7.4.1 Anomalies of SSR and CLT

The annual mean anomaly of the period 2021-2050 for summer months (JJA) with respect to the reference period 1971-2000 is represented in figures 7.1 and 7.2 first for shortwave solar irradiation, SSR, and then for total cloud cover, CLT.

Annual mean anomaly for summer months shows an increase in Europe, more relevant in Central-Europe, for climate models with evolving aerosols in scenarios (ALADIN5.3 AND RACMO22E) although there are differences in the magnitude of the changes. ALADIN5.3 presents highest anomaly in the mentioned area, which correlates with the negative anomaly and spatial pattern of aod of this model, as can be seen in figure 7.3. For RACMO, the aod anomaly has similar spatial pattern but the magnitude is slightly smaller. The projected SSR anomalies in RACMO are positive for a large area among the domain although they do not reach the higher value of the ALADIN5.3. Mean values for the whole domain are positive for ALADIN ( $7.15 W/m^2$ ) and RACMO ( $2.60 W/m^2$ ) and negative for the others.

The rest of the RCMs from the first and second family presents a similar anomaly in SSR among them, with a slightly decrease in SSR with the exception of southern and western Europe that shows a small increase.

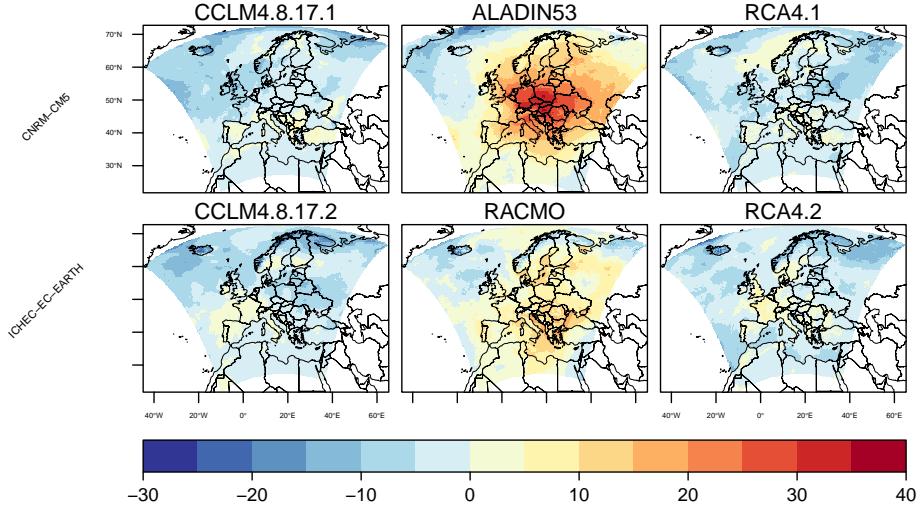


FIGURE 7.1:

With respect to the CLT anomalies it is noteworthy that they are not able to explain anomalies in SSR. For both families and the whole group of simulations the anomaly in CLT is close to zero as it can be seen in table ???. The spatial pattern of CLT anomaly in ALADIN53 shows an increase in the area where SSR anomaly is higher (Central Europe), which is remarkable. For ALADIN5.3 and RACMO22E the spatial correlation is very low between CLT and SSR, -0.26 and -0.15 respectively whereas it increases between SSR and AOD, -0.94 for ALADIN5.3 and -0.65 for RACMO22E.

\*\* Hablar de que la anomalía de nubosidad es parecida entre modelos independientemente de su GCM?

#### 7.4.2 Projected changes in PV production

The photovoltaic yearly productivity, defined as the power output by the power capacity installed (it is consider a 1 kW system for each grid cell), is calculated for each pixel of land in the domain of EURO-CORDEX. Then, the averaged by country of the

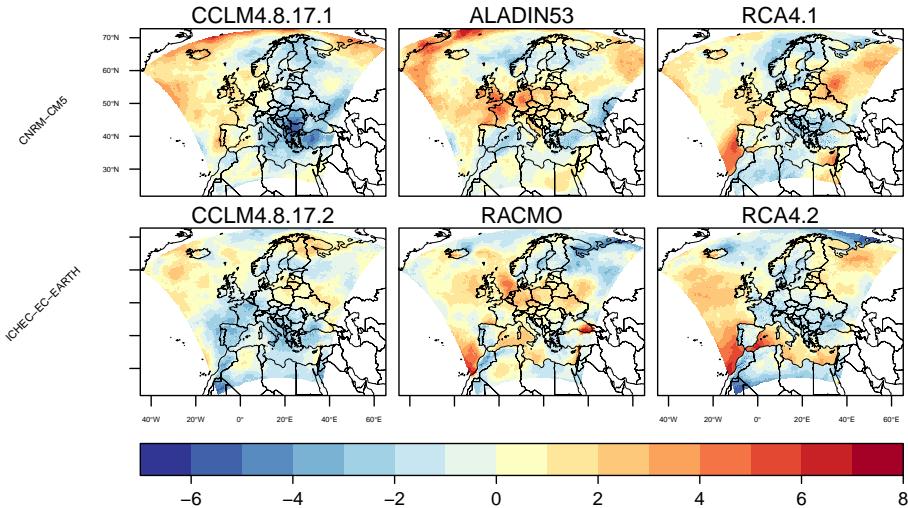


FIGURE 7.2:

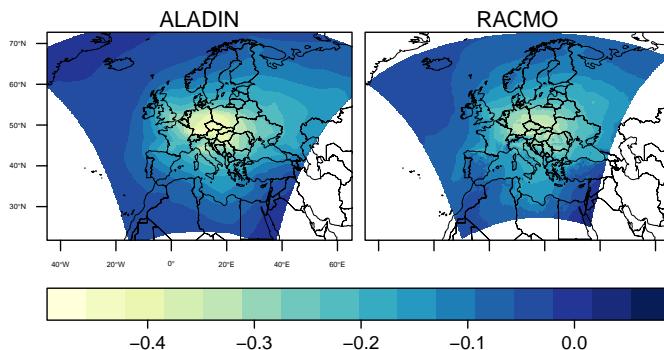


FIGURE 7.3:

TABLE 7.3: *Spatial correlation between SSR, CLT and AOD anomalies maps.*

CMIP5 GCM	RCM	$\rho_{SSR,CLT}$	$\rho_{SSR,AOD}$	$\Delta SSR$	$\Delta CLT$	$\Delta AOD$	$\Delta PV_{annual}$
CNRM-CM5	ALADIN5.3	-0.26	-0.94	7.15	0.60	-0.10	3.5%
	CCLM4-8-17	-0.83	-	-4.17	-0.15	-	-2.8%
	RCA4	-0.71	-	-3.55	0.21	-	-2.6%
ICHEC-EC-EARTH	RACMO	-0.15	-0.65	2.60	0.13	-0.11	-0.7%
	CCLM4-8-17	-0.80	-	-4.40	-0.51	-	-3.9%
	RCA4	-0.59	-	-3.44	0.09	-	-3.6%

relative difference with respect to the reference period 1971-2000 is represented in next figure.

#### Anomalía para los modelos con aerosoles y los modelos sin aerosoles

The geographical dependence of the PV anomaly is due to the spatial pattern of SSR anomaly, that is close related with the evolution of AOD in central Europe projected by ALADIN5.3 and RACMO22E. It is important to notice that for the ALADIN53 simulation including aerosols the sign in PV power output anomaly is positive and negative in the other cases. Besides, the RACMO anomaly is close to zero (although negative) whereas the rest of the simulation has the same magnitude order of the anomaly: -2 to -4%

## 7.5 Discussion

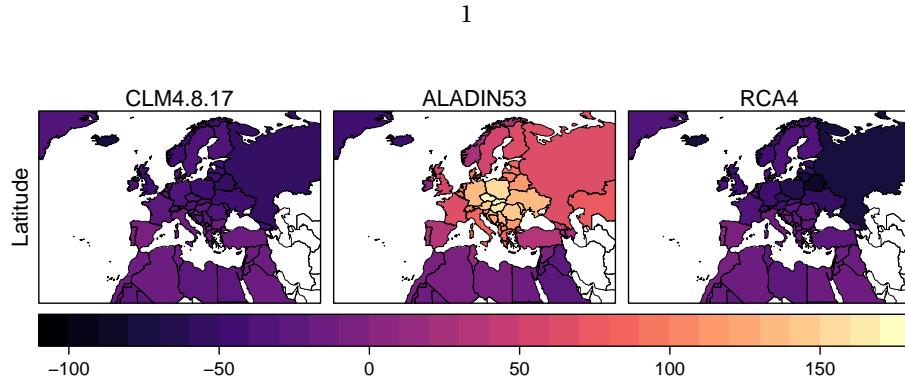
To determine the uncertainty in climate projections is one of the main issues in climate science because the information underlying the simulations is difficult to communicate. In order to isolate the uncertainty sources in a multi-model workit is necessary to design a sensitivity test for each of the models in the study. Up to now, there are not enough simulations from the EURO-CORDEX ensemble including evolving aerosols and for those that include them, there is not exist the same simulation excluding the aerosols forcing, in order to quantify its impact.

For that reason, this preliminary study, although it highlights very important issues, is the firs step in order to understand the role of aerosols in the RCM projections.

## 7.6 Conclusion

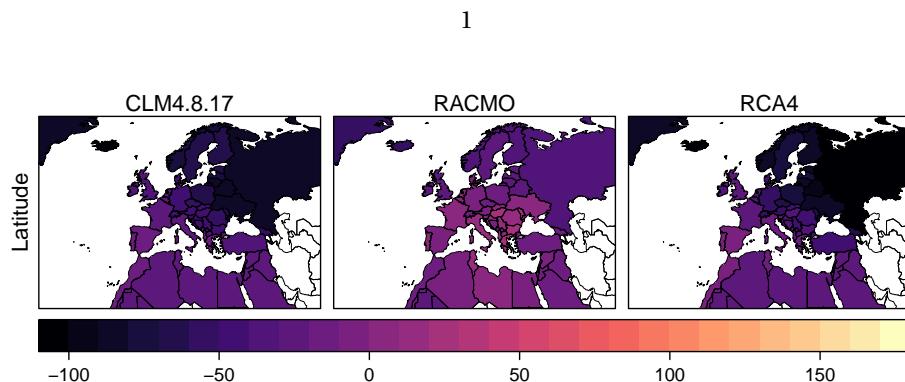
The study shows that regional climate models with evolving aerosols in the scenarios behaves different than those that has a fixed climatology. In the case of ALADIN53 the sign of the anomaly is reversed, agreeing with the positive signal in PV potential projected by GCMs.

It has been shown also that the anomaly of SSR is not directly linked with CLT anomaly in the case of the RCMs simulation that include evolving aerosols. The spa-



---

FIGURE 7.4: *a*



---

FIGURE 7.5: *b*

---

FIGURE 7.6:

tial correlation between SSR and CLT in this models is very low in comparison to the correlation between SSR and AOD anomaly.

The spatial pattern of the aerosols anomaly impact the future projections of PV differently in each country. The results show a general decresase for RCMs with no-evolving aerosols, more important in higher latitudes. However, for ALADIN53 and RACMO22E simulations the Central-Europe countries present a positive anomaly, which is relevant in terms of energy resource assessment.



---

## Bibliography

- [KL+13] Gert König-Langlo, Rainer Sieger, Holger Schmithüsen, et al. *The Baseline Surface Radiation Network and its World Radiation Monitoring Centre at the Alfred Wegener Institute*. 2013.