Green Energy and Technology

Richard Perez *Editor*

Wind Field and Solar Radiation Characterization and Forecasting

A Numerical Approach for Complex Terrain

Springer

Green Energy and Technology

More information about this series at http://www.springer.com/series/8059

Richard Perez
Editor

Wind Field and Solar Radiation Characterization and Forecasting

A Numerical Approach for Complex Terrain

Springer

Editor
Richard Perez
Atmospheric Sciences Research Center
State University of New York
Albany, NY
USA

ISSN 1865-3529 ISSN 1865-3537 (electronic)
Green Energy and Technology
ISBN 978-3-319-76875-5 ISBN 978-3-319-76876-2 (eBook)
https://doi.org/10.1007/978-3-319-76876-2

Library of Congress Control Number: 2018937343

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

A central challenge in assessing renewable energy resources consists in elaborating operational previsions of their temporal and regional fluctuations. Without adequate forecasting methods, the degree of penetration of these energies in the power grid will not meet the expectations triggered by climate change and geopolitical issues. Consequently, the present volume describes recent advances in both deterministic and probabilistic forecasting of wind fields and solar radiation, with focus on complex terrain conditions that are frequently encountered in practice.

The first two chapters expose and characterize the geophysical and meteorological data and parameters upon which the methods presented rely: requirements for spatial and temporal resolution, and availability of multichannel remote sensing from space as operational data sources. Low altitude orbiting satellites provide land cover and albedo data, as well as digital terrain elevation grids, where traditional land-based sources do not exist. Geostationary satellites provide operational meteorological data such as cloud cover with high temporal resolution, as well as atmospheric soundings used as input for numerical weather prediction models, that, in turn, yield the primary data for wind field forecasts.

Orographic representation is one of the features to which wind simulations are very sensitive. Hence, Chapter "Discretization of the Region of Interest" introduces a technique for shaping all involved solids, natural and built, to the required scale in a given region. Chapter "Wind Field Diagnostic Model" describes a wind field diagnostic model for downscaling initial, irregular, or coarsely gridded fields, and underlines the methods' sensitivity to its input parameters. Deterministic wind field prediction is addressed in Chapter "Wind Field Deterministic Forecasting", where regional numerical weather prediction models, with resolution between 1 and 3 km, are presented as adequate providers of the requested information. The probabilistic prediction of wind fields is defined in Chapter "Wind Field Probabilistic Forecasting" as an ensemble methodology designed for the downscaling wind model introduced in Chapter "Wind Field Diagnostic Model". The procedure leans on ensemble forecasting, a methodology used by major meteorological services for dealing with the uncertainties inherent to numerical weather predictions.

An example of the approach for Gran Canaria Island, validated with ground observations, concludes the wind resource part of this inspiring survey.

The, not less forward-looking, solar radiation part begins with Chapter "Solar Resource Variability", where the instability of the resource over different scales is quantified by mean of a predictable variability-smoothing space-time continuum. Its implications for solar penetration on the power grid and variability mitigation strategies are also discussed. Chapter "Solar Radiation Forecasting with Statistical Models" deals with deterministic forecasting and focuses on several statistical methods for intraday time horizons. Probabilistic forecasting is introduced in Chapter "Solar Radiation Probabilistic Forecasting" with two approaches for producing intraday predictions with lead times ranging from one to 6 hours. Verification metrics developed by the meteorological verification community are used to evaluate their performances. Chapter "Solar Radiation Maps" concludes the survey, focusing on the assessment of the resource per se, rather than the assessment of its variability. Terrain-following mapping of surface irradiation is presented as an established discipline, in which advances mainly rely on the growth of computing power. This procedure is applicable for any point on the earth's surface and for any desired timescale.

Emeritus, MeteoSwiss Dr. Antoine Zelenka
February 2018

Acknowledgements

This work has been partially supported by the following contracts and organizations:

- Spanish Government, “Secretaría de Estado de Investigación, Desarrollo e Innovación”, “Ministerio de Economía y Competitividad”, and FEDER, under grant contract: CTM2014-55014-C3-1-R.
- FEDER and the Spanish Government, “Ministerio de Economía y Competitividad” under grand contracts CTM2014-55014-C3-1-R, CTM2014-55014-C3-3-R, the Department of Education, Junta of Castilla y León, under grand contract SA020U16 and by CONACYT-SENER “Fondo Sectorial CONACYT SENER HIDROCARBUROS” under grand contract 163723.
- The authors would like to thank the Institute IUSIANI from the University of Las Palmas de Gran Canaria, State Meteorological Agency (AEMET), Red Eléctrica de España (REE), Canary Islands Technological Institute (ITC), MINES ParisTech/ARMINES, Copernicus Atmosphere Monitoring Service (CAMS), The Satellite Application Facility on Climate Monitoring (CM SAF), and Agroclimatic Information System for Irrigation (SIAR) for providing wind and solar radiation data.
- The authors would like to thank the laboratory PIMENT from the University of La Reunion, the laboratory LARGE from the University of Guadeloupe, the National Renewable Energy Laboratory (NREL), and the meteorological network SURFRAD for providing their ground measurements. They also would like to thank the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing forecast data.

Contents

Contributors

Javier Calvo Sánchez NWP Group of the State Meteorological Agency (AEMET), Madrid, Spain

J. Manuel Cascón Deparment of Economics and Economic History, Faculty of Economic and Business, University of Salamanca. Edificio FES, Salamanca, Spain

Mathieu David Laboratoire PIMENT (Physique et Ingénierie Mathématique pour l'Énergie, l'environnemeNt et le bâtimenT), Université de La Réunion, Saint-Denis Cedex, France

Felipe Díaz University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain

José María Escobar University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Las Palmas de Gran Canaria, Spain

Thomas E. Hoff Clean Power Research, Napa, CA, USA

Sergey Kivalov Atmospheric Sciences Research Center, State University of New York, Albany, NY, USA

Philippe Lauret Laboratoire PIMENT (Physique et Ingénierie Mathématique pour l'Énergie, l'environnemeNt et le bâtimenT), Université de La Réunion, Saint-Denis Cedex, France

Luis Mazorra-Aguiar University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain

Rafael Montenegro University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Las Palmas de Gran Canaria, Spain

Gustavo Montero University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Las Palmas de Gran Canaria, Spain

Gema Morales Martín NWP Group of the State Meteorological Agency (AEMET), Madrid, Spain

Albert Oliver University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Las Palmas de Gran Canaria, Spain

Marc Perez Clean Power Research, Napa, CA, USA

Richard Perez Atmospheric Sciences Research Center, State University of New York, Albany, NY, USA

Jesús Polo Photovoltaic Solar Energy Unit, Renewable Energy Division (Energy Department) of CIEMAT, Madrid, Spain

Eduardo Rodríguez University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Las Palmas de Gran Canaria, Spain

Acronyms

ACF	Autocorrelation Function
AERONET	Aerosol Robotic Network
ALADIN	Aire Limite Adaptation dynamique Developpment INternational
ANN	Artificial Neural Networks
AOD_s	Aerosol Optical Depths
APOLLO	AVHRR Processing Scheme Over cLouds, Land and Ocean
AR	Autoregressive Model
ARD	Automatic Relevance Determination
ARIMA	Autoregressive Integrated Moving Average Model
ARM	Atmospheric Radiation Measurement
ARMA	Autoregressive Moving Average Model
AROME	Applications of Research to Operations at MesoscalE
ARPEGE	Action de Recherche Petite Echelle Grande Echelle
ATOVS	Advanced TIROS Operational Vertical Sounder
AVHRR	Advanced Very High Resolution Radiometer
BDRF	Bidirectional Reflectance Distribution Function
BIC	Bayesian Information Criterion
BSRN	Baseline Surface Radiation Network
CAMS	Copernicus Atmosphere Monitoring Service
CAPE	Convective Available Potential Environment
CBL	Convective Boundary-Layer
CDF	Cumulative Density Function
CFD	Computational Fluid Dynamics
CLC	CORINE Land Cover
CM SAF	Satellite Application Facility on Climate Monitoring
CNRM	National Centre for Meteorological Research
COAMPS	Coupled Ocean/Atmosphere Mesoscale Prediction System
CORINE	Coordination of Information on the Environment
COSMO	Consortium for Small-Scale Modeling
CRPS	Continuous Rank Probability Score

DEM	Digital Elevation Model
DHI_c	Diffuse Horizontal Irradiance on Earth's Surface for Clear-Sky Model
DHI	Diffuse Horizontal Irradiance on Earth's Surface
DNI_c	Direct Normal Irradiance on Earth's Surface for Clear-Sky Model
DNI	Direct Normal Irradiance on Earth's Surface
DTU	Danmark Tekniske Universitet
ECMWF	European Centre Mesoscale Weather Forecast
EDMF	Eddy Diffusivity Mass Lux
EMOS	Ensemble Model Output Statistic
EPS	Ensemble Prediction System
ESRA	European Solar Radiation Atlas
ETRS89	European Terrestrial Reference System 1989
GARCH	Generalized Autoregressive Conditional heteroskedasticity Model
GB	Gradient Boosting Technique
GEFCom	Global Energy Forecasting Competition
GEWEX	Global Energy and Water Experiment
GFS	Global Forecast System
GHI_0	Global Horizontal Irradiance on Earth's Surface for Top of the Atmosphere
GHI_c	Global Horizontal Irradiance on Earth's Surface for Clear-Sky Model
GHI	Global Horizontal Irradiance
GNSS	Global Navigation Satellite System
GOES	Geostationary Operational Environmental Satellite
GOS	Global Observing System
GRS	Global Telecommunication System
HC3v5	Helioclim-3 Database Version 5
HIRLAM	High Resolution Limited Area Models
IEA	International Energy Agency
IFS	Integrated Forecasting System
IODC	Indian Ocean Data Coverage
IrSOLAV	Investigaciones y recursos Solares Avanzados
ISBA	Interaction Sol-Biosphere-Atmosphere
K^*t	Clear-Sky Index
K_t	Clearness Index
LAI	Leaf Area Index
LAM	Limited Area Models
LIDAR	Laser Imaging Detection and Ranging
LMQR	Linear Model in Quantile Regression
MACC	Monitoring Atmospheric Composition and Climate
MAGIC	Mesoscale Atmospheric Global Irradiance Code
MCA	Map of Crops and Utilization
MFE	Spanish Forest Map
MLP	Multilayer Perceptron

MM5	Pennsylvania State University/National Center for Atmospheric Research Mesoscale Numerical Model
MMU	Minimum Mapping Unit
MODIS	Moderate Resolution Imaging Spectroradiometer
MVIRI	Meteosat Visible Infrared Imager
NAM	North American Mesoscale Forecast System
NASA	National Aeronautics and Space Administration
NCEF-GFS	National Centers for Environmental Prediction-Global Forecasting System
NLCD	National Land Cover Database
NOAA	US National Oceanic and Atmospheric Administration
NREL	National Renewable Energy Laboratory
NSD	National Satellite Data
NSRDB	National Solar Resource Data Bases
NWP	Numerical Weather Prediction
OLI	Operational Land Imager
PACF	Partial Autocorrelation Function
PBL	Planetary Boundary Layers
PDF	Probabilistic Density Function
PNOA	National Plan of Aerial Orthophotography
PSU/NCAR	Pennsylvania State University/National Center for Atmospheric Research
PV	Photovoltaic
PVGIS	Photovoltaic Geographical Information System
PVPS	Photovoltaic Power System Programme
PYREX	Pyrenean Experiment
QRF	Quantile Regression Forest
QRNN	Quantile Regression Neural Network
RBCNs	Regional Basic Climatological Networks
RBSNs	Regional Basic Synoptic Networks
RMIB	Royal Meteorological Institute of Belgium
RMSE	Root Mean Square Error
SARAH	Surface Solar Radiation Data Set—Heliosat
SBL	Stable Boundary-Layer
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SHC	Solar Heating and Cooling Programme
SIA	Integrate Water Information System
SIGPAC	Geographical Information System of Agricultural Parcel
SIOSE	Information System of Land Cover of Spain
SIS	Surface Incoming Shortwave Radiation Data
SOLEMI	Solar Energy Mining
SRML	Solar Radiation Monitoring Laboratory
SRRA	Indian Solar Radiation Resource Assessment
SSO	Subgrid Scale Orography
SSRD	Surface Solar Radiation Downwards

SUNY	State University of New York
SURFEX	Surface Externalise
TCC	Total Cloud Cover
TEB	Town Energy Budget
THORPEX	The Observing System Research and Predictability Experiment
TIGGE	THORPEX Interactive Grand Global Ensemble
TIRS	Thermal Infrared Sensor
TLS	Thick Level Set
TMY	Typical Meteorological Year
TOA	Top of the Atmosphere
UKMO-UM	United Kingdom Met Office—Unified Model
UTM	Universal Transverse Mercator
VD	Variant Deficit
WCRP	World Climate Research Programme
WMO	World Meteorological Organization
WQR	Weighted Quantile Regression
WRF	Weather Research and Forecasting Model
WVM	Wavelet Variability Model
WWW	World Weather Watch

Part I
Analysis and Characterization of Geographical and Meteorological Data

The aim of this book is to describe different methodologies for solar radiation and wind fields characterization and forecasting. In Part II and III of this book it is possible to find wind fields and solar radiation methodologies respectively. These Chapters describe variability models, deterministic forecasting, probabilistic forecasting or the methodology for generating natural resources maps.

The Chapters in this Part are dedicated to introduce different data and tools needed for the explained models. Forecasting and variability models need accurate ground measurement for modeling performance and for calculating accuracy. Indeed, solar radiation and wind field ground measurement gauges, Numerical Weather Prediction data and Satellite derived data are described, as well as quality check procedures. On the other hand, to elaborate natural resources maps, some geographical and meteorological information involved in the numerical modeling need to be presented. Moreover, the meccano method developed to construct simultane- ously tetrahedral meshes and volumetric parameterizations of solids is introduced focusing on the application of the method to build tetrahedral meshes over complex terrain, that are interesting for simulation of environmental processes.

Acquisition and Analysis of Meteorological Data

Javier Calvo Sánchez, Gema Morales Martín and Jesús Polo

Abstract Wind and solar radiation observations are required for renewable energy modeling and forecasting. High-quality ground measurements are essential for renewable energy studies. Ideal wind measurement devices should respond to slightest breezes, be strong enough to stand up high winds, give a fast and accurate answer for turbulent fluctuations, and have a linear output and a simple dynamic performance. The solar radiation reaching the earth's surface contains several components, beam, diffuse, and reflected (albedo) radiation, as a consequence of the interaction with atmospheric particles. The instruments for measuring solar radiation are classified according to the working principle (mainly thermoelectric or photoelectric sensors) and to the component of solar radiation to be measured. Ground measurements give information about solar radiation or wind fields in a specific location. On the other hand, in many applications for the modeling and prediction of natural resources, meteorological data with a greater spatial distribution are needed. Satellite models offer meteorological data estimated from satellite images with a high spatial and temporary resolution. In the same way, Numerical Weather Predictions models give information about several meteorological variables with a great spatial resolution.

1 Ground Measurement Data

Wind observations are required in a wide range of activities. For example, monitoring wind speed and wind direction is a vital issue when it comes to safe transportation. Adverse weather conditions affect all types of transport: private, public, and industrial

J. Calvo Sánchez (✉) · G. Morales Martín
NWP Group of the State Meteorological Agency (AEMET), c/ Leonardo Prieto Castro, 28040 Madrid, Spain
e-mail: fcalvos@aemet.es

G. Morales Martín
e-mail: gmoralesm@aemet.es

J. Polo
Photovoltaic Solar Energy Unit, Renewable Energy Division (Energy Department) of CIEMAT, Avda Computense 40, 28040 Madrid, Spain
e-mail: jesus.polo@ciemat.es

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_1

or business transport. High winds pose a particular threat, especially for tall, high-sided vehicles such as the lorries that use road systems to deliver.

Buildings may be damaged by stronger wind events for which they were designed. However, except for tornadoes, most damage occurs because various building elements have limited wind resistance due to inadequate design, application, material deterioration, or roof system abuse. Wind data is also needed to run transport models and for the production of weather forecasts based on numerical weather prediction (NWP) models. In agricultural purposes, strong winds affect the plant's life both mechanically and physiologically and, as a consequence, affect crops production. Wind also has its importance in the dispersion of air pollution conditions, which is crucial related to health warnings to population. In sport and other leisure activities, the wind can be an issue as it may prevent from developing particular activities.

1.1 Wind Data

Following [55], one of the definitions of wind is a three-dimensional vector quantity with small-scale random fluctuations in space and time superimposed upon a larger scale organized flow.

As the vertical component of the wind speed is usually small near the earth's surface, in most applications, the measured wind is the horizontal component. Sometimes the instruments determine the two horizontal north and east components. Alternatively, it is possible to measure speed and direction directly, from which we can derive the components if necessary. In any case, the direction of the wind is referred to the direction from which the wind is blowing measured in degrees clockwise from north. In specific applications such as turbulence research, it is important measuring the vertical component of the wind too. Then, the wind might be expressed using the three components of the vector or, alternatively, the speed, the horizontal direction, and the elevation angle.

As it can be inferred from the definition of wind, this meteorological variable can be described in terms of mean flow and gustiness or variation about the mean. For practical purposes, more useful concepts derived from the general definition of wind are the following:

- *Averaged quantities* are those wind quantities averaged over a period of time. Depending on the application, the period used often range from 10 min, for forecasting purposes, up to hours or days for climatological studies. Some measurements averaged below 10 min are sometimes required in certain aeronautical activities.
- *Peak gust*, which is the maximum observed wind speed over a specified time interval. This quantity is very important for transport activities, specially for aircraft take-off and landing.
- *Gust duration*, which is the duration of the observed peak gust.

According to the *World Meteorological Organization* (WMO) standards, wind speed should be measured to the nearest unit (meters per second, kilometers per hour, or knots) and wind direction should be measured in degrees from north direction and reported to the nearest 10 degrees. The meteorological requirements are summarized in [1] but for horizontal speed, an accuracy of 0.5 ms^{-1} below 5 ms^{-1} and better than 10 percent above 5 ms^{-1} is usually sufficient. For wind direction, it should be measured with an accuracy of 5°.

1.1.1 Measurement Devices

As mentioned in [7], the ideal wind-measuring instrument would respond to the slightest breeze yet be strong enough to stand up in hurricane-force winds, respond to rapidly changing turbulent fluctuations, have a linear output, and exhibit simple dynamic performance. Of course, it is almost impossible to build instruments that meet all these requirements. Rather, the sensors are designed to work optimally for particular exposure conditions or specific needs.

Traditional wind measurements

Although the best approach for measuring wind is the use of instruments, sometimes measurements are not available for different reasons: the equipment is not working temporarily, the place is not accessible to install the instrumentation, or the measurement is not representative of an extended area. In these cases, indirect ways to estimate the wind have to be used and, unfortunately, errors in values might be large.

For wind speed, the estimation is made by observing the effects of wind in objects, people, or buildings around the observer. The Beaufort scale is an empirical measure for describing wind intensity according to the observation of surrounding conditions (see Table 5.1 and 5.2 of Part I in [1]). Similarly, the wind direction might be estimated by observing the drift of smoke from a chimney, the movement of flags, leaves, etc.

However, for nearly all applications, a precise value of wind speed and direction is necessary, particularly useful are the averaged quantities. Thus, a wind-measuring system is usually needed, comprising a sensor itself as well as a processing and recording system.

The simplest, and low cost, equipment consists of handheld anemometers for wind speed and simple vanes, mounted on a pole, for wind direction. The wind velocity measurement should be taken from a point well exposed to the wind, far apart from obstructions. Regarding the vane, the measurement alignment should be the one pointing the nearest compass direction. Both cases need a human observer to read the values (Fig. 1).

Following the same principle than the simplest devices, cup and propeller sensors are the most common instruments to provide wind velocity. The wind velocity is determined as a function of the angular velocity of the cup or propeller rotor (Fig. 2). Cup and propeller anemometers are linear over most of their range, with a notable exception at the lower end. Very small winds could not overcome the internal friction of the gauges to start the rotation. Consequently, there is a threshold wind speed,

Fig. 1 Left. Handheld anemometer which should face directly to the wind. Right. Simple vane indicating where the wind comes from

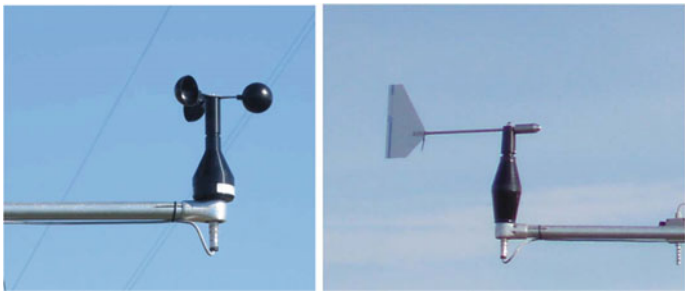


Fig. 2 Cup anemometer (left) and propeller vane (right)

below which the anemometer will not turn. There is also an upper limit for the maximum wind speed above which the anemometer cannot keep going without any damage to the probe.

The response of the cup and propeller wind sensors to changes in wind speed is characterized by the response length, which depends on the geometry of the device and the materials used. Usually, the response is faster for acceleration than for deceleration, so that the average speed of these devices overestimates the actual average wind speed [9, 14]. It even depends on the vertical wind conditions. This overspeeding, which in some cases might be as much as 10% of the measurement, has to be quantified as it might be considered for some applications.

As mentioned before, for wind direction, more or less sophisticated vanes are used. They are flat plate or airfoil that can rotate about a pole and, in static equilibrium, is oriented along the wind vector. There is a counterweight to balance the vane about the bar where it is mounted [7]. The vanes are devices that have to ensure two premises:

it must not have a preferred direction and the instrument has to be well aligned to the true north. It is a concern the response of the vane to rapid changes in wind direction. Usually, it is more problematic for low winds where the measurements might oscillate considerably about the true direction but not so noticeable for large winds. More details about this limitation can be found in [50, 55].

Other wind sensors

There are other sensors more sophisticated, which provide wind values derived from different physical principles. They use faster response sensors so the wind can be sampled at higher frequencies. Thus, they are suitable for turbulence measurements which usually need to sample data at higher rates than 1 measure per second.

Pitot tube anemometers They measure overpressure in a tube that is kept aligned with the wind vector (Fig. 3). The wind speed is related to the root square of the overpressure value. Actually, its major requirement is the need for a high-quality differential pressure sensor to convert it to wind values. Pitot tubes are simple, reliable, inexpensive, and suited for a variety of environmental conditions, including extremely high temperatures and a wide range of pressures. However, they provide erroneous values for very low and very high winds. They are used on aircraft as speedometers. A full description of this sensor can be found in [32, 79].

Sonic anemometers They measure the time between emission and reception of an ultrasonic pulse traveling over a fixed distance [42]. Their response to wind changes is linear but, as it depends on the sound velocity in air, measurements are sensitive to changes in temperature and humidity. Further information about how sonic anemometers derive the components of wind may be found in [7]. The sonic anemometer is fairly expensive compared to a simple mechanical sensor and requires considerably more power. There can be signal loss due to heavy rain or wet snow. However, these gauges are able to take measurements at very high frequencies (usually at 20 Hz) which makes them suitable for fun-

Fig. 3 Pitot tube anemometer mounted in an aircraft

damental research on turbulence. Its use is widely spread all over the scientific community.

Hot-wire anemometers They measure the cooling on thin heated wires because of the action of the wind (Fig. 4). The cooling depends on the mass flow rate, i.e., the speed and density of the flow past the sensing element. Thermal anemometry became popular because the technique involves the use of very small probes. They offer very high spatial resolution and excellent frequency response characteristics. The basic principles of the technique are relatively straightforward and the probes are difficult to damage if reasonable care is taken. These sensors are well suited to measurement of atmospheric turbulence because they work at a high rate. However, they are quite sensitive to the external conditions so their use is not very common, specially in wet environments. In these situations, they tend to overestimate the wind speed. Another shortcoming is that hot-wires are expensive and they must be calibrated as they are susceptible to drift. They also have the problem that they are too sensitive at low wind speeds (extended reading on this can be found in [7]).

Hot-disk anemometers They measure the gradient of temperature across a chip device. It provides both wind speed and direction. Their use is not wide extended since they are recently developed and their limitations are similar to the hot-wire anemometers.

Remote wind sensors Devices such as the sodar, lidar, and radar, measure distant winds from the place where the equipment is set. Their methodology is based in sound, light, and electromagnetic wave propagation and backscatter detection (Doppler shift). They are quite useful for those places of interest where no devices can be installed for different reasons. Additionally, they are essential to measure upper winds.

Fig. 4 Hot-wire anemometer

Usually, we need the combination of two different devices in order to measure wind speed and direction. A cup anemometer and vane combination or a propeller-vane equipment is often the preferable systems, which can be used also to derive the velocity components from the wind speed and direction. If needed, sometimes the association of two propellers in orthogonal configuration is used for the same purpose. A more sophisticated instrument is the sonic anemometer with the transducers pointing at the three directions of space (Fig. 5). The disadvantage of measuring the wind components using the same instrument is that if the device stops working, wind information is missed completely. On the other hand, having just one gauge makes the equipment maintenance easier (Fig. 6).

At the end, it is a user decision to install a particular measurement equipment according to their needs.

Fig. 5 Two different types of sonic anemometers to measure the three components of wind

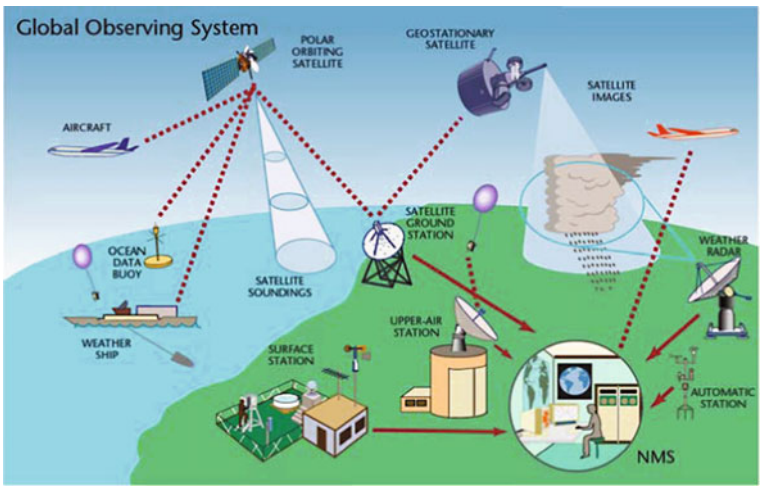


Fig. 6 Global Observing System of the World Weather Watch Program comprises all types of observation available to monitor the atmosphere and the environment. *Source* http://www.wmo.int

1.1.2 Data Processing Methods

Taking apart turbulent studies, typical meteorological operations need averaged wind quantities more than instantaneous wind values. Although from a mathematical point of view, the average of a sample is a simple concept, there are some considerations to be taken into account when the wind measurements are averaged. The mean vector speed in the averaged wind direction is less than the average of all instantaneous wind speeds [50, 92]. Usually, the differences between these two quantities are small. However, in situations such as strong turbulent conditions close to the ground, it might be a concern the way the wind is averaged. It also may depend on the sensor dynamic response characteristics or the sensor sensitivity to relative wind direction.

Consequently, the optimum wind system is the one which provides the two components of the horizontal wind, preferably using fast-response sensors like sonics or thermal wind gauges. It calculates the time average on every component of the wind vector and then compound the averaged wind speed and direction.

Not only the averaged speed has to be calculated cautiously, it is also a concern for the wind direction the discontinuity between 0 and 360°. It causes a problem when we average directions around north point, leading sometimes to erroneous values.[1] This can be solved by applying a correction to the direction record (adding or subtracting 360° when necessary, for example) to overcome this problem.

In some cases, the interesting magnitude to process is the peak gust of winds. In transportation activities, for example, when values of wind gusts are above a threshold, it may be decided to interrupt the service because of security reasons. Sometimes peak gusts are more dangerous than sustainable high winds. Other areas where gust values are relevant is in construction. The standard deviation of wind direction and wind speed is used as a complement of the gust information because it provides an estimation of how the wind fluctuates. Unfortunately, the computation of both, the gust and the standard deviation, are extremely sensitive to the dynamic response of the measurement system. Slowly responding systems flatten out the extremes and provide gusts with small amplitude. On the contrary, fast-response systems record high and narrow peaks. Standard deviation of wind direction and wind speed can be computed by an equipment taking samples at least at a rate of 1 Hz [1].

1.1.3 Installation and Recommendations

According to the *World Meteorological Organization* (WMO), the standard height for wind measurements is at 10 m. At this height, the optimum wind observation location is the one where the observed wind is representative over an area of at least a few kilometers. Ideally, the anemometer should be installed in an open terrain such as the distance between the sensors, and any obstruction is at least 10 times the height

[1] For example, the average of two measurements of 10° and 350° is 180°, which is just the opposite direction to the actual wind.

of the obstruction (preferably 20 times) (see [1]). However, sometimes the terrain is not so flat or homogeneous for considering the measurement representative. In these cases, it should be chosen the site where the perturbation of the measurements is minimum. There are methods to determine how the local topography affect the wind values and how to make some corrections in order to derive the best estimation of the wind at 10 m [91]. The measurement errors, associated with a not suitable site according to the WMO recommendations, are usually much larger than those coming from the intrinsic characteristics of the probe. Because of this, it is very important that including the observations, a good description of the place around the measurement station should be available [93]. In this way, we can assess whether there is any preferable direction for the wind or the existence of any obstacle that might perturb the wind speed.

Not only the characteristics of the site where the equipment is installed may prevent from taking good measurements. The masts or the tower itself where the instruments are anchored can perturb the measurement as well. In those cases where the device cannot be installed at the top of the mast, it should be suspended in a side boom with a length of at least three mast or tower widths [31]. Finally, when the anemometer must be installed in a building, it should be placed on the top of it and additionally raised in a mast.

Some extra considerations should be taken when meteorological conditions are adverse. The well-designed equipment are robust enough to work properly in a wide range of weather conditions. However, the sensors are specially affected by freezing weather. The snow or hail may deteriorate the exposed parts of the device and might become out of operation. For these situations, there are special wind instruments with shields to minimize the damage.

Anemometers installed at sea deserve special attention. Measurements taken at 10 m at sea account for additional inconveniences. It is almost impracticable to take data at this height as it depends on the state of the sea and the influence of the tides. Despite this fact, in [21], it is concluded that it is much more important than the standard height the influence of the ship, buoy, or platform where the instrument is mounted.

1.1.4 Maintenance and Calibration

When the wind sensors are used at the open-air, they are prone to deteriorate over time. The continuous exposure of the equipment to the friction of air, high temperatures, high wind gusts, humidity, dust, etc., turns out to deterioration of plastic parts, fatigue of the sensor components, corrosion, and other physical issues that change the original design of the probe. All these problems necessarily spoil the quality of the measurements and the users should be aware of it.

Inconveniences might be incorrect zero of the sensor, noise and low sensitivity, and irregular or reduced variability of recorded wind. In those more sophisticated devices, there are also electronic failures that may produce erroneous measurements Thus, it is recommended to inspect the equipment for physical damages and periodically check

out the functionality of the sensors. The fully reliable calibration of an anemometer should be done in a wind tunnel along with a suitable reference instrument. This wind tunnel must be large enough to provide a steady flow that is uniform across the tunnel. Sometimes these reference instruments are used in situ to calibrate other devices and check the functionality of the sensors without uninstalling the operating system.

1.1.5 Surface Measurement Networks

One of the core programs of the WMO is the *World Weather Watch* (WWW) (http://www.wmo.int/pages/prog/www/index_en.html). Established in 1963, its main goal is to share meteorological information around the world, joining resources and efforts, to monitor the atmosphere and the environment. One of the components of this program is the *Global Observing System* (GOS) (http://www.wmo.int/pages/prog/www/OSY/GOS.html), which intends to coordinate methods and facilities for making meteorological and other environmental observations on a global scale.

The GOS comprises observing facilities on land, at sea, in the air, and in outer space. These facilities are owned and operated by the Member countries of WMO. These countries are committed to operate the observing systems following the WMO recommendations in order to have a global network of comparable measurements (further information about the station requirements is detailed in [2]). There are about 11,000 stations on land making observations at or near the earth's surface, at least every 3 hours and often hourly, of meteorological parameters such as atmospheric pressure, wind speed and direction, air temperature, and relative humidity. Some 4000 of these stations comprise the *Regional Basic Synoptic Networks* (RBSNs) and over 3000 stations comprise the *Regional Basic Climatological Networks* (RBCNs). Data from these stations are exchanged globally in real time.

The observation systems that take part of these networks should follow the guidelines stated by the WMO according to installation, calibration, and maintenance specifications.

Apart from this global measuring network, every meteorological service has its own observation network which should be set up following the WMO standards. However, not always it is possible to follow the rules when a measurement device is installed. That is the reason why there is a classification ranging from 1 to 5, for every variable and every site, to describe how far the location is from the WMO guides. Class 1 corresponds to those places following the recommendations, whereas class 5 describes a site inappropriate for standard observations. Usually, these classes are associated with an estimation of the measured error.

1.2 Solar Radiation Data

The solar radiation reaching the earth's surface contains several components as a consequence of the interaction with the atmospheric constituents. The solar global irradiance incoming to a specific surface is the sum of the direct normal irradiance (DNI) that comes from the sun disk without suffering any interaction through the media, the diffuse irradiance from the whole sky as a result of scattering, and the reflected irradiance from the surrounding surfaces (which becomes zero for horizontal surfaces). When DNI is referred to as the solar irradiance measured by a pyrheliometer (which normally subtends a viewing angle slightly larger than the sun disk), it contains both the radiation from the sun and a narrow annulus of sky around the sun disk which is commonly called circumsolar radiation [5]. The circumsolar radiation results from the Mie scattering processes along the sun director vector. The instruments for measuring solar radiation are classified according to the working principle (mainly thermoelectric or photoelectric sensors) and to the component of solar radiation to be measured [28].

1.2.1 Thermopile Devices

Thermopile radiometers measure solar irradiation essentially through detecting the difference between the temperature of an illuminated and shadowed area (Seebeck effect). They have a very flat spectral response to the irradiance in the wavelength range of around 300–3000 nm. For broadband hemispherical solar radiation, the thermopile instrument is called thermopile pyranometer. The thermopile sensor is protected by one or two glass domes. Pyranometers can be used for measuring global horizontal and tilted irradiance and also for diffuse irradiance if a proper shadowing device is used for blocking the direct normal radiation. Figure 7 shows several models of thermopile pyranometers.

DNI is measured by a pyrheliometer which is a thermopile radiometer with a 5 field of view and a flat window. This instrument requires continuous tracking of the solar disk in the sky and thus is mounted on a solar tracker system. Figure 8 shows several models of pyrheliometers mounted in a solar tracking device.

According to ISO-9060, thermopile pyranometers can be classified into three groups in precision decreasing order: secondary standard, first class, and second class pyranometers. The uncertainty of thermopile radiometers is around 5% for global irradiance, 3% for direct normal irradiance, and 7% for diffuse irradiance if maintenance requirements (regular cleaning of domes and windows and good alignment for pyrheliometers) are kept [90].

Fig. 7 Thermopile pyranometers

Fig. 8 Pyrheliometers

1.2.2 Photovoltaic Devices

Silicon-based sensors essentially measure the short-circuit current, which is proportional to the solar irradiance. These sensors can be grouped into two categories: photodiode pyranometer and reference cells. They have a faster response time and lower cost compared to thermopile radiometers but lower accuracy and limited spectral responsivity (around 300–1100 nm). Therefore, the reference cells

in outdoor conditions need to account for temperature corrections and spectral mismatch. Figure 9 shows pictures of different silicon-based sensors.

1.2.3 Surface Measurement Networks

There are many different networks of ground stations of different nature with measurements of solar irradiance or insolation. Most countries usually have a national network of solar radiation measuring stations managed by the national meteorological institutions. The World Radiation Data Center (WRDC, wrdc.mgo.rssi.ru), sponsored by the World Meteorological Organization (WMO), is a repository of most of the data coming from the national meteorological institutions to ensure the availability of these data for research by the international scientific community. The WRDC collects data since 1964 to present for over 1000 stations primarily as daily

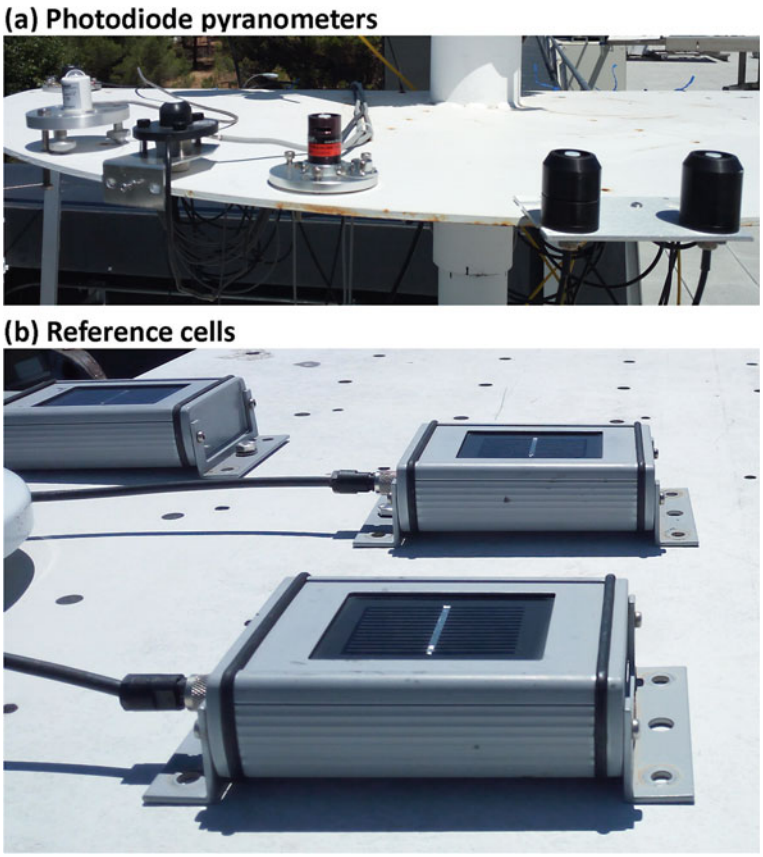


Fig. 9 Silicon-based sensors for measuring solar radiation

sums of global horizontal irradiance (GHI) and sunshine duration. At the top level of international solar radiation networks are the Baseline Surface Radiation Network (BSRN, bsrn.awi.de) and the Atmospheric Radiation Measurement (ARM) program. The BSRN network is a project of the World Climate Research Program (WCRP) and the Global Energy and Water Experiment (GEWEX). BSRN ground data includes in most cases GHI, DNI, and diffuse (DHI) irradiance data of very high quality up to 1-min. Figure 10 shows a map of the locations of BSRN stations.

The ARM program was created by the US Department of Energy to develop several highly instrumented ground stations to study cloud formation processes and radiative transfer phenomena [62]. A complete and automated algorithm for quality check is used to detect erroneous data [48]. There are also some local networks that offer high-quality solar radiation data. For instance, the Solar Radiation Monitoring Laboratory (SRML) of Oregon University supplies high-quality solar resource data on at 30 sites in the Northwest US (http://solardata.uoregon.edu/).

1.2.4 Quality Checks and Procedures for Solar Radiation Data

The quality considerations of a solar radiation measuring system refer to both the design and setup phase and the measuring period. Proper setup is essential because even the best quality checks cannot improve poorly observed data [27]. Therefore, proper sitting, calibration procedures, sampling and averaging, maintenance, and recording of data are as important as the accuracy of the instruments. Once the

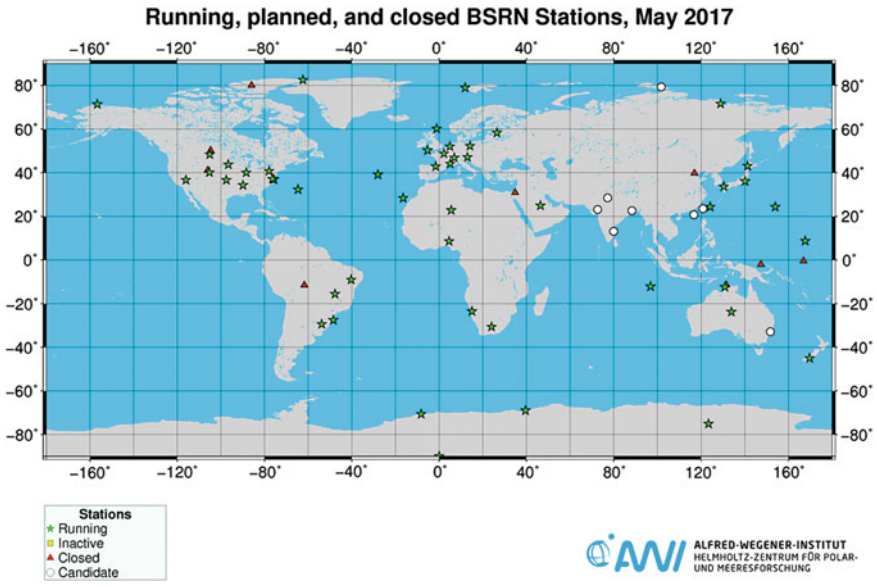


Fig. 10 BSRN ground stations (*Source* bsrn.awi.de) [44]

measuring station is working complete quality checks must be done before the use of the data, and a good flagging criterion must be selected to help the best exploitation of the recoded data [25]. The BSRN establishes a standard quality check of solar radiation measurements divided into three levels [82]: physically possible limits, extremely rare limits, and cross-comparison among the different components. In a similar way, the EU-MESOR projects proposed a quality check based on three levels [34]. Another well-known procedure for quality of solar radiation measurements was developed by NREL and referred to as SERI QC [54]. Recently, additional procedures and software have been developed for complete quality assurance of solar radiation data including filling the gaps. One example of this procedure is proposed by the Royal Meteorological Institute of Belgium RMIB [41]. Another example that follows the international best practices was recently developed for the Indian Solar Radiation Resource Assessment (SRRA) network [85]. Sengupta et al., within the framework of the Tasks 36 and 46 of IEA-SHC, made a thorough review of solar radiation measurements including instrument selection, installation, and operation and maintenance [87].

2 Satellite-Derived Data

The need of long-term solar radiation data and the limitations on the availability of solar radiation measurements produced the development of methods and models for estimating solar radiation components from weather satellites. These methods started to appear in parallel to the availability of meteorological satellites and have evolved a lot as a result of more than 30 years of experience. Thus, the methodologies for computing solar irradiance from meteorological satellite imagery have acquired high maturity and reliability, being nowadays a widely used source of datasets for many solar energy, environmental, and climatology studies. Geostationary satellites are observing and monitoring the earth in operational way for over 30 years. A geostationary satellite circles around the earth at an orbit more than 36000 km above the equator. The geostationary orbit allows the continuous monitoring of clouds over a wide geographical extent, and thus, the imaging sensors aboard geostationary satellites are the major source of satellite-derived solar radiation data. The temporal resolution is very high (up to 15 min) and the spatial resolution can reach 1 km at nadir. Figure 11 shows the position of the main geostationary weather satellites available at NOAA's geostationary satellite server.

Geostationary weather satellites have onboard sensors, named radiometers, which essentially make measurements of the upwelling electromagnetic radiance in several spectral ranges: visible channel (0.5–1.1 m), thermal infrared band (10.5–12.5 m), and water vapor infrared channel (5.7–7.1 m). The spectral channels of the radiometers have been increasing with the new generation of instruments. Thus, Meteosat first generation MVIRI sensor has three spectral channels, second generation SEVIRI radiometer includes 12, and third generation will expand to a set of six satellites (four imaging and two sounder, www.eumetsat.int). Likewise, the last

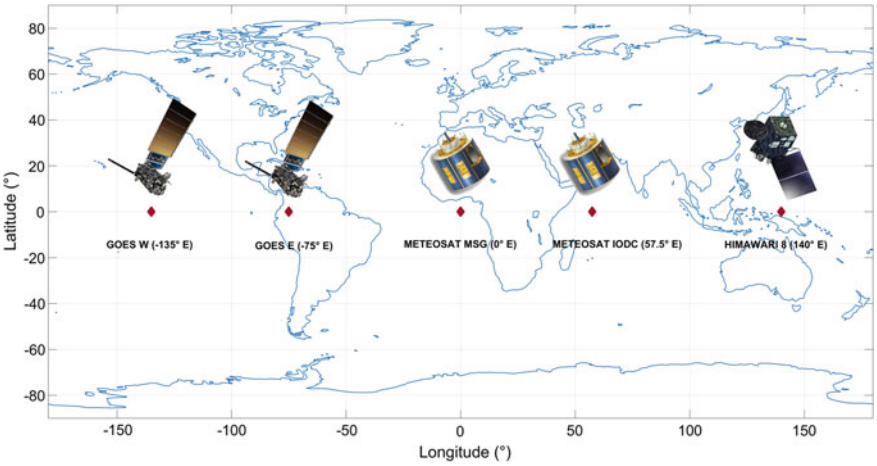


Fig. 11 Position of geostationary weather satellites available at NOAA's server

generation of GOES satellite includes a sounder with 19-channel radiometer and the GOES imager five channel radiometer (http://noaasis.noaa.gov/NOAASIS/). Similarly, the Japanese satellite Himawari-8 and -9 have the Advanced Baseline Imager instrument with 16 spectral channels in different bands from 0.43 to 13.4 m (www.data.jma.go.jp). There are many methods proposed elsewhere for deriving solar radiation data from satellite images using both geostationary and polar-orbiting satellites. However, those based on the use of geostationary satellites provide long-term time series of solar irradiance at high temporal resolution (up to 15 min) with broad spatial coverage. First methodologies proposed the use of normalized reflectance measured by the satellite sensor to determine de cloud transmission [10, 56, 69]. The different models have been historically classified as pure statistical and physical satellite models depending on how they model the solar radiation attenuation and solve the radiative transfer through the earth's atmosphere [59, 60]. Nevertheless, many of the recently proposed and currently used methodologies might be defined as semi-empirical methods since they normally use physical transmittance models for computing solar irradiance under clear-sky conditions [75]. A satellite image is essentially a measure of the upwelling radiation traveling from the earth-atmosphere system to the onboard sensor at a specific time instant and within a spectral range determined by the sensor channel. This radiation consists of both the reflected radiation from the earth-atmosphere system and the backscattered radiation from the atmosphere, Fig. 12. Therefore, the calibrated information from the image can be related with the state of the atmosphere from clear sky to overcast situations and incident solar irradiance can be retrieved accordingly. This is the fundamentals of the semi-empirical models for satellite-derived solar radiation.

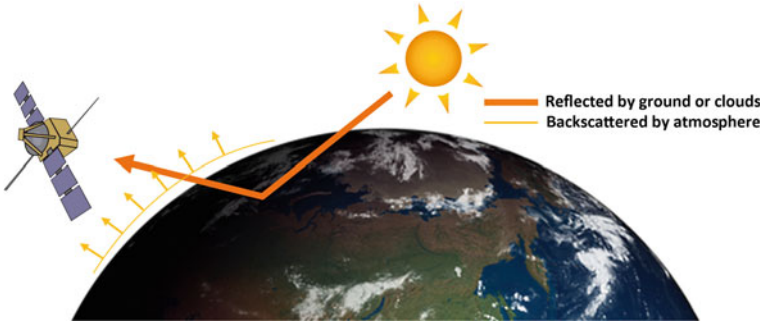


Fig. 12 Components of solar radiation measured by a satellite sensor

2.1 Semi-empirical Models

The most well-known semi-empirical models for estimating solar irradiance from satellite imagery are the Heliosat method developed initially to work with Meteosat visible images [10, 19, 80] and the SUNY (State University of New York) model that was formerly conceived to work with GOES imagery [63, 66]. Both are based on the cloud index concept as a parameter for normalizing the reflectance measured by the satellite sensor. The cloud index is defined as

$$n = \frac{\rho - \rho_g}{\rho_c - \rho_g} \tag{1}$$

where ρ is the instantaneous reflectance measured by the satellite sensor, ρ_g is the ground albedo (i.e., the reflectance of the darkest pixel), and ρ_c is the cloud albedo (i.e., the reflectance of the brightest pixel). The denominator in the above definition for the cloud index is denoted as dynamic range, Fig. 13; thus, the cloud index represents a way to normalize the satellite reflectance (offset corrected) by the dynamic range. Consequently, for complete overcast conditions, the satellite albedo tends to be close to the cloud albedo and the cloud index should be the unity; on the contrary, for clear-sky conditions, cloud index should be zero because the satellite albedo equals the ground albedo. Cloud index is related to the clear-sky index (K_t^*), defined as the ratio of global irradiance at ground to the global irradiance under clear-sky conditions, to estimate solar global irradiance, see Chap. 8. The first Heliosat version proposed a linear relationship between clearness index and cloud index. Later on, the use of clear-sky index was better proposed instead of clearness index [4]. Thus, in Heliosat-2 model, the proposed relationship was [33, 80].

$$K_t^* = \frac{GHI}{GHI_c} = \begin{Bmatrix} 1.2 & n < -0.2 \\ 1 - n & -0.2 \leq n < 0.8 \\ 2.0667 - 3.6667n + 1.6667n^2 & 0.8 \leq n < 1.1 \\ 0.05 & 1.1 \leq n \end{Bmatrix} \tag{2}$$

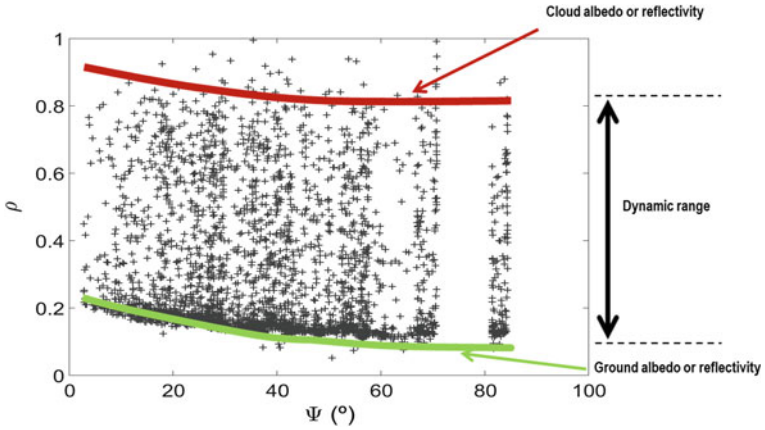


Fig. 13 Satellite sensor reflectance versus the co-scattering angle

Several modifications were proposed in the last years on both the Heliosat-2 method scheme [11, 15, 84] and on the correlation between clear-sky and cloud index [94], where GHI_c corresponds to global solar radiation for a horizontal surface for a clear-sky model.

New formulations for cloud index and clear-sky transmittance calculations were proposed in the Heliosat-3 model as a revision and improvement of the earlier methods for adapting the methodology to Meteosat Second Generation visible channel [16, 17, 57]. Some of the main novelties of Heliosat-3 method were the computation of the instantaneous satellite albedo taking into account the backscattered radiation from the atmosphere as a function of the co-scattering angle (the angle subtended by the sun and satellite director vectors) and the algorithm for ground albedo computation. The Heliosat-3 method was later modified by proposing a dynamic model for estimating the ground albedo as a function of the co-scattering angle (improving thus the limitations of earlier versions for high reflective areas) and extending the application to other geostationary satellites [71, 73]. All these modifications and the possibility of using different clear-sky models and atmospheric-derived input resulted in a model called Intisat Lib [72]. The MagicSol algorithm used in the Satellite Application Facility on Climate Monitoring (CM SAF, www.cmsaf.eu) surface radiation product is also a modified version of Heliosat-2 method [77, 83]. The main modifications are the implementation of a self-calibration method to account for satellite switches and sensor degradation and the use of MAGIC (Mesoscale Atmospheric Global Irradiance Code) for fast clear-sky irradiance computation using look-up tables. Latest version of this method includes improvements in the radiative transfer modeling (SPECMAGIC model) and in cloud index estimations (denoted as effective cloud albedo) to deliver a new dataset product called SARAH that covers over 30 years of observations based on Meteosat first and second generation [3, 58, 81]. This methodology is also integrated into the PVGIS (http://re.jrc.ec.europa.eu/pvgis/) web service offering solar resource and performance of photovoltaic technologies [3, 35]. The

SUNY model uses also the cloud index concept and it calculates it from the pixel brightness that is considered proportional to the earth's radiance sensed by the satellite [65, 66]. The main difference with Heliosat family models, in terms of cloud index concept, is that SUNY method works directly with the raw pixel brightness instead of estimate the satellite sensor reflectance. The raw pixel is first normalized by the cosine of solar zenith angle to account for first-order solar geometry effect. The relationship proposed for the clear-sky index was

$$K_t^* = \frac{GHI}{GHI_c} = 0.02 + 0.98(1 - n) \tag{3}$$

The second version of the SUNY model was used to produce the National Solar Resource Data Bases (NSRDB) of the National Renewable Energy Lab (NREL) and the model was also adapted to be used with Meteosat IODC images over the Indian Ocean [68]. The SUNY model version 3 makes use of both visible and infrared channels imagery [20]. The latest version is the fourth with notable performance improvement over the preceding versions, which have been improved in better source of aerosol optical depth, short-term forecast scheme integrated into the model, and a better empirical method for estimating the dynamic range [67]. The SUNY model is also the basis of the SolarAnywhere database. On the other hand, the SolarGIS model is based also on SUNY incorporating additional features for terrain effects, clear-sky index calibration adapted to each satellite characteristics, and dynamic computation of the dynamic range upper bound [11, 12, 64].

2.2 Physical Models

Physical model approach consists of solving the radiative transfer equation in the atmosphere and requires detailed information of the main atmospheric constituents. They use satellite data for retrieving cloud properties or cloud optical depth. Nevertheless, despite this conceptual different philosophy, the differences with semi-empirical models are not so significant for practical purposes. The first physical-based method was based on energy conservation within an earth-atmosphere column where the cloud effects were calculated from the measured satellite visible brightness [29]. Later a model based on the radiation budget for the tropical Western Pacific Ocean was proposed by [61], which is the base of a modified method focused on estimating solar radiation from satellite imagery in Southeast Asia [38–40]. BRASIL-SR is a physical model that employed the two-stream approach to solve the radiative transfer equation. The information on cloud optical depth was obtained from satellite imagery. Incident global horizontal irradiance is obtained from

$$GHI = GHI_0[(\tau_c - \tau_{cloud})(1 - n) + \tau_{cloud}] \tag{4}$$

where GHI_0 is the irradiance at the top of the atmosphere, τ_c is the clear-sky transmittance, and τ_{cloud} is the overcast sky transmittance [51, 52]. This model was later modified to be adapted to the specific conditions of Chile creating a version named Chile-SR [24]. The latest version of Heliosat family, Heliosat-4, is a fully physical model composed of two models mostly based on look-up tables [78]: McClear model for solar irradiance under clear-sky and McCloud model for determining the extinction of irradiance due to clouds. The McClear model is the result of multiple radiative transfer computations with libRadtran [22, 23] for selected values of the inputs [47]. It basically consists of look-up tables and interpolation functions. The input data regarding aerosol optical depth, aerosol type, ozone column, and water vapor is obtained from CAMS (Copernicus Atmosphere Monitoring Service, https://atmosphere.copernicus.eu/) which is a freely available operation service evolved from the former MACC (Monitoring Atmospheric Composition and Climate) project. McCloud model is aimed at computing clear-sky index for both global and beam irradiance as a function of cloud properties. The inputs to McCloud are essentially the bidirectional reflectance distribution function (BDRF) derived from MODIS [6], and several cloud properties (cloud optical depth, cloud type, and cloud coverage) that are derived from satellite images using the APOLLO (AVHRR Processing scheme Over cLouds, Land and Ocean) scheme [45].

2.3 Satellite-Derived Solar Radiation Databases and Products

The high maturity level and large experience in methods for deriving solar radiation from satellite imagery can be measured by a large number of services and databases (both private and public) that are available nowadays. Table 1 lists some of the most well-known services for solar radiation data. Most of these databases are continuously evolving to updated or modified versions and expanding the spatial coverage by using additional satellite imagery.

2.4 Model Assessment and Uncertainties

Satellite-derived solar radiation methods have been being evaluated in parallel to their development. However, the assessment of models is limited to the availability of quality ground data which is still limited both in spatial distribution of ground station and in temporal coverage. The comparison of uncertainties in hourly satellite estimations with interpolation of neighboring ground station showed that satellite-derived information was more accurate than the use of 20–30 km distance stations [95]. A thorough assessment study of several satellite products and databases was performed under the framework of the Task 36 SHC-IEA (Solar Heating and Cooling program of the International Energy Agency) with quality ground data from several BSRN (BaseLine Surface Radiation Network) stations concluding that uncertainties

Table 1 Services and databases of satellite-derived solar radiation data

Name	Time basis	Coverage	Website
NASA SRB	3-hourly	World	http://gewex-srb.larc.nasa.gov/
DLR-ISIS	3-hourly	World	http://www.pa.op.dlr.de/ISIS/
HelioClim	hourly	Europe–Africa	http://www.soda-is.com/eng/helioclim/
SOLEMI	hourly	Europe–Africa–Asia	http://wdc.dlr.de/data-products/SERVICES/SOLARENERGY/
SolarGIS	30-min	World	http://solargis.info/
EnMetSol	hourly	Europe–Africa	https://www.uni-oldenburg.de/en/physics/research/ehf/energiemeteorology/enmetsol/
IrSOLaV	hourly	World	http://irsolav.com/
CM SAF (SARAH)	hourly	Europe–Africa	http://www.cmsaf.eu/
SolarAnywhere	30-min	North America	http://www.solaranywhere.com/
CAMS	15-min	World	http://atmosphere.copernicus.eu/catalogue/
PVGIS	hourly	Europe–Africa–Asia	http://re.jrc.ec.europa.eu/pvgis/
Vaisala	hourly	World	http://www.vaisala.com
Australian Bureau of Meteorology	hourly	Australia	http://www.bom.gov.au/climate/data-services/solar-information.shtml

for hourly values were around 17% for global and 34% for direct normal irradiance [37]. Nevertheless, the uncertainties may vary significantly among the sites and climatology.

Extensive validation of other satellite products is also found in the literature for lower temporal resolution (monthly or daily) due to the limitations of ground available data [83]. In this sense, comparisons among several datasets and numerical models also are available elsewhere [76]. In addition, sensitivity analysis has shown the most influencing input parameters in some models [26, 70]. Even though the uncertainties in the satellite-based models may vary among the different models, in general, difficulties come mostly from the treatment of clouds (radiative effect and optical properties are not completely well defined), aerosol optical depth, and other atmospheric constituents that input the model, terrain effects and high reflec-

tive albedo of deserts and snow, and spatial and temporal resolution [11, 12]. These sources of uncertainty produce systematic deviations, bias, or seasonal errors that cannot be avoided in most cases. In the last few years, several authors have proposed and developed methods for correcting bias or systematic errors in satellite retrievals with the use of short-term measurements at site [74]. Thus, with the help of a short period of ground measurements, it is possible to find a correction method to improve the accuracy (and consequently the bankability) of long-term time series of satellite-derived solar irradiance for a specific site.

3 Numerical Weather Prediction

Before the computer era, in the 1920s Richardson developed the first weather prediction model and had the dream that in the future numerical weather prediction models would be able to produce forecasts faster than the movement of the weather systems [49]. The first computer-based forecast was produced in 1950 and the biggest improvements came in the 1970s when computer capacity started to grow rapidly. There have been significant improvements over the past decades due to the increase of computer power, the improvement of the initial state by using many different types of observations, including remote sense observations and the scientific developments of the models. Although the equations describing the atmosphere are well known, they do not have analytic solution so they need to be approximated numerically. The level of complexity retained in the approximated equations depends on the purpose of the model.

The prediction process starts with an estimation of the initial state of the atmosphere using observations; this is known as analysis. The observations covering the globe are received by the processing centers almost in real time and, for operational weather prediction models, the integrations should be finished in a short period so that the predictions can be useful. The combination of analysis and forecast is often called a deterministic run of the model.

The error in the estimation of the initial state and the errors introduced by the different approximations can be amplified rapidly by the nonlinear equations describing the atmosphere. In other words, there is a limit in the predictability of the atmosphere that depends on the day. In order to estimate the predictability of an atmospheric state, ensemble approaches are used. This means that several integrations of the model are performed including perturbations of the initial state and also perturbing the model either by using different configurations, the model tendencies or using different boundary conditions. The results from the ensemble runs need to be interpreted in a probabilistic way.

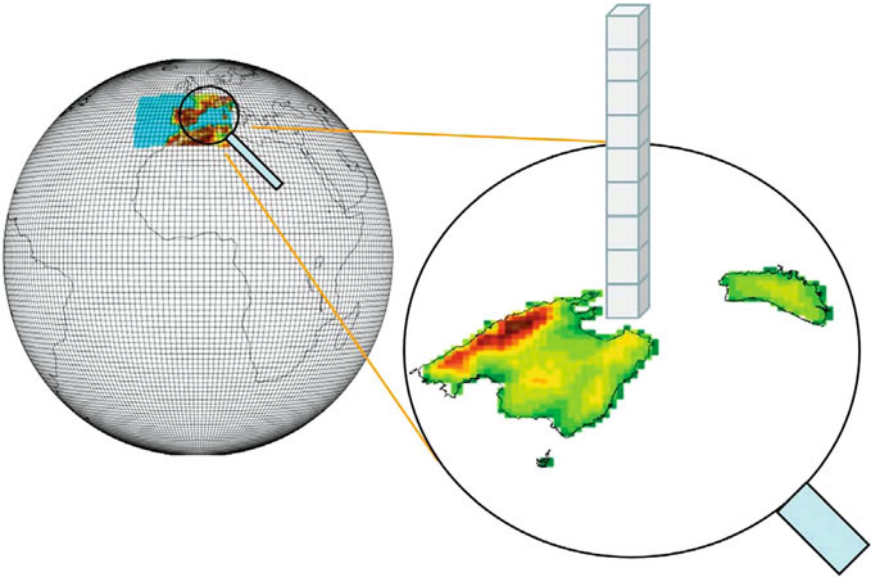


Fig. 14 The global model discretizes the atmosphere over the whole globe whereas Limited Area Model focus in a regional domain allowing an increase of the resolution. The calculations in the vertical are performed over independent discretized columns

3.1 Global and Regional Models

There are two big groups of atmospheric models, the global models covering the whole globe (Fig. 14) and the Limited Area Models (LAM), covering regional domains, which allow an increase of the resolution. Global models are needed for medium and longer forecast ranges. Besides, different target forecast lengths need different complexity of the processes. For instance, medium-range forecast models need more complex radiation parametrization than short range models because in the later the assimilation is able to correct the possible biases of the models. Another example is the need of a coupled oceanic model that is very important for long and seasonal ranges. Currently, there is a tendency to build seamless models capable of been run at different resolutions and for different purposes. This approach is achieved by including a few scale-dependent parameters.

In order to run global models operationally, big computer resources are needed and only a few processing centers have this capability. Among the global models with highest resolution are *National Centers for Environmental Prediction—Global Forecasting System* (NCEP-GFS, 13 km horizontal resolution), *United Kingdom Met Office—Unified Model* (UKMO-UM, 10 km) and the *European Center for Medium-Range Weather Forecasts—Integrated Forecasting System* (ECMWF-IFS, 9 km).

On the other hand, most LAM are currently run at convection-permitting scales using resolutions of 4 km and below. These models are Non-Hydrostatic and deep

convection is either treated explicitly or including a special parametrization to take into account that the parametrized and the resolved convection is not well separated. LAM tend to be run at regional scales and can include local data in their assimilation system as radar data. Usually, they wait less time for observation arrival because it is important to deliver their output rapidly in order to use them for warning purposes. In principle, these regional models improve the simulation of severe weather. One of the limitations of the LAM is that they need boundary conditions in the border of their domains that should be provided by a global model and this is always a source of error. For this reason, it is better to use large domains in order to have boundaries far from the area of interest and it is not worth to run LAMs for very long forecast lengths because after some time boundary conditions would dominate the model solution. Typical forecast lengths for LAMs are 48 h.

3.2 Data Assimilation

In order to launch the model, an estimation of the initial state of the atmosphere is needed. A huge amount of observations spread over the globe are used for this purpose. These observations are distributed almost in real time or with short delivery times, and they are exchanged between the different processing centers through a *Global Telecommunication System* (GTS). A schematic view of the observations used by the models can be seen in Fig. 6. Several times per day, sounding balloons are launched; many different types of satellites, either geostationary or polar, provide huge amount of data from different sensors; aircrafts provide observations during their fly; and there is a dense surface network whose observations are shared through the GTS.

Despite the huge amount of observations, they are not enough to estimate the initial state of the atmosphere in each grid point of the model. The models discretize the atmosphere in thousands of small elements going from the surface to the stratosphere and the value of the atmospheric variables in all these elements should be set. The method used to solve this problem is to use a previous short-term integration from the model as *first guess* and then correct it using the available observations. Data assimilation is the process of combining statistically the observations with the first guess, producing an *analysis* that will be used to initiate the model [18, 43]. The big amount of different observations needing specific treatment and quality control and the complex process of combining properly all the information they provide with the model data make the assimilation processes a very complex issue. It should be taken into account that observations are irregularly distributed and that many observations are not in model space (for instance, satellite radiances) so specific observatory operators should be constructed in order to transform from model space to observation space. In the analysis process, the statistics of model errors and observation errors should be taken into account.

Operative atmospheric models usually have implemented one of the two following methods:

3DVar method Many models, specially LAM, use this method in the assimilation process. The analysis is started after some waiting time in order to collect the observations (cutoff time). This cutoff time depends on the purpose of the model. Medium-range forecast models may wait six hours for the observations whereas nowcasting models may just wait 10 min. In order to estimate the atmospheric state, a cost function, $J(x)$, is minimized

$$J(x) = \underbrace{\left(x - x^b\right) B^{-1} \left(x - x^b\right)}_{J_B} + \underbrace{\left(y_0 - H(x)\right) R^{-1} \left(y_0 - H(x)\right)}_{J_0} \tag{5}$$

where
y = observation vector,
x = model state vector,
H = observation operator,
B = covariance matrix for the first guess errors, and
R = covariance matrix for the observation errors.

The cost function relates the distance between the first guess and the model state weighted with the precision of the first guess (B) plus the distance between the observations and the model state weighted with the precision of the observations (R). A problem is that the B matrix is static, so the same one is used for all the atmospheric conditions, and it is known that the model errors depend on the atmospheric flow. In order to assimilate more observations and that these are close to the analysis time, frequent assimilation cycles are performed (Fig. 15).

4DVar method It is a more accurate method because the observations can be used at their exact time. An optimization processes is performed in which the model is run back and forth in time in order to obtain a forecast which adjust optimally to the observations. The final analysis is the one that achieves the better adjustment. Forward integrations are performed with the complete model whereas backward integrations are performed with a simplified linearized version of the model. Many global models use 4DVar in their assimilation system.

3.3 Dynamics

The equations describing the atmospheric flow are well known [13]. Generally, the equations used are the Euler equations, the thermodynamic equation (energy conservation), the continuity equation (mass conservation), and the perfect gas law. These equations are quite general and are capable of describing a great variety of scales and atmospheric processes including non-hydrostatic processes. In order to describe moist processes, other conservation equations are added for water vapor, liquid, and ice water and for different precipitation species. Depending on the purpose of the model, different approximations are considered as the hydrostatic approximation.

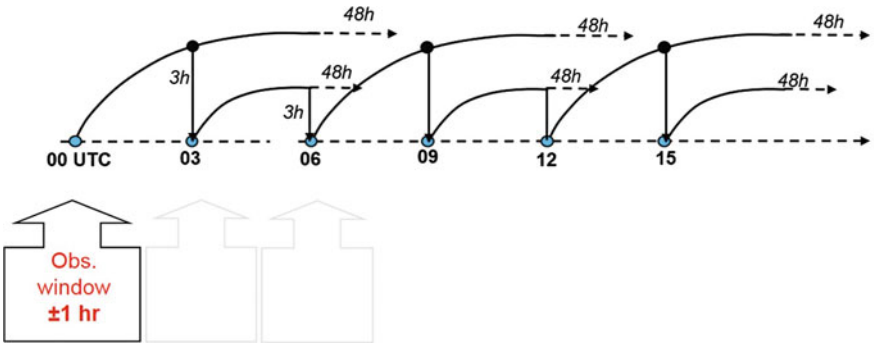


Fig. 15 Example of 3DVar assimilation cycles every 3 h. All the observations performed within a time window are considered valid for the analysis time. The analysis is built using the first guess from the previous cycle

The primitive system of equations The equations generally used by global models are the primitive equations that simplify the Euler equations taken into account the hydrostatic approximation (see [13]),

$$\frac{d\mathbf{V}_h}{dt} = -fk \times \mathbf{V}_h - \frac{RT}{p}\mathbf{C}p \tag{6}$$

$$\frac{dT}{dt} = \frac{RT}{C_p p}\frac{dp}{dt} \tag{7}$$

$$\frac{dp}{dt} = \frac{p}{1 - R/C_p}\left(V \cdot \mathbf{V}_h + \frac{\partial w}{\partial z}\right) \tag{8}$$

$$\frac{\partial p}{\partial z} = \frac{p}{RT}g \tag{9}$$

where the temporal derivatives (prediction variables) allow the estimation of the evolution of the atmosphere. With the hydrostatic approximation, the vertical velocity is not a prediction variable, and the pressure is estimated using the hydrostatic equation.

Non-hydrostatic equations Most regional models are currently operating at resolutions of a few kilometers. At these scales, hydrostatic hypothesis is no longer valid. There are different approaches to take into account non-hydrostatic (NH) effects [89]. Some widely used are:

- Laprise method [46].
- Anelastic approximation. The atmospheric variables are made function of basic hydrostatic state and deviations from this state. Density fluctuations are neglected except in the buoyancy terms [36].

In the Laprise method, the continuity equation [13] for a general vertical coordinate can be written as

$$\frac{\partial}{\partial t}\left(\rho\frac{\partial z}{\partial s}\right) = -div_s\left(\rho\frac{\partial z}{\partial s}\mathbf{V}\right) - \frac{\partial}{\partial s}\left(\rho\frac{\partial z}{\partial s}\frac{\partial s}{\partial t}\right) \tag{10}$$

Choosing a coordinate, s, for which $\rho\frac{\partial z}{\partial s}$ is constant, the continuity equation will take a diagnostic form. A coordinate with this property is the hydrostatic pressure

$$\frac{\partial \pi}{\partial s} = -\rho g \tag{11}$$

Using π as vertical coordinate, the Euler equations, which include compressibility effects, have a similar form as the primitive equation in p coordinates. The algorithms and methods used to solve primitive equations can be used for the non-hydrostatic case. In the NH equations, the pressure is a prognostic variable as is the vertical velocity. For stability reasons, instead of using p and w as prognostic variables, the following variables are used:

$$P = \frac{(p-\pi)}{\pi} \tag{12}$$

$$d = -\frac{gp}{RT}\left(\frac{\partial w}{\partial z} - \nabla\Phi\frac{\partial s}{\partial \pi}\right) \tag{13}$$

where Φ is the geopotential. This method [8] is used in ALADIN, AROME, and currently in the NH version of the IFS model. There are many different methods to solve these equations depending on the selected geometry, the vertical coordinate, and the spatial and temporal discretization methods.

Finite differences method It is frequently used for spatial and temporal discretization of the atmospheric equations. The partial derivatives of a function, $f(x, y, z, t)$ in a point are computed using the values of the function in the neighbor points.

$$\frac{\partial f}{\partial x} = \frac{f(x+\Delta x) - f(x-\Delta x)}{\Delta x} \tag{14}$$

where Δx is the length between grid points.

Spectral method Instead of using the value of the model variables in the computation grid, the variables are expressed as a linear combination of some basic functions chosen to facilitate the resolution of the differential equations. The spectral method tends to be applied in the horizontal, and the functions chosen are the spherical harmonics. Any variable could be expressed in the following form:

$$f(\lambda, \mu, t) = \sum_{n,m} a_{n,m} Y_{n,m}((\lambda, \mu) \tag{15}$$

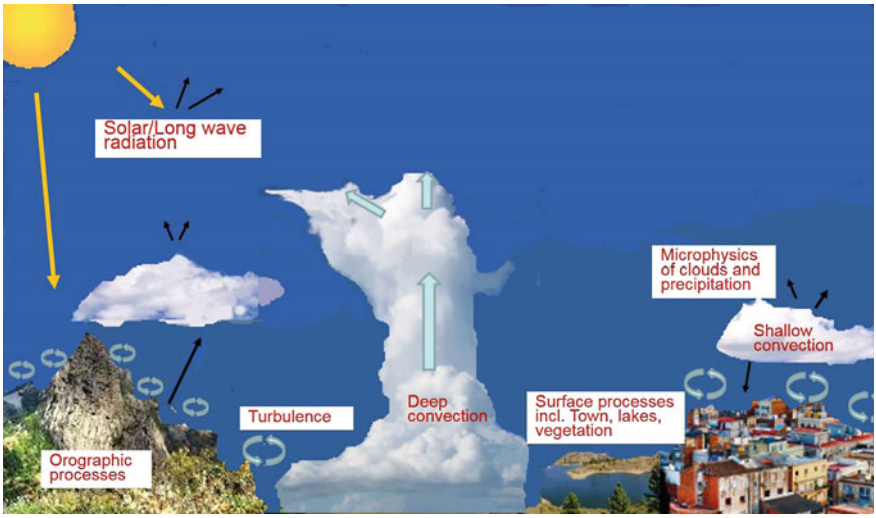


Fig. 16 Processes parameterized in most atmospheric numerical weather prediction models

For the actual computations, a finite number of basic functions should be retained what it is known as model truncation. For a model with triangular truncation T1279, the maximum wave number would be 1279. The spectral approach is used by most global models. Although less common, LAM can also use the spectral approximation. For this purpose, a halo should be added to the domain of interest in order to achieve bi-periodization.

3.4 Parameterization of Physical Processes

The processes that occur at spatial or temporal scales not solved by the model dynamics should be treated in an approximated manner. The goal is to represent the collective effects of these unresolved processes in the variables resolved by the model, something that is known as physical parameterizations [88]. The type of processes and the complexity that need to be included depend on the scale and purpose of the model. Some processes as the radiation are also parameterized (simplified) because they are too complex for being treated in full detail. The main physical parameterizations included in atmospheric models are turbulence, radiation, surface processes, large-scale clouds, convective clouds, and orographic processes (Fig. 16). Generally, the effects of the different processes are computed individually using the information in a model column (one-dimensional parameterizations), and then the different tendencies produced are added to the model equations.

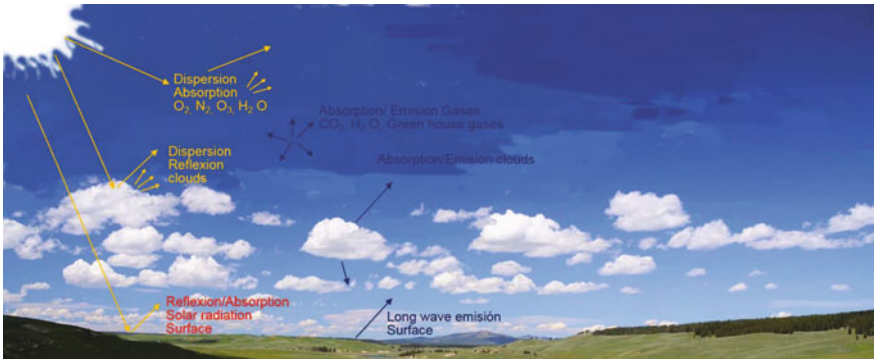


Fig. 17 Radiation parametrization. The computations are divided into short wave (yellow) and long wave (blue) emitted by the earth surface, the atmosphere, and the clouds. The biggest uncertainty comes from the cloud prediction

Radiation It estimates the radiative fluxes between the different model levels and the surface. The solar radiation entering the atmosphere is partially absorbed by atmospheric gases, mainly oxygen, nitrogen, and ozone, and also is dispersed by the atmospheric gases and the aerosols. The clouds produce diffusion and reflection of the radiation and absorb a small part of the solar radiation. Part of the radiation that finally reaches the surface is absorbed whereas the rest is reflected back to the atmosphere (depending on the surface albedo).

Besides, the surface, the atmospheric gases, and the clouds emit long-wave radiation depending on their temperature and emissivity. In contrast with solar radiation, most part of the long-wave radiation is absorbed by the atmosphere mainly due to the carbon dioxide, the water vapor, and the ozone.

The radiative transfer equations are well known and the radiative fluxes could be computed accurately but with a computer cost that it is not feasible for an operational model. The goal of the radiation parametrization is to find efficient approximations. Among the atmospheric particles that influence in the radiation computations, the water vapor and the clouds are predicted by the model whereas for other gases, climatological values are considered. Currently, there is a tendency to include ozone and aerosols as forecast variables initialized in the assimilation process.

The radiation computations are divided into short wave and long wave (Fig. 17). Besides only the main spectral bands of absorption and emission are considered. Despite the approximations, the errors in clear sky are small, and the biggest source of error in the radiation prediction comes from the errors in cloud forecasting.

Orographic processes Mountains exert a significant influence on the atmospheric flow. In general, models use a mean orography that depends on its horizontal resolution, underestimating the height of the orographic obstacles and the terrain irregularities. Many models parameterize the orographic effects which are not resolved explicitly such as the orographic gravity waves, the blocking effect, and

small-scale turbulence produced by the terrain variability. As model resolution increases, the mesoscale orographic effects are resolved by the model, and generally below 4 km, only the small-scale turbulence needs to be parameterized.

Surface processes The surface parameterization estimates heat, humidity, and momentum fluxes, which are boundary conditions for the atmosphere. These fluxes have a significant impact on important meteorological variables as 2 m temperature and 10 m wind, and module other atmospheric processes such as turbulence, radiation, and clouds. Surface processes play an important role in climatic models that require more sophisticated schemes than operational models. In order to compute the fluxes, the schemes need to compute the evolution of the temperature and humidity in the different possible soil and vegetation types (Fig. 18). The models include also assimilation for the surface variables using satellite data and surface observations or inferring the surface values from the screen level observations.

Turbulence The lower part of the troposphere is strongly influenced by the earth surface and responds rapidly to the surface drag and the heat and moisture fluxes. The turbulent eddies generated by these exchanges condition significantly the atmospheric evolution. Near the ground, the wind shows a logarithmic profile crossing the isobars toward lower pressures, and above the surface, the influence decreases. In order to estimate the surface fluxes and the profiles of the variables near the surface, the Monin–Obukhov Similarity theory is generally used. Above the surface layer, a method widely used is the turbulent kinetic energy parametrization.

Besides, the fluxes are functions of the atmospheric stability. During the night or when the surface temperature is lower than the atmospheric temperature, the boundary layer stabilizes producing only small eddies and small turbulent fluxes. On the other hand, during the day or with a surface warmer than the atmosphere, big eddies are generated that can modify the atmosphere up to several hundreds of meters. When the wind is strong, also big turbulent fluxes are generated due to the shear generated by the surface drag. Overall, the biggest parametrization errors are produced in stable situations. This is the reason why fog is very difficult to represent in the models.

Clouds Clouds are one of the most important atmospheric phenomena but its simulation limits the atmospheric predictability. Atmospheric models split the cloud representation into two parts:

Clouds in layers. Also called large-scale or stratiform clouds. They include the large atmospheric systems that take place at scales resolved by the model so they can be simulated by the model dynamics. Nonetheless, the microphysical processes that module the different condensation processes and the evolution of the water species need to be parametrized (microphysics parametrization, Fig. 19).

Convective clouds. An example is the thunderstorm clouds. They are generated because the atmosphere is unstable so vertical currents are generated that tend to stabilize the atmosphere and may produce big precipitation amounts in a short time period. Convection plays a very important role in the atmospheric flow, producing big heat, moisture, and momentum fluxes. Besides, convection develops in

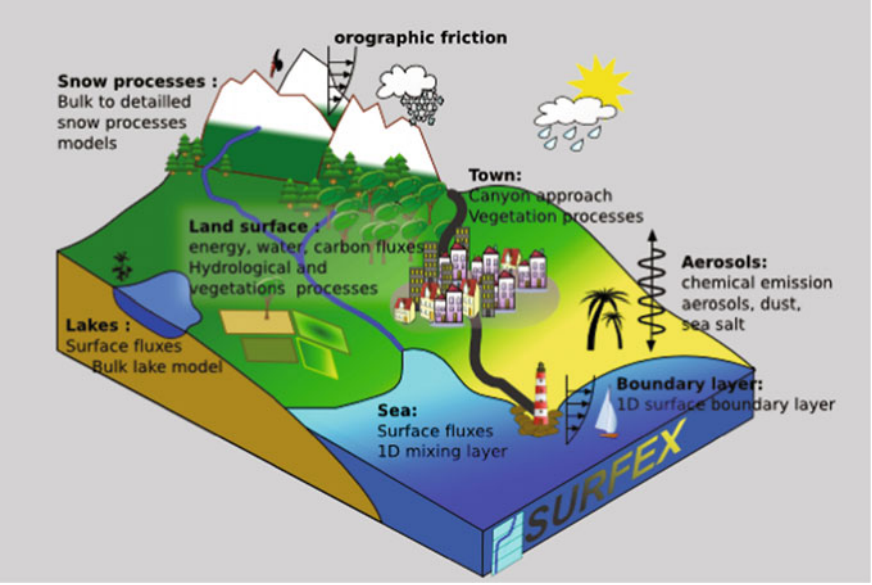


Fig. 18 Components of the surface scheme SURFEX which includes different models for different surfaces types [53]. *Source* CNRM: http://www.umr-cnrm.fr/surfex

a great variety of scales, from shallow cumulus to hurricanes. Most models parameterize convection using a mass flux approach. In this type of parametrization, the convective circulations are simplified considering the updrafts, the downdrafts associated with precipitation evaporation, and the subsidence in the cloud environment. The convection consumes the Convective Available Potential Energy

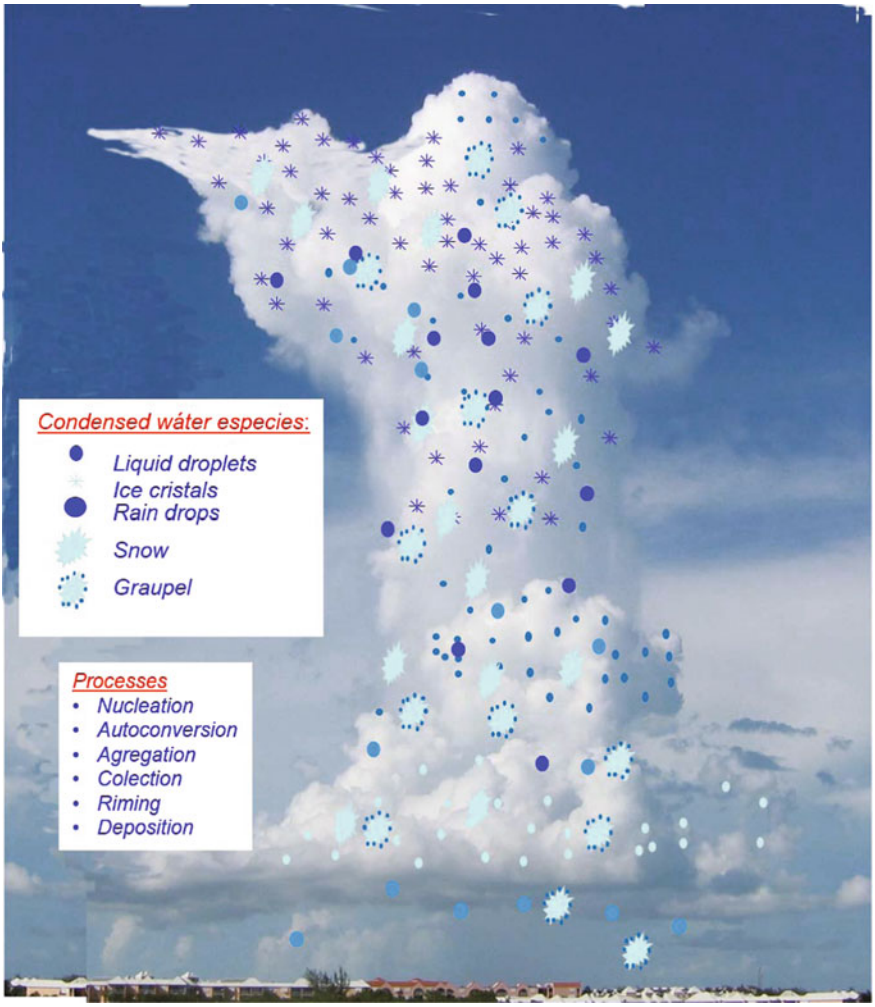


Fig. 19 Microphysical processes parameterized in the AROME model [86]

(CAPE) and may produce very intense and localized precipitation. There are two possible approaches of convective clouds depending on the resolution of the model:

- Convection-permitting models. As model resolution increases, the deep convective clouds start being resolved by the model, and generally, deep convection parameterization is deactivated so that convective clouds are represented explicitly by the dynamics and the microphysics. At what resolution convective clouds are resolved is not clear but probably, at least 1 km resolution is needed. Nonetheless, many models already switch off deep convection parameterization at 4 km resolution.

- Grey convection scales. In the range 2 to 10 km resolution, it can be considered that deep convection is partially resolved by the model but not completely. For these scales, a few convection schemes have been developed that are able to take into account this effect (see, for instance, [30]). These parameterizations are much more complex than the ones used in larger scales and should include new prognostic variables such as the convective cloud fraction and the vertical velocity in the convective currents.

3.5 Modeling Consortia

Numerical Weather Prediction models are the main tool for weather prediction and they have improved significantly over the past decades. Development and maintenance of these complex models represent a huge effort and in order to achieve this, the international collaboration is a key aspect. An example of this collaboration is the ECMWF which is an international independent organization funded by 34 countries and which has a clear leadership in medium-range weather forecast. Another way to strengthen the collaboration in NWP is by creating international Consortia. There are several such Consortia in Europe devoted mainly to LAM developing for operational use. ALADIN and HIRLAM are examples of such Consortia (http://hirlam.org, http://www.umr-cnrm.fr/aladin)

References

1. Part I. Measurements of meteorological variables (2008). In: Guide to meteorological instruments and methods of observation, chap. 5. World Meteorological Organization
2. Global aspects. annex v to the wmo technical regulations (2015). In: Manual on the global observing system, vol. I. World Meteorological Organization
3. Amillo AG, Huld T, Müller R (2014) A new database of global and direct solar radiation using the eastern meteosat satellite, models and validation. Remote Sens 6(9):8165–8189
4. Beyer HG, Costanzo C, Heinemann D (1996) Modifications of the heliosat procedure for irradiance estimates from satellite images. Sol Energy 56(3):207–212
5. Blanc P, Espinar B, Geuder N, Gueymard C, Meyer R, Pitz-Paal R, Reinhardt B, Renné D, Sengupta M, Wald L et al (2014) Direct normal irradiance related definitions and applications: the circumsolar issue. Sol Energy 110:561–577
6. Blanc P, Gschwind B, Lefevre M, Wald L (2014) Twelve monthly maps of ground albedo parameters derived from modis data sets. In: IEEE international geoscience and remote sensing symposium (IGARSS). IEEE, pp 3270–3272
7. Brock FV, Richardson SJ (2001) Meteorological measurement systems. Oxford University Press
8. Bubnová R, Hello G, Bénard P, Geleyn JF (1995) Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the arpege/aladin nwp system. Mon Weather Rev 123(2):515–535
9. Busch N, Kristensen L (1976) Cup anemometer overspeeding. J Appl Meteorol 15:1328–1332

10. Cano D, Monget JM, Albuisson M, Guillard H, Regas N, Wald L (1986) A method for the determination of the global solar radiation from meteorological satellite data. Sol Energy 37(1):31–39
11. Cebecauer T, Suri M (2010) Accuracy improvements of satellite-derived solar resource based on gems re-analysis aerosols. In: SolarPACES conference
12. Cebecauer T, Suri M, Gueymard C (2011) Uncertainty sources in satellite-derived direct normal irradiance: how can prediction accuracy be improved globally. In: Proceedings of the SolarPACES conference, Granada, Spain, vol 2023
13. Coiffier J (2011) Fundamentals of numerical weather prediction. Cambridge University Press
14. Coppin P (1982) An examintaion of cup anemometer overspeeding. Meteorolog Rundsch 35:1–11
15. Cros S, Albuisson M, Wald L (2006) Simulating meteosat-7 broadband radiances using two visible channels of meteosat-8. Sol Energy 80(3):361–367
16. Dagestad KF (2004) Mean bias deviation of the heliosat algorithm for varying cloud properties and sun-ground-satellite geometry. Theor Appl Climatol 79(3):215–224
17. Dagestad KF, Olseth JA (2007) A modified algorithm for calculating the cloud index. Sol Energy 81(2):280–289
18. Daley R (1993) Atmospheric data analysis. Cambridge University Press, Cambridge Atmospheric and Space Science Series
19. Diabaté L, Demarcq H, Michaud-Regas N, Wald L (1987) Estimating incident solar radiation at the surface from images of the earth transmitted by geostationary satellites: the heliosat project. Int J Sol Energy 5(5–6):261–278
20. Djebbar R, Morris R, Thevenard D, Perez R, Schlemmer J (2012) Assessment of suny version 3 global horizontal and direct normal solar irradiance in canada. Energy Procedia 30:1274–1283
21. Dobson F (1981) Review of reference height for and averaging time of surface wind measurements at sea. Technical Report 3, World Meteorological Organization. Marine Meteorology and Related Oceanographic Activities
22. Emde C, Buras-Schnell R, Kylling A, Mayer B, Gasteiger J, Hamann U, Kylling J, Richter B, Pause C, Dowling T et al (2015) The libradtran software package for radiative transfer calculations (version 2.0). Geosci Model Dev Discuss 8(12):10,237–10,303
23. Emde C, Buras-Schnell R, Kylling A, Mayer B, Gasteiger J, Hamann U, Kylling J, Richter B, Pause C, Dowling T et al (2016) The libradtran software package for radiative transfer calculations (version 2.0. 1). Geosci Model Dev 9(5):1647–1672
24. Escobar RA, Cortés C, Pino A, Pereira EB, Martins FR, Cardemil JM (2014) Solar energy resource assessment in chile: satellite estimation and ground station measurements. Renew Energy 71:324–332
25. Espinar B, Blanc P, Wald L, Hoyer-Klick C, Homscheidt MS, Wanderer T (2012) On quality control procedures for solar radiation and meteorological measures, from subhourly to montly average time periods. In: EGU general assembly
26. Espinar B, Ramirez L, Polo J, Zarzalejo L, Wald L (2009) Analysis of the influences of uncertainties in input variables on the outcomes of the heliosat-2 method. Sol Energy 83(9):1731–1741
27. Fiebrich CA, Morgan CR, McCombs AG, Hall PK Jr, McPherson RA (2010) Quality assurance procedures for mesoscale meteorological data. J Atmos Ocean Technol 27(10):1565–1582
28. Garg H, Garg S (1993) Measurement of solar radiationi. radiation instruments. Renew Energy 3(4–5):321–333
29. Gautier C, Diak G, Masse S (1980) A simple physical model to estimate incident solar radiation at the surface from goes satellite data. J Appl Meteorol 19(8):1005–1012
30. Gerard L (2015) Bulk mass-flux perturbation formulation for a unified approach of deep convection at high resolution. Mon Weather Rev 143(10):4038–4063
31. Gill G, Olsson L, Sela J, Suda M (1967) Accuracy of wind measurements on towers or stacks. Bull Am Meteorol Soc 48:665–674
32. Gold E (1936) Wind in britain: the dines anemometer and some notable records during the last 40 years. Q J Roy Meteorol Soc 62:167–206

33. Hammer A, Heinemann D, Hoyer C, Kuhlemann R, Lorenz E, Müller R, Beyer HG (2003) Solar energy assessment using remote sensing technologies. Remote Sens Environ 86(3):423–432
34. Hoyer-Klick C, Dumortier D, Tsvetkov A, Polo J, Torres J, Kurz C, Ineichen P (2009) Mesor existing ground data sets
35. Huld T, Müller R, Gambardella A (2012) A new solar radiation database for estimating pv performance in europe and africa. Sol Energy 86(6):1803–1815
36. Ikawa M (1988) Comparison of some schemes for nonhydrostatic models with orography. J Meteorol Soc Jpn Ser. II 66(5):753–776
37. Ineichen P (2014) Long term satellite global, beam and diffuse irradiance validation. Energy Procedia 48:1586–1596
38. Janjai S, Laksanaboonsong J, Nunez M, Thongsathitya A (2005) Development of a method for generating operational solar radiation maps from satellite data for a tropical environment. Sol Energy 78(6):739–751
39. Janjai S, Masiri I, Laksanaboonsong J (2013) Satellite-derived solar resource maps for myanmar. Renew Energy 53:132–140
40. Janjai S, Pankaew P, Laksanaboonsong J, Kitichantaropas P (2011) Estimation of solar radiation over cambodia from long-term satellite data. Renew Energy 36(4):1214–1220
41. Journée M, Bertrand C (2011) Quality control of solar radiation data within the rmib solar measurements network. Sol Energy 85(1):72–86
42. Kaimal J (1980) Sonic anemometers. In: Dobson LHF, Davis R (eds) Air-Sea interaction: instruments and methods. New York, pp 81–96
43. Kalnay E (2003) Atmospheric modeling, data assimilation and predictability. Cambridge University Press
44. König-Langlo G, Sieger R, Schmithüsen H, Bücker A, Richter F, Dutton E (2013) The baseline surface radiation network and its world radiation monitoring centre at the alfred wegener institute
45. Kriebel K, Gesell G, Kästner M, Mannstein H (2003) The cloud analysis tool apollo: improvements and validations. Int J Remote Sens 24(12):2389–2408
46. Laprise R (1992) The euler equations of motion with hydrostatic pressure as an independent variable. Mon Weather Rev 120(1):197–207
47. Lefevre M, Oumbe A, Blanc P, Espinar B, Gschwind B, Qu Z, Wald L, Homscheidt MS, Hoyer-Klick C, Arola A et al (2013) Mcclear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. Atmos Measure Techn 6:2403–2418
48. Long CN, Shi Y (2008) An automated quality assessment and control algorithm for surface radiation measurements. Open Atmos Sci J 2(1)
49. Lynch P (2006) The emergence of numerical weather prediction: richardson's dream. Cambridge University Press (2006)
50. MacCready P (1966) Mean wind speed measurements in turbulence. J Appl Meteorol 5:219–225
51. Martins FR, Pereira EB, Abreu S (2007) Satellite-derived solar resource maps for brazil under swera project. Sol Energy 81(4):517–528
52. Martins FR, Pereira EB, Abreu SL, Beyer HG, Colle S, Perez R, Heinemann D (2003) Cross validation of satellite radiation transfer models during swera project in brazil. In: Proceedings of ISES solar world congress
53. Masson V, Le Moigne P, Martin E, Faroux S, Alias A, Alkama R, Belamari S, Barbu A, Boone A, Bouyssel F, Brousseau P, Brun E, Calvet JC, Carrer D, Decharme B, Delire C, Donier S, Essaouini K, Gibelin AL, Giordani H, Habets F, Jidane M, Kerdraon G, Kourzeneva E, Lafaysse M, Lafont S, Lebeaupin Brossier C, Lemonsu A, Mahfouf JF, Marguinaud P, Mokhtari M, Morin S, Pigeon G, Salgado R, Seity Y, Taillefer F, Tanguy G, Tulet P, Vincendon B, Vionnet V, Voldoire A (2013) The surfexv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geosci Model Dev 6(4):929–960. https://doi.org/10.5194/gmd-6-929-2013. https://www.geosci-model-dev.net/6/929/2013/
54. Maxwell E, Wilcox S, Rymes M (1993) Users manual for seri qc software, assessing the quality of solar radiation data. Solar Energy Research Institute, Golden, CO. http://www.osti.gov/bridge

55. Mazarella D (1972) An inventory of specifications for wind-measuring instruments. Bull Am Meteorol Soc 53:860–871
56. Möser W, Raschke E (1984) Incident solar radiation over europe estimated from meteosat data. J Clim Appl Meteorol 23(1):166–170
57. Mueller R, Dagestad KF, Ineichen P, Schroedter-Homscheidt M, Cros S, Dumortier D, Kuhlemann R, Olseth J, Piernavieja G, Reise C et al (2004) Rethinking satellite-based solar irradiance modelling: the solis clear-sky module. Remote Sens Environ 91(2):160–174
58. Mueller R, Pfeifroth U, Traeger-Chatterjee C (2015) Towards optimal aerosol information for the retrieval of solar surface radiation using heliosat. Atmosphere 6(7):863–878
59. Noia M, Ratto C, Festa R (1993) Solar irradiance estimation from geostationary satellite data: I. statistical models. Sol Energy 51(6):449–456
60. Noia M, Ratto C, Festa R (1993) Solar irradiance estimation from geostationary satellite data: Ii. physical models. Sol Energy 51(6):457–465
61. Nunez M (1993) The development of a satellite-based insolation model for the tropical western pacific ocean. Int J Climatol 13(6):607–627
62. Peppler RA, Long C, Sisterson D, Turner D, Bahrmann C, Christensen SW, Doty K, Eagan R, Halter T, Iveyh M et al (2008) An overview of arm program climate research facility data quality assurance. Open Atmos Sci J 2(1):192–216
63. Perez R (2002) Time specific irradiances derived from geostationary satellite images. J Sol Energy Eng 124(1):1–1
64. Perez R, Cebecauer T, Šúri M (2013) Semi-empirical satellite models. Solar energy forecasting and resource assessment, Academic Press, Oxford, UK, pp 21–48
65. Perez R, Ineichen P, Kmiecik M, Moore K, Renne D, George R (2004) Producing satellite-derived irradiances in complex arid terrain. Sol Energy 77(4):367–371
66. Perez R, Ineichen P, Moore K, Kmiecik M, Chain C, George R, Vignola F (2002) A new operational model for satellite-derived irradiances: description and validation. Sol Energy 73(5):307–317
67. Perez R, Schlemmer J, Hemker K, Kivalov S, Kankiewicz A, Gueymard C (2015) Satellite-to-irradiance modeling-a new version of the suny model. In: IEEE 42nd photovoltaic specialist conference (PVSC), IEEE, pp 1–7
68. Perez R, Schlemmer J, Renne D, Cowlin S, George R, Bandyopadhyay B (2009) Validation of the suny satellite model in a meteosat environment. In: Proceeding of ASES annual conference, Buffalo, New York
69. Pinker R, Laszlo I (1991) Effects of spatial sampling of satellite data on derived surface solar irradiance. J Atmos Ocean Technol 8(1):96–107
70. Polo J, Antonanzas-Torres F, Vindel J, Ramirez L (2014) Sensitivity of satellite-based methods for deriving solar radiation to different choice of aerosol input and models. Renew Energy 68:785–792
71. Polo J, Martin L, Cony M (2012) Revision of ground albedo estimation in heliosat scheme for deriving solar radiation from seviri hrv channel of meteosat satellite. Sol Energy 86(1):275–282
72. Polo J, Martín L, Vindel J (2015) Correcting satellite derived dni with systematic and seasonal deviations: application to india. Renew Energy 80:238–243
73. Polo J, Vindel J, Martín L (2013) Angular dependence of the albedo estimated in models for solar radiation derived from geostationary satellites. Sol Energy 93:256–266
74. Polo J, Wilbert S, Ruiz-Arias JA, Meyer R, Gueymard C, Suri M, Martín L, Mieslinger T, Blanc P, Grant I et al (2016) Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. Sol Energy 132:25–37
75. Polo J, Zarzalejo LF, Ramírez L (2008) Solar radiation derived from satellite images. In: Modeling solar radiation at the earth's surface, pp 449–462
76. Posselt R, Mueller R, Stöckli R, Trentmann J (2012) Remote sensing of solar surface radiation for climate monitoringthe cm-saf retrieval in international comparison. Remote Sens Environ 118:186–198
77. Posselt R, Müller R, Trentmann J, Stockli R, Liniger MA (2014) A surface radiation climatology across two meteosat satellite generations. Remote Sens Environ 142:103–110

78. Qu Z, Oumbe A, Blanc P, Espinar B, Gesell G, Gschwind B, Klüser L, Lefèvre M, Saboret L, Schroedter-Homscheidt M et al (2017) Fast radiative transfer parameterisation for assessing the surface solar irradiance: the heliosat-4 method. Meteorol Z 26(1):33–57
79. Retallack B (1984) In: Compendium of lecture notes for training class IV meteorological personnel, vol II, World Meteorological Organization
80. Rigollier C, Lefèvre M, Wald L (2004) The method heliosat-2 for deriving shortwave solar radiation from satellite images. Sol Energy 77(2):159–169
81. Riihelä A, Carlund T, Trentmann J, Müller R, Lindfors AV (2015) Validation of cm saf surface solar radiation datasets over finland and sweden. Remote Sens 7(6):6663–6682
82. Roesch A, Wild M, Ohmura A, Dutton EG, Long CN, Zhang T (2011) Assessment of bsrn radiation records for the computation of monthly means. Atmos Measure Tech 4(2):339–354
83. Sanchez-Lorenzo A, Wild M, Trentmann J (2013) Validation and stability assessment of the monthly mean cm saf surface solar radiation dataset over europe against a homogenized surface dataset (1983–2005). Remote Sens Environ 134:355–366
84. Schillings C, Meyer R, Mannstein H (2004) Validation of a method for deriving high resolution direct normal irradiance from satellite data and application for the arabian peninsula. Sol Energy 76(4):485–497
85. Schwandt M, Chhatbar K, Meyer R, Mitra I, Vashistha R, Giridhar G, Gomathinayagam S, Kumar A (2014) Quality check procedures and statistics for the indian srra solar radiation measurement network. Energy Procedia 57:1227–1236
86. Seity Y, Brousseau P, Malardel S, Hello G, Bénard P, Bouttier F, Lac C, Masson V (2011) The arome-france convective-scale operational model. Mon Weather Rev 139(3):976–991
87. Sengupta M, Habte A, Kurtz S, Dobos A, Wilbert S, Lorenz E, Stoffel T, Renné D, Gueymard C, Myers D et al (2015) Best practices handbook for the collection and use of solar resource data for solar energy applications. NREL
88. Stensrud D (2009) Parameterization schemes: keys to understanding numerical weather prediction models. Cambridge University Press
89. Steppeler J, Hess R, Schättler U, Bonaventura L (2003) Review of numerical methods for nonhydrostatic weather prediction models. Meteorol Atmos Phys 82(1):287–301
90. Vignola F, Grover C, Lemon N, McMahan A (2012) Building a bankable solar radiation dataset. Sol Energy 86(8):2218–2229
91. Walmsley J, Troen I, Lalas D, Mason P (1990) Surface-layer flow in complex terrain: comparison of models and full-scale observation. Bound Layer Meteorol 52:259–281
92. Wieringa J (1980) A revaluation of the kansas mast influence on measurements of stress and cup anemometer overspeeding. Bound Layer Meteorol 18:411–430
93. Wieringa J (1983) Description requirements for assessment of non-ideal wind stations, for example aachen. J Wind Eng Ind Aerodyn 11:121–131
94. Zarzalejo LF, Polo J, Martín L, Ramírez L, Espinar B (2009) A new statistical approach for deriving global solar radiation from satellite images. Sol Energy 83(4):480–484
95. Zelenka A, Perez R, Seals R, Renné D (1999) Effective accuracy of satellite-derived hourly irradiances. Theor Appl Climatol 62(3):199–207

Characterization of Geographical and Meteorological Parameters

Gustavo Montero, Eduardo Rodríguez and Albert Oliver

Abstract This chapter is devoted to the introduction of some geographical and meteorological information involved in the numerical modeling of wind fields and solar radiation. First, a brief description of the topographical data given by a Digital Elevation Model and Land Cover databases is provided. In particular, the Information System of Land Cover of Spain (SIOSE) is considered. The study is focused on the roughness length and the displacement height parameters that appear in the logarithmic wind profile, as well as in the albedo related to solar radiation computation. An extended literature review and characterization of both parameters are reported. Next, the concept of atmospheric stability is introduced from the Monin–Obukhov similarity theory to the recent revision of Zilitinkevich of the Neutral and Stable Boundary Layers (SBL). The latter considers the effect of the free-flow static stability and baroclinicity on the turbulent transport of momentum and of the Convective Boundary Layers (CBL), more precisely, the scalars in the boundary layer, as well as the model of turbulent entrainment.

1 Geographical Data

The main geographical information for wind and solar radiation modeling may be classified into two general databases, the topographical data related to the orography of the region to be studied and the land cover databases containing the information of the land uses. In this section, both are introduced.

G. Montero (✉) · E. Rodríguez · A. Oliver
University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain
e-mail: gustavo.montero@ulpgc.es

E. Rodríguez
e-mail: eduardo.rodriguez@ulpgc.es

A. Oliver
e-mail: albert.oliver@ulpgc.es

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_2

1.1 Topographical Information

To study the orography of a surface, it is usual to start from a Digital Elevation Model (DEM) that contains elevation data on a uniform grid (height map). For example, the National Geographic Institute of Spain provides a 25 m $\times$ 25 m grid with a precision of 5 m in height for all the national territory (MDT25). In Spain, the geodetic Cartesian reference frame used is the European Terrestrial Reference System 1989 (ETRS89) in the Peninsula, Balearic Islands, Ceuta and Melilla, and REGCAN95 in the Canary Islands (both systems are compatible with WGS84). UTM projection in the corresponding time zone is also applied, with an extended time zone 30 for sheets of time zones 29 and 31. This DEM was obtained by interpolation from land cover data obtained with LIDAR of the National Plan of Aerial Orthophotography (PNOA), except for the sheets of Ceuta, Melilla, and Alboran Island (1110, 1111, 1078B). They were constructed by automatic stereo-correlation of photogrammetric flies (PNOA) with a resolution from 25 to 50 cm/pixel, revised and interpolated with break lines where it was viable.

1.2 Land Cover Databases

The characterization of both the aerodynamic roughness length (z_0) and the displacement height (d) is critical when modeling the wind field using the log vertical profile. It is known that the values of these parameters depend on weather conditions and land coverage. Thus, many authors have studied its relationship, providing typical values for each land cover. In this chapter, we have performed a comprehensive literature review to collect the intervals of z_0 and d values for each land coverage [60]. In particular, we have focused on the coverages present in the "Information System of Land Cover of Spain" (SIOSE).

Shape factors and atmospheric conditions can influence the aerodynamic parameters of surfaces and, hence, the vertical wind profiles (see, e.g., [54] showing the influence of z_0 in wind speed). Therefore, it is essential to know the roughness length and the displacement height to define the wind state in numerous applications, such as the wind field, air quality, and forest fire spread modeling. The values of z_0 and d are generally related to the vegetation and topographical characteristics so that they are affected by the land coverage variations; for example, the change of season (especially in vegetation cover), construction or demolition of buildings, etc. In addition, for each coverage, z_0 and d estimations may vary according to the wind speed and direction, and the atmospheric stability, see, e.g., [7]. Under this assumption, maps of z_0 and d are built for each weather conditions. As a rule of thumb, we can compute their values as a function of the height of the surface morphology characteristics (h). For instance, for a crop or forest canopy [10] proposes a value of d between 0.67 and 0.75 h, and a value of z_0 about 0.12 h. However, these approximations cannot be

applied if the surface is not homogeneous. In such cases, a more detailed analysis of the land coverage is required [47, 48].

For this reason, during the last five decades, many authors have proposed parameterizations for several land coverages. In these first approaches, the characteristics of the surface elements were used to estimate the roughness parameters for the various canopies. For example, with regard to crop canopies, [42] estimated z_0 according to h and d for harvested wheat; [93] obtained z_0 and d as a function of h for olives orchards; and [41] parameterized z_0 from wind measures using the wind profile. In forests, [100] estimated z_0 and d according to the tree crown and structure, while [63] obtained them from wind observations and canopy structure. On the other hand, [20] performed a parameterization from measures in a desert with artificial vegetation, whereas [99] did it from wind speed, temperature, and turbulent flow measures in bare soils. Moreover, in urban terrains, [30] compared several formulae of these parameters, and [56] carried out a parameterization of z_0 and d in a heterogeneous surface that was validated using a wind tunnel and empirical data. In wetlands, [61] obtained z_0 and d from the minimization of a least square difference function based on the log wind profile equation for near-neutral stability; in wet grasslands and reed beds, [1] estimated z_0 from eddy correlation measurements; and [78] used specific parameterizations for a Siberian bog. For water surfaces, some approaches were gathered, such as the estimation of z_0 for the sea that was validated in the laboratory by [39]; the comparison of two parameterizations for oceans by [22]; and the parameterization of z_0 for the sea by [25]. Finally, some authors have also parameterized z_0 and d for general land coverage: [92] proposed changes in Raupach parameterization using a list of values from the literature; [29] compared three previous methods and proposed to use the median; and, in the model of the European Wind Atlas [90], a parameterization of z_0 and d was introduced in four classes of coverages.

Some other authors estimated z_0 and d values according to canopy form. For example, [8] used catastral databases, whereas [14] applied a characterization of buildings in urban terrains. In particular, [84] estimated d in Tokyo considering buildings of different heights. Most of these methods require costly field works. However, the methodology proposed here avoids measuring problems related to the evaluation of roughness parameters. In other approaches, the wind profile is directly used to estimate z_0 and d from wind measurements at different heights over a homogeneous surface from within the inertial sublayer. In their work, [52] estimated z_0 and d for cotton, orchards, and desert covers from wind measures using the wind profile; [36] obtained z_0 and d from measures of different instruments; [5] calculated z_0 and d from wind measures and wind profile with radar in desert; [68, 73] used anemometer measures at different heights for cliff coverage; [11] estimated z_0 and d from field measures in peatlands; [83] obtained z_0 and d from friction velocity ($u^\star$) measures using the wind profile; [13] used different measure equipment and compared results with other authors in desert; [67] presented a regression of data obtained with radio-wind probes in forest; and [45] obtained d from simultaneous scintillation measurements at two heights. Both parameters can be estimated by solving the nonlinear wind profile equations: see, e.g., [26] for seas and land, and [18] for forest canopy.

Some applications of particular models and their database have been considered in this study. In [82], z_0 and d were estimated with several numerical models and measured data for cotton, scrub, and grass canopies; [57] used a computational fluid dynamics (CFD) simulation and the land cover database of the National Land Agency of Japan; [38] performed a simulation with COAMPS W-UCM for several episodes in New York; [44] used the model COSMO to parameterize z_0 in urban terrains; [91] applied the LGN3 database in Rotterdam; [55] compared results of z_0 and d from the National Land Cover Database (NLCD) for the Conterminous United States [24] with field measures in floodplain surfaces; and [97] presented a project for evaluating the annual wind energy production with a CFD code, a digital land model and three land cover databases. Relevant summaries extracted from several sources may be found in [34], where an extended review of many estimations of z_0 and d is provided, and in [28], where a useful list of parameter values is presented.

Special attention is paid to studies involving the use of remote sensing. On the one hand, the use of aircraft Lidar surveys was presented in [31] where z_0 was obtained for roads; in [37], describing a totally automated approach to the generation of z_0 values from Lidar terrain data; in [88], using a combination of low and high density airborne Lidar and satellite SPOT-5 HRG data, in conjunction with ground measurements of forest structure, to parameterize four models for d and z_0 over cool-temperate forests in an inland river basin; and also in [15] for an application in an inland river basin. On the other hand, in the last 25 years, several projects on land cover mapping have been developed mainly using satellite images. They characterize both parameters for each surface type. For example, the LGN7 model that uses several databases in The Netherlands with NSD (National Satellite Data) and aerial photos [35]; the LGN3 land cover database of The Netherlands that combines satellite and ancillary data [98]; the NLCD [24]; the CORINE land cover database [6]; and the SIOSE land cover database of Spain [64] that has been used in this work.

In general, most of the previous methods estimate fixed values of z_0 and d of particular land coverages. The final aim of constructing roughness parameter maps is to improve the mesoscale predictions using a downscaling model, allowing to obtain a more accurate estimation of the wind power that may be generated in a region.

1.3 SIOSE Land Cover Database

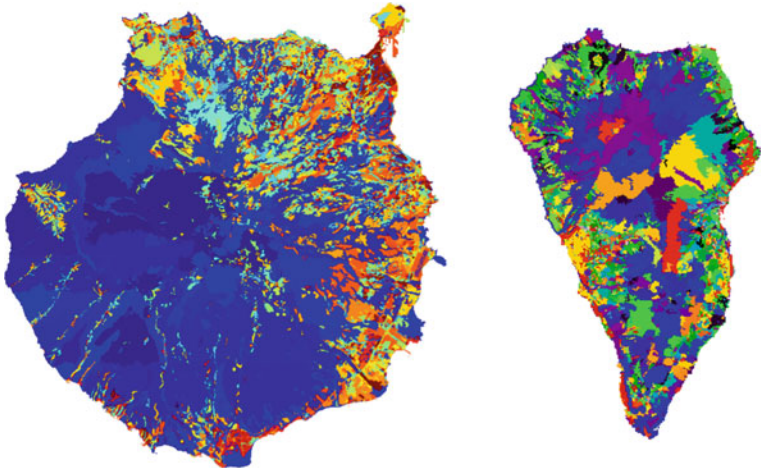
In 1990, the first land cover database encompassing the whole national territory was constructed in Spain on a scale of 1 : 100.000. It was developed in the framework of the CORINE Land Cover (CLC) European project. In the year 2000, there was a need to update the database to homogenize and improve its utility for developing territorial analysis and European policies. The resulting database is known as Image and CORINE Land Cover 2000. This update led to a new land cover database for the entire European continent. Other CLC updates have been produced in 2006 and 2012. In short, it consists of an inventory of land cover in 44 classes. The CLC uses

a Minimum Mapping Unit (MMU) of 25 hectares (ha) for areal phenomena and a minimum width of 100 m for linear phenomena. The time series are complemented by change layers, which highlight changes in land cover with an MMU of 5 ha. The CLC is produced by the majority of countries, by visual interpretation of high-resolution satellite imagery. In a few countries, semiautomatic solutions are applied, using national in situ data, satellite image processing, GIS integration, and generalization. In comparison with the NLCD of the United States of America, the CLC has the advantage that it contains vectorized polygon data instead of the raster data of the NLCD. Also, some layers of the CLC, like fruit trees or olive groves, are missing in the NLCD. However, the NLCD brings some better aspects like up-to-date data, a very high resolution of 30 m/pixel and an available companion dataset with canopy density.

The project SIOSE (Spanish acronym for Information System of Land Cover of Spain) was created in 2005 by the National Reference Center on Land Cover and on Land Use and Spatial Planning to integrate the local information available from the Autonomous Communities and the General State Administration. It uses the geodetic Cartesian reference frame ETRS89 with UTM Projection on time zones 28, 29, 30 and 31, and INSPIRE Directive. Since the requirements at the Spanish national level were higher than those supplied by the European project, the SIOSE generated a new land cover database for all the country on a $1:25.000$ scale. It was based on reference images from 2005, with a MUM of 0.5 to 2 ha (SIOSE 2005) and a planimetric accuracy of 5 m or better. The project was updated in 2009 and 2011, see [64]. Other important differences with the CLC are the land classification and the hierarchy levels, which are much more simplified in the CLC than in the SIOSE.

The SIOSE inventory is based on reference information, satellite SPOT5 imagery, as geometrical and time reference, and orthophotographs of the National Aerial Orthophotography Plan (PNOA). It uses the cadastre; the Integrated Water Information System (SIA); the Geographical Information System of Agricultural Parcel (SIGPAC); the database of boundary lines between Autonomous Communities from the Central Register of Cartography of the National Geographic Institute; orthophotographs and satellite imagery; databases and thematic maps related to the land cover; the Spanish Forest Map (MFE); and the Map of Crops and Utilization (MCA) provided by the Autonomous Communities, previously approved by the Project National Direction.

The SIOSE database consists of different basic and compound coverages. A compound coverage is made up of a combination of basic or compound coverages. Specifically, it considers eight general groups of basic coverages (Crops, Grassland, Forest, Scrubs, No Vegetation, Artificial Coverage, Wet Coverage, and Water Coverage) that are further refined into forty specific classes of basic land coverage, see, e.g., [65]. Therefore, at any point of the terrain, the land cover is defined as a weighted average of these forty basic coverages. The spatial unit is the polygon. Each polygon must contain a basic or compound coverage. The coverages that represent at least 5% of the polygon surface should be considered. In practice, each point of the DEM grid is labeled with the types of SIOSE polygon where it belongs. Figure 1 shows the polygons of the land cover classes in Gran Canaria (6, 983) and La Palma (2, 470).



(a) Gran Canaria (b) La Palma

Fig. 1 SIOSE land cover polygons for the Islands of Gran Canaria and La Palma, respectively

1.4 Roughness Length and Displacement Height: Literature Review

The search spaces of z_0 and d must be defined to obtain the appropriate values of them. Here, we present a methodology to generate a table with the ranges of z_0 and d values for each land coverage. Particularly, the methodology is applied to Gran Canaria and La Palma Islands, but it is suitable to any other location. To find the ranges of possible z_0 and d values for each land cover, we have carried out a literature review. Table 1 summarizes it and the specific references are listed in the caption. The first and second columns show the SIOSE code and a description for each of the distinct land coverages. The third and fourth, and the fifth and sixth columns present the nominal value and the range of the parameter z_0 and d, respectively. When data are not available, we have used the rule of thumb to obtain the z_0 and d values from the canopy height h, see [10]. Also, it is worth remarking that d is assumed to be zero for water surfaces (ACU, AEM, AES, ALC, ALG, AMO, and LAA classes).

Table 1 considers both general and local characteristics of the land coverage. On the one hand, for the most common coverage types in Gran Canaria, we have used the largest interval from those proposed by a wide list of authors. On the other hand, due to the particular characteristics of some land coverages in Gran Canaria, we have used a more specific study in some cases. For instance, the works of [24, 73] have been employed to characterize sea cliffs. The former study proposes a value of $d = 3.3$ m for a cliff of 40 m. Using this same ratio, the value for the Andén Verde (690 m) and the Risco de Faneque (1027 m) cliffs are 57 m and 85 m, respectively. These values have been taken as lower bound (3.3 m), nominal value (57 m), and upper bound (85 m). Another specific land coverage of Gran Canaria is the LOC

class (Other Woody Crops). In Gran Canaria, it refers to Aloe Vera plantations. In this case, [58] studied the typical canopy height ranges for the Aloe Vera. Using these ranges and the rule of thumb, we have obtained the z_0 and d ranges and nominal values.

We have searched in the literature the minimum and the maximum values of each parameter and canopy. In addition, the values more used by different authors have been selected as nominal values, but these are not used in our approach. In fact, we used the ranges of Table 1 as searching space for the solution of each parameter in each characteristic wind situations. Then, we performed an extended literature review on the assigned values of the roughness parameters of each coverage. For this reason, in general, we think that the proposed ranges completely cover the variation interval of z_0 and d corresponding to each coverage, respectively, not only for the region studied in this paper but also for any region if the specific coverage is included in that literature review. Nevertheless, some ranges have been defined according to the local characteristics of certain coverages. This is the case of the high sea cliffs in Gran Canaria, for example, where the range of z_0 and d have been adapted to such heights. Another case is Other Woody Crops related to Aloe Vera plantations on the island. For the latter, we have used the standard morphological characteristics. So, in general, these are the only particular cases to review for the application of ranges given in Table 1 to any other region.

The bibliography of Table 1 may be classified according to the procedure used to obtain z_0 and d. Several authors proposed parameterizations of the roughness parameters from measures of the wind and other physical magnitudes. This is the case of [10] which is used here in many coverages; [90] in water surfaces; [51] in screes; [79] in screes, conifers, and citrus fruit trees; [86, 87] in conifers; [46] in rice crops; [30] in low buildings; [93] in olive groves; [40] in vineyards; [74] in water surfaces; and [32] in artificial coverages.

Some other works are based on the canopy morphology, such as [12] used in water and saline coverages; [53] in deciduous forests, wetlands, and scrubs; and [58] in Aloe Vera crops. Another extended approach is the use of measurements and wind profile: [73] used in cliffs; [13] in screes; [5] in Quaternary lava flow; [18, 36, 67] in conifers; [83] in rice crops; [41] in crops different from rice; [70] in no citrus fruit trees; [72, 95] in vineyards; [26] in scrubs; and [96] in soils without vegetation and artificial coverages.

Also in Table 1, there are some values of z_0 and d arising from applications of specific numerical models with their land cover databases. For example, [75] used them in lakes and lagoons; [101] in rocky outcrops and rocks, evergreen forests, salt marshes, wetlands, bare soils, and artificial coverages; [21] in evergreen forests; [57] used in no citrus fruit trees; [82] in grasslands; and [81] in ravine and road, parking, or unvegetated pedestrian areas.

Some papers dealing with collections of data from other authors, in particular from old publications, were useful too. In particular, the one by [4] used in estuaries, seas and oceans, and soils without vegetation; [23] in crops different from rice and artificial coverage; [89] in buildings, evergreen forests, glaciers and perpetual snow, scrubs, meadows, grasslands and artificial green area and urban trees; [28] in deciduous and

evergreen forests; and [59] in crops different from rice, woody crops, meadows and grasslands. Also, some early publications were used for constructing Table 1, such as [16], used in crops different from rice; [17] in glaciers and perpetual snow, meadows, and grasslands; and [9, 43] in citrus fruit trees.

Finally, some of the current land cover databases based on remote sensing surveys (specifically, aircraft lidar and satellite images) that were taken into account are [24]: the NLCD database used in cliffs; [3]: the 1-km land cover data set DISCOVER (IGBP-DIS) for watercourses; [6]: the CLC database applied in many coverages; and [98]: the LGN3 database used in rocky outcrops and rocks, salt marshes, soils without vegetation, and artificial coverages.

1.5 Roughness Length and Displacement Height Characterization

The SIOSE project uses a vectorial format, but, for convenience, we will translate it to a raster format. For this, we will define a grid with n_p points and, for each point, we will look for the mean value of basic coverages. Once we have the values of z_0 and d for each basic coverage, we can compute the specific z_0 and d values at any point using an appropriate weighted mean. This way, the SIOSE database will let us create a matrix with the percentage of the basic coverages at any point. This matrix is defined as follows: let M be an $n_p \times n_b$ matrix, with components $m_{i,j}$, where n_b is the number of basic coverages. For each row i of M, $m_{i,j}$ is the fraction of the basic coverage j at the point n_i ($m_{i,j} < 1$ and $\sum_{j=1}^{n_b} m_{i,j} = 1$).

When the ranges of z_0 and d are set for each basic land coverage, we can compute its values at any point of the terrain. Assuming that the values of z_0 and d are a certain mean of the values of the basic coverages z_{0j} and d_j, $j = 1, \ldots, n_b$, we can compute their values at any point. A simple weighted average may produce differences with the effective roughness of one order of magnitude. The study by [76] for the effective roughness length improves the formula proposed by [85], taking into account a nondimensional patchiness parameter to consider the textural information about the spatial dependence of the primitives (regions with specific properties) characterizing the surface inhomogeneity. However, generally the coverage information provided by the land cover databases (e.g., SIOSE) does not include the spatial distribution of basic coverages in a composed one, but only the fraction of the area covered by each surface type. So, in this case, only Taylor's formula was applied; in this approach, for a coverage i composed by n_b basic canopies with roughness length z_{0j}; $j = 1, \ldots, n_b$ on a fraction m_{ij} of the area, respectively, an approximation to the effective roughness length is given by computing the weighted geometric mean roughness length z_0:

$$z_0 = \prod_{j=1}^{n_b} z_{0j}^{m_{ij}}. \tag{1}$$

Regarding the displacement height, d, [76, 85] did not study its estimation. In any case, the above mean formula is not appropriate for d, since, in the case of a basic coverage with displacement height equal to zero, it would produce a mean $z_0 = 0$ independently of the z_{0j} values of the other basic coverages. Some other works have studied the variation of d in several specific coverages, e.g., in urban [56] and vegetation [62, 100] canopies. One important conclusion is that the effective displacement height of a heterogeneous coverage can exceed the surface mean canopy height significantly. Taking this into account, we propose to use a weighted root mean square to obtain the highest mean value:

$$d = \sqrt{\sum_{j=1}^{n_b} m_{ij} d_j^2}, \tag{2}$$

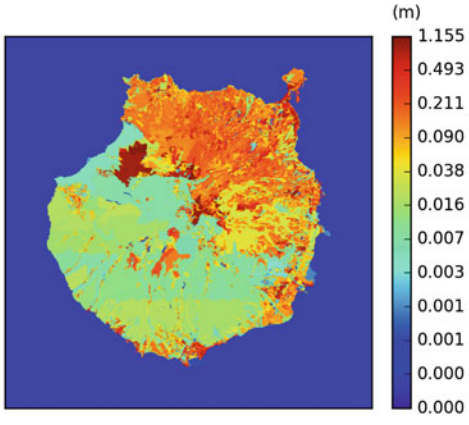
where d_j and m_{ij} are the displacement height and the fraction of the basic coverage j in the composed one i, respectively. Figures 2a and b show the resulting composed z_0 and d values in Gran Canaria and La Palma Islands, considering the nominal values of the basic coverages given in Table 1. We remark that a weighted average version of that proposed by [69]

$$d = exp\left[\sum_{j=1}^{n_b} \frac{m_{ij}}{\ln d_j}\right]^{-1}, \tag{3}$$

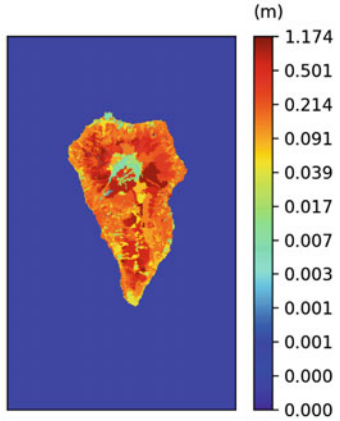
which must be evaluated in the limit if any $d_j = 0$ or $d_j = 1$, is not appropriate in the case of any $d_j = 1$ since the average result of d would always be equal to 1, independently of the other d_j values. Similar conclusions may be reached from its application in the calculation of z_0 if any $z_{0j} = 1$. It can be noted that the application of this methodology to another database is straightforward.

1.6 Literature Review and Characterization of Albedo

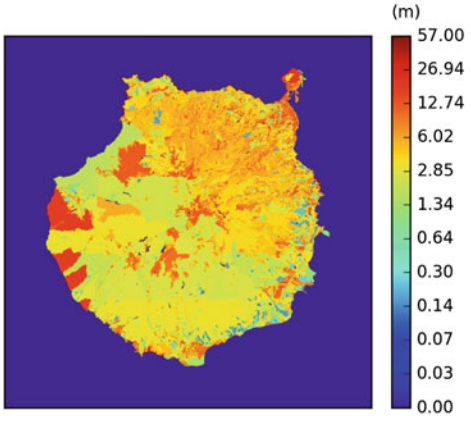
The ranges of albedo ρ for each land cover have been obtained from a literature review. Table 2 summarizes it, and the specific references are listed in the caption. Here again, the first and second columns show the SIOSE code and a description for each land coverages. The third and fourth columns present the nominal value and the range of ρ. Following the same procedure as for z_0 and d, we have taken the largest interval from those proposed by a wide list of authors.



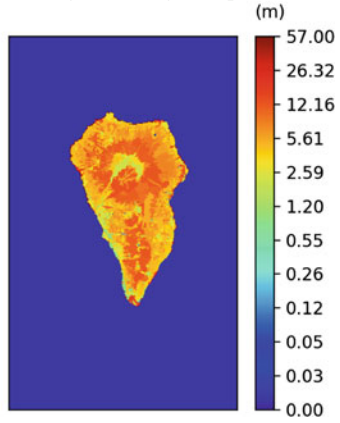
(a) Roughness length map of Gran Canaria



(b) Roughness length map of La Palma



(c) Displacement height map of Gran Canaria



(d) Displacement height map of La Palma

Fig. 2 Roughness length and displacement height maps of Gran Canaria and La Palma islands (m) corresponding to the nominal values stated in Table 1 and using the mean values given in (1) and (2), respectively

Using the same nomenclature as in the roughness parameter characterization, we propose to compute the effective albedo at any point of a SIOSE polygon with an average of the mean values of the basic coverage albedo,

$$\rho = \sum_{j=1}^{n_b} m_{ij}\rho_j, \tag{4}$$

where ρ_j and m_{ij} are the albedo and the fraction of the basic coverage j in the composed one i, respectively.

Table 1 Nominal values and ranges of z_0 and d for the land cover classes provided by SIOSE. The superindex indicates the source: (a) [24], (b) [73], (c) [3], (d) [90], (e) [6], (f) [10], (g) [12], (h) [4], (i) [75], (j) [98], (k) [101], (l) [51], (m) [13], (n) [79], (o) [5], (p) [18], (q) [36], (r) [67], (s) [86, 87], (t) [83], (u) [46], (v) [16], (w) [41], (x) [23], (y) [74], (z) [89], (α) [30], (β) [53], (γ) [28], (δ) [21], (ϵ) [17], (ζ) [57], (η) [9], (θ) [43], (ι) [59], (κ) [70], (λ) [58], (μ) [93], (ν) [72], (ξ) [40], (σ) [95], (τ) [26], (ϕ) [32], (χ) [82], (ψ) [81], (ω) [96]

Code	Land cover	z_0 (m)	z_{0min}	z_{0max}	d (m)	d_{min}–d_{max}
ACM	Sea Cliffs	$0.05^{(a)}$	$0.05^{(b)}$	$0.19^{(b)}$	$57^{(b)}$	$3.3^{(b)}$–$85^{(b)}$
ACU	Watercourses	$0.00025^{(c)}$	$0.0001^{(d)}$	$0.01^{(e)}$	$0^{(f,y)}$	–
AEM	Water Body. Reservoirs	$0.00025^{(d)}$	$0.0001^{(d)}$	$0.005^{(g)}$	$0^{(f,y)}$	–
AES	Estuaries	$0.0002^{(h)}$	$0.0001^{(d)}$	$0.01^{(e)}$	$0^{(f,y)}$	–
ALC	Coastal Lagoons	$0.005^{(g)}$	$0.0001^{(d)}$	$0.01^{(e)}$	$0^{(f,y)}$	–
ALG	Water Body. Lakes and Lagoons	$0.0005^{(i)}$	$0.0001^{(d)}$	$0.005^{(g)}$	$0^{(f,y)}$	–
AMO	Seas and Oceans	$0.0002^{(h)}$	$0.0001^{(d)}$	$0.03^{(a)}$	$0^{(f,y)}$	–
ARR	Rocky Outcrops and Rocks	$0.005^{(e)}$	$0.0003^{(j)}$	$0.18^{(k)}$	$0.03^{(f)}$	$0^{(f)}$–$0.96^{(f)}$
CCH	Screes	$0.1^{(a)}$	$0.05^{(l)}$	$0.15^{(m)}$	$0.6^{(n)}$	$0.56^{(n)}$–$0.66^{(n)}$
CLC	Quaternary Lava Flow	$0.0286^{(o)}$	$0.0013^{(o)}$	$0.0735^{(o)}$	$0.15^{(f)}$	$0^{(f)}$–$0.4^{(f)}$
CNF	Forest. Conifers	$1.28^{(p)}$	$0.25^{(q)}$	$1.93^{(r)}$	$13.1^{(s)}$	$4.87^{(r)}$–$22^{(n)}$
CHA	Herbaceous Crops. Rice	$0.072^{(t)}$	$0.001^{(u)}$	$0.11^{(t)}$	$0.85^{(t)}$	$0.1^{(t)}$–$1.55^{(t)}$
CHL	Herbaceous Crops. Different from Rice	$0.1^{(v)}$	$0.004^{(w)}$	$0.74^{(x)}$	$0.25^{(w)}$	$0.1^{(w)}$–$3^{(\iota)}$
EDF	Artificial Coverage. Buildings	$1.5^{(z)}$	$0.7^{(z)}$	$3.7^{(x)}$	$14^{(z)}$	$7^{(z)}$–$19.73^{(f)}$
FDC	Forest. Leafy. Deciduous	$1^{(\beta)}$	$0.18^{(\beta)}$	$1.4^{(a)}$	$11.8^{(\gamma)}$	$3^{(\gamma)}$–$21.6^{(\gamma)}$
FDP	Forest. Leafy. Evergreen	$0.72^{(k)}$	$0.6^{(c)}$	$2.65^{(\delta)}$	$9.7^{(\gamma)}$	$3^{(\gamma)}$–$31^{(z)}$
GNP	No Vegetation. Glaciers and Perpetual Snow	$0.001^{(e)}$	$0.00001^{(\epsilon)}$	$0.012^{(z)}$	$0.01^{(f)}$	$0^{(f)}$–$0.06^{(f)}$
HMA	Salt Marshes	$0.11^{(k)}$	$0.0002^{(j)}$	$0.17^{(j)}$	$0.6^{(f)}$	$0^{(f)}$–$0.93^{(f)}$
HPA	Wetlands	$0.1^{(e)}$	$0.005^{(\beta)}$	$0.55^{(k)}$	$0.55^{(f)}$	$0.03^{(f)}$–$3^{(f)}$
HSA	Salt Mines	$0.01^{(e)}$	$0.0005^{(e)}$	$0.04^{(g)}$	$0.05^{(f)}$	$0^{(f)}$–$0.22^{(f)}$
HSM	Salt Lakes	$0.01^{(e)}$	$0.0005^{(e)}$	$0.04^{(g)}$	$0.05^{(f)}$	$0^{(f)}$–$0.22^{(f)}$
HTU	Peat Bogs	$0.03^{(e)}$	$0.0005^{(e)}$	$0.03^{(e)}$	$0.16^{(f)}$	$0^{(f)}$–$0.16^{(f)}$

(continued)

Table 1 (continued)

Code	Land cover	z_0 (m)	z_{0min}	z_{0max}	d (m)	d_{min}–d_{max}
LAA	Artificial Coverage. Artificial Water Body	$0.0001^{(e)}$	$0.0001^{(e)}$	$0.005^{(g)}$	$0^{(e,y)}$	–
LFC	Woody Crops. Citrus Fruit Trees	$0.31^{(\eta)}$	$0.03^{(d)}$	$0.4^{(\theta)}$	$3^{(n)}$	$0^{(\iota)}$–$4^{(\iota)}$
LFN	Woody Crops. No Citrus Fruit Trees	$0.25^{(e)}$	$0.03^{(d)}$	$1^{(\zeta)}$	$0.92^{(\kappa)}$	$0^{(\iota)}$–$4^{(\iota)}$
LOC	Other Woody Crops	$0.0615^{(\lambda,f)}$	$0.0369^{(\lambda,f)}$	$0.0861^{(\lambda,f)}$	$0.33^{(\lambda,f)}$	$0.2^{(\lambda,f)}$–$0.47^{(\lambda,f)}$
LOL	Olive Groves	$0.48^{(\mu)}$	$0.25^{(e)}$	$0.61^{(\mu)}$	$2.67^{(\mu)}$	$2^{(\mu)}$–$3^{(\mu)}$
LVI	Vineyards	$0.2^{(\upsilon)}$	$0.08^{(\xi)}$	$0.55^{(\upsilon)}$	$0.75^{(\upsilon)}$	$0.31^{(\xi)}$–$1.4^{(\sigma)}$
MTR	Scrubs	$0.16^{(\beta)}$	$0.016^{(\beta)}$	$1^{(a)}$	$4.8^{(\tau)}$	$0.9^{(z)}$–$7.1^{(\tau)}$
OCT	Artificial Coverage. Other Buildings	$0.5^{(e)}$	$0.06^{(k)}$	$1^{(e)}$	$4^{(\alpha)}$	$2^{(\alpha)}$–$14^{(z)}$
PDA	No Vegetation. Beaches, Dunes and Sandy Areas	$0.0003^{(e)}$	$0.0003^{(e)}$	$0.06^{(j)}$	$0^{(f)}$	$0^{(f)}$–$0.33^{(f)}$
PRD	Crops. Meadows	$0.03^{(e)}$	$0.001^{(\epsilon)}$	$0.1^{(e)}$	$0.013^{(z)}$	$0.007^{(\iota)}$–$0.035^{(z)}$
PST	Grasslands	$0.09^{(\epsilon)}$	$0.001^{(\epsilon)}$	$0.15^{(\epsilon)}$	$0.171^{(\chi)}$	$0.013^{(z)}$–$0.66^{(\iota)}$
RMB	No Vegetation. Ravines	$0.0012^{(\psi)}$	$0.0003^{(d)}$	$0.005^{(\omega)}$	$0.03^{(h,f)}$	$0^{(f)}$–$0.03^{(f)}$
SDN	No Vegetation. Bare Soil	$0.001^{(j)}$	$0.0002^{(\omega)}$	$0.04^{(k)}$	$0.03^{(h,f)}$	$0^{(f)}$–$0.22^{(f)}$
SNE	Artificial Coverage. Unbuilt Land	$0.0003^{(j)}$	$0.0002^{(\omega)}$	$0.04^{(a)}$	$0^{(f)}$	$0^{(f)}$–$0.22^{(f)}$
VAP	Artificial Coverage. Road, Parking or Unvegetated Pedestrian Areas	$0.03^{(e)}$	$0.0035^{(\psi)}$	$0.5^{(e)}$	$1^{(y,\phi)}$	$0.02^{(\phi)}$–$2.5^{(\phi)}$
ZAU	Artificial Coverage. Artificial Green Area and Urban Trees	$0.4^{(d)}$	$0.03^{(j)}$	$1.3^{(x)}$	$3.5^{(y,\phi)}$	$3.5^{(z)}$–$14^{(z)}$

(continued)

Table 1 (continued)

Code	Land cover	z_0 (m)	z_{0min}	z_{0max}	d (m)	d_{min}–d_{max}
ZEV	Artificial Coverage. Extraction or Waste Areas	$0.1^{(e)}$	$0.0003^{(j)}$	$0.18^{(k)}$	$0.16^{(\omega,f)}$	$0^{(f)}$–$1^{(f)}$
ZQM	No Vegetation. Burnt Areas	$0.6^{(e)}$	$0.1^{(e)}$	$1.1^{(j)}$	$3.27^{(f)}$	$0.54^{(f)}$–$6^{(f)}$

Table 2 Albedo nominal values and ranges for SIOSE land cover classes. The superindex indicates the source: (*a*) [2], (*b*) [50], (*c*) [80], (*d*) [66], (*e*) [33], (*f*) [94], (*g*) [19], (*h*) [49], (*i*) [27]

Code	Land cover	ρ	ρ_{min}	ρ_{max}
ACM	Sea Cliffs	$0.2^{(a)}$	$0.05^{(b)}$	$0.7^{(b)}$
ACU	Watercourses	$0.14^{(c)}$	$0.03^{(d)}$	$1^{(d)}$
AEM	Water Body. Reservoirs	$0.14^{(c)}$	$0.03^{(d)}$	$1^{(d)}$
AES	Estuaries	$0.1^{(a)}$	$0.03^{(d)}$	$1^{(d)}$
ALC	Coastal Lagoons	$0.1^{(a)}$	$0.03^{(d)}$	$1^{(d)}$
ALG	Water Body. Lakes and Lagoons	$0.14^{(c)}$	$0.03^{(d)}$	$1^{(d)}$
AMO	Seas and Oceans	$0.1^{(a)}$	$0.03^{(d)}$	$1^{(e)}$
ARR	Rocky Outcrops and Rocks	$0.2^{(a)}$	$0.1^{(b)}$	$0.4^{(b)}$
CCH	Screes	$0.2^{(a)}$	$0.05^{(b)}$	$0.7^{(b)}$
CLC	Quaternary Lava Flow	$0.1^{(e)}$	$0.05^{(f)}$	$0.15^{(f)}$
CNF	Forest. Conifers	$0.14^{(e)}$	$0.05^{(d)}$	$0.15^{(d)}$
CHA	Herbaceous Crops. Rice	$0.12^{(e)}$	$0.11^{(e)}$	$0.25^{(d)}$
CHL	Herbaceous Crops. Different from Rice	$0.2^{(a)}$	$0.18^{(d)}$	$0.25^{(d)}$
EDF	Artificial Coverage. Buildings	$0.18^{(a)}$	$0.18^{(a)}$	$0.35^{(a)}$
FDC	Forest. Leafy. Deciduous	$0.16^{(a)}$	$0.10^{(e)}$	$0.5^{(a)}$
FDP	Forest. Leafy. Evergreen	$0.12^{(a)}$	$0.09^{(c)}$	$0.35^{(a)}$
GNP	No Vegetation. Glaciers and Perpetual Snow	$0.6^{(a)}$	$0.2^{(d)}$	$0.95^{(d)}$
HMA	Salt Marshes	$0.14^{(c)}$	$0.14^{(a)}$	$0.3^{(a)}$
HPA	Wetlands	$0.14^{(c)}$	$0.14^{(a)}$	$0.3^{(a)}$
HSA	Salt Mines	$0.5^{(g)}$	$0.166^{(e)}$	$0.5^{(g)}$
HSM	Salt Lakes	$0.5^{(g)}$	$0.166^{(e)}$	$0.5^{(g)}$
HTU	Peat Bogs	$0.14^{(c)}$	$0.14^{(a)}$	$0.3^{(a)}$
LAA	Artificial Coverage. Artificial Water Body	$0.14^{(c)}$	$0.03^{(d)}$	$1^{(a)}$
LFC	Woody Crops. Citrus Fruit Trees	$0.18^{(a)}$	$0.13^{(h)}$	$0.22^{(h)}$
LFN	Woody Crops. No Citrus Fruit Trees	$0.18^{(a)}$	$0.13^{(h)}$	$0.22^{(h)}$
LOC	Other Woody Crops	$0.18^{(a)}$	$0.13^{(h)}$	$0.22^{(h)}$
LOL	Olive Groves	$0.18^{(a)}$	$0.13^{(h)}$	$0.22^{(h)}$
LVI	Vineyards	$0.18^{(a)}$	$0.14^{(a)}$	$0.5^{(a)}$
MTR	Scrubs	$0.25^{(a)}$	$0.14^{(c)}$	$0.5^{(a)}$

(continued)

Table 2 (continued)

Code	Land cover	ρ	ρ_{min}	ρ_{max}
OCT	Artificial Coverage. Other Buildings	$0.16^{(a)}$	$0.16^{(a)}$	$0.45^{(a)}$
PDA	No Vegetation. Beaches, Dunes and Sandy Areas	$0.35^{(f)}$	$0.15^{(f)}$	$0.45^{(g)}$
PRD	Crops. Meadows	$0.2^{(a)}$	$0.1^{(e)}$	$0.6^{(a)}$
PST	Grasslands	$0.18^{(a)}$	$0.08^{(e)}$	$0.6^{(a)}$
RMB	No Vegetation. Ravines	$0.16^{(c)}$	$0.147^{(i)}$	$0.173^{(i)}$
SDN	No Vegetation. Bare Soil	$0.16^{(c)}$	$0.147^{(i)}$	$0.173^{(i)}$
SNE	Artificial Coverage. Unbuilt Land	$0.18^{(a)}$	$0.15^{(a)}$	$0.6^{(a)}$
VAP	Artificial Coverage. Road, Parking or Unvegetated Pedestrian Areas	$0.18^{(a)}$	$0.18^{(a)}$	$0.35^{(a)}$
ZAU	Artificial Coverage. Artificial Green Area and Urban Trees	$0.15^{(c)}$	$0.15^{(a)}$	$0.6^{(a)}$
ZEV	Artificial Coverage. Extraction or Waste Areas	$0.2^{(a)}$	$0.13^{(a)}$	$0.6^{(a)}$
ZQM	No Vegetation. Burnt Areas	$0.097^{(i)}$	$0.089^{(i)}$	$0.098^{(i)}$

2 Meteorological Parameters

The wind is produced, firstly, as a consequence of spatial differences of barometric pressure, generally caused by the absorption of the solar radiation. In a horizontal plane, the wind flows from high-pressure zones to low-pressure ones, while vertically from low-pressure zones to high-pressure ones. Wind speed is proportional to the pressure variation per unity of length or pressure gradient. Zones of same pressures are represented in the weather maps joined by imaginary lines (isobars). The closer the isobars are to each other, the stronger the wind is.

A second factor that affects the air movement is the Coriolis force, caused by the Earth rotation. The term $f = 2\Theta \, sen \, \phi_l$ is called Coriolis parameter, where $\Theta = 7.292 \times 10^{-5} \, s^{-1}$ is the Earth rotation velocity and ϕ_l is the latitude. It is considered positive in the North Hemisphere, null in the Equator, and negative in the South Hemisphere.

Thirdly, a centripetal acceleration may appear when the wind turns around a center. Finally, also the friction because of the air movement must be considered. Winds affected by the pressure gradient and the Coriolis force are called geostrophic winds.

The Monin–Obukhov similarity theory leads to a division of the lowest layer of the atmosphere into several sublayers where the vertical wind profile is constructed in different ways. So, the planetary boundary layer is located at a height z_{pbl} over the terrain, and it is the layer under the free atmosphere that is directly affected by the friction of the Earth surface. In this case, z_{pbl} is defined such that wind velocity is considered to be constant over it [77],

Table 3 Pasquill stability classes depending on surface wind speeds and isolation. Strong isolation corresponds to the typical sunny noon of the middle summer in England; light isolation, to similar conditions in middle winter. Night is referred to the period between one hour before sunset and one hour after sunrise. Neutral class D should also be used for overcast skies during the day or the night, and for any sky condition during the preceding and following night hours defined above

Pasquill stability class					
	Isolation			Night	
Surface wind speed (m/s)	Strong	Moderate	Light	Overcast ≥4/8 clouds	≤3/8 clouds
< 2	A	A-B	B	–	–
$2-3$	A-B	B	C	E	F
$3-5$	B	B-C	C	D	E
$5-6$	C	C-D	D	D	D
> 6	C	D	D	D	D
For A-B, take the average of A and B values, etc.					

$$z_{pbl} = \frac{\gamma\,|\mathbf{v}^*|}{f}, \tag{5}$$

where γ is a constant between 0.15 and 0.45 that depends on the atmospheric stability (Table 3), and $\mathbf{v}^*$ is the friction velocity that is computed from wind measures or predictions.

The mixing layer, also called convective layer, is the atmospheric layer affected by convective phenomena caused by the surface heat. The air is well mixed, that is, the wind and the potential temperature are almost constant with height. The height of the mixing layer h_m is approximated by

$$h_m = \gamma' \sqrt{\frac{|\mathbf{v}^*|\,L}{f}}, \tag{6}$$

where usually $\gamma' = 0.4$ [102] and L is the Monin–Obukhov length, that is computed with the Liu formulae [71],

$$\frac{1}{L} = a z_0^b, \tag{7}$$

where a and b are defined by the Pasquill stability class (see Table 4).

The surface layer, located at a height z_{sl} over de terrain, is the lowest layer of the planetary boundary layer, just joining the terrain surface layer, where the drag friction force is dominant. The height of the surface layer is usually obtained [102],

$$z_{sl} = \frac{h_m}{10}. \tag{8}$$

Table 4 Values of the parameters a and b to calculate Monin–Obukhov length depending on the Pasquill stability class

Pasquill stability class	a	b
A (Extremely unstable)	−0.08750	−0.1029
B (Moderately unstable)	−0.03849	−0.1714
C (Lightly unstable)	−0.00807	−0.3049
D (Neutral)	0.00000	0.0000
E (Lightly stable)	0.00807	−0.3049
F (Moderately stable)	0.03849	−0.1714

The atmospheric stability is related to the atmospheric turbulence as well as with the temperature gradient and the thermal inversion. It provides a qualitative measure of the air density variations because of pressure and temperature changes, and other phenomena that affect certain atmospheric movements.

The stability of the atmosphere may be classified as follows:

- **Stable atmosphere**. If a mass of air goes up, it will be surrounded by hotter air and thus, less dense than it. This will make go down. If It goes down, it will be surrounded by colder air (denser) and will tend to go up. This air trend of staying in the same layer is called stability of the atmospheric stratification.
- **Unstable atmosphere**. Under unstable conditions, the potential temperature decreases with height, increasing the vertical movements, that is, if the air goes up, it will be surrounded by colder and denser air and it will tend to continue going up; and if it goes down, it will find hotter and lighter air, and it will tend to continue going down.
- **Neutral atmosphere**. If a mass of air (after a vertical movement in an atmospheric layer without mixing with the surrounding air) experiments a null vertical net force, the ascending movements will not be affected by the thermal gradient, and the atmosphere layer is assumed to be neutrally stratified. Under such conditions, this mass of air does not tend to regret to its original position (stable atmosphere) nor accelerates going away from it (unstable stratification).

The atmospheric stability was usually characterized by the Table 3 including the Pasquill stability classes.

Usually, the anemometers provide measures of turbulence intensity that may help to complete the information about the type of atmospheric stability of the region. The turbulence intensity i is defined as the square root of the sum of the variances, σ_u^2, σ_v^2, σ_w^2, of the three components of the wind velocity u_0, v_0, w_0, respectively, divided by the average of the measured wind speeds,

$$i = \frac{\sqrt{\sigma_u^2 + \sigma_v^2 + \sigma_w^2}}{|\mathbf{v}_0|}. \tag{9}$$

Table 5 Pasquill stability classes depending on the surface wind speed and the turbulence intensity

Pasquill stability class							
Surface wind speed (m/s)	Isolation				Night		
	$i > 0.35$	$0.35 \geq i > 0.25$	$0.25 \geq i > 0.15$	$0.15 \geq i$	$i > 0.075$	$0.075 \geq i > 0.03$	$0.03 \geq i$
$\|v_0\| < 2$	A	B	B	B	F	F	F
$2 \leq \|v_0\| < 3$	A	B	C	C	E	E	F
$3 \leq \|v_0\| < 5$	B	B	C	C	D	E	E
$\|v_0\| \geq 5$	C	C	C	D	D	D	D

In practice, we only have measures of intensity variations of the speed but not of the direction. In such cases, Eq. (9) yields:

$$i = \frac{\sigma_{v_0}}{|\mathbf{v}_0|}, \tag{10}$$

where σ_{v_0} represents the standard deviation of the measured wind intensities.

An unstable atmosphere means a high level of turbulence, with a range of turbulence intensities between 0.2 and 0.4, approximately. However, a stable atmosphere, with a small turbulence or an almost null one, is characterized by intensities between 0.05 and 0.01. Table 5 illustrates the relation of the turbulence intensity and the atmospheric stability.

The concepts of Neutral and Stable Boundary Layers (SBL) have been revised to consider the effect of the free-flow static stability and baroclinicity on the turbulent transport of momentum and scalars in the boundary layer, as well as the model of turbulent entrainment for Convective Boundary Layers (CBL) [104]. Accordingly, different types of SBL regimes can be distinguished: truly neutral (absence of any buoyancy effects throughout the PBL); conditionally neutral (buoyancy flux at the surface is negligible); short-lived nocturnal (separated from the free atmosphere by near-neutral residual layers); and long-lived (immediately adjoining the stably stratified free atmosphere). The SBL height, h, may be evaluated according to the expression recently introduced by [103]. This expression represents a multi-limit equation for the equilibrium PBL height that covers the types mentioned earlier of neutral and stable conditions in the atmosphere. In contrast to the stable and neutral cases, the estimation of the CBL height is not straightforward, since our model is diagnostic and the recommended parameterizations are prognostic. To overcome this problem, some prognostic data must be used. In this approach, the CBL height is obtained from the results of the mesoscale model, e.g., HARMONIE-AROME model.

The estimation of the PBL height is calculated separately for stable/neutral conditions and convective conditions. For this, the Brunt–Väisälä frequency N in the PBL and the surface buoyancy flux B_s allow to characterize atmospheric stability:

$$N^2 = \frac{g}{T}\left(\frac{\partial T}{\partial z} + \Gamma_d\right), \tag{11}$$

g being the gravity acceleration, T a reference value of the air absolute temperature, z the height variable, and $\Gamma_d = 9.8 \times 10^{-3}$ K/m the dry adiabatic lapse rate. If $N^2 \geq 0$, the atmosphere is considered stable/neutral. However, $N^2 < 0$ indicates a CBL.

In stable/neutral atmosphere (SBL), the formula to compute the SBL height was proposed by [103]

$$h = \gamma\, u_*/f, \tag{12}$$

where u_* is the surface friction velocity, f the Coriolis parameter defined as $f = 2\omega \sin\phi$ (ω is the Earth rotation and ϕ the latitude), and γ is a function of the imposed-stability parameter $\mu_N = N_{2h-h}/f$ in the free atmosphere:

$$\gamma = \gamma_0 \left(1 + \frac{\gamma_0^2 C_{uN}}{C_s^2}\mu_N\right)^{-1/2}. \tag{13}$$

Experimental data suggest $\gamma_0 = 0.5$, $C_{uN}/C_s^2 = 0.6$; N_{2h-h} is the free-flow Brunt-Väisälä frequency in the free atmosphere immediately above the SBL ($h < z < 2h$). In particular, the stable/neutral PBL may be classified as: truly neutral (TN) at $\mu_N = 0$ (B_s and $N_{2h-h} = 0$); conditionally neutral (CN) usually at $0.5 \times 10^2 < \mu_N < 3 \times 10^2$ ($B_s \geq 0$ and $N_{2h-h} > 0$); nocturnal stable (NS) at ($B_s < 0$ and $N_{2h-h} = 0$); and long-lived stable (LS) at ($B_s < 0$ and $N_{2h-h} > 0$).

3 Conclusions

Some essential data and parameter definitions for wind and solar radiation modeling have been introduced in this chapter. From the geographical point of view, one must have information about the orography of the surface involved in the simulation and the land cover distribution of such surface. On the one hand, nowadays, DEM allows to construct meshes adapted to a terrain surface to be used in the discretization of a problem. On the other hand, the study of the roughness parameters associated to an accurate knowledge of the land coverages provides a valuable tool for a better wind field description. In the same way, the setting up of each land coverage albedo is important to obtain reliable values of reflection in solar radiation modeling.

References

1. Acreman MC, Harding RJ, Lloyd CR, McNeil DD (2003) Evaporation characteristics of wetlands: experience from a wet grassland and a reedbed using eddy correlation measurements. Hydrol Earth Syst Sci 7(1):11–21. https://doi.org/10.5194/hess-7-11-2003
2. Adrian G, Fiedler F (1991) Simulation of unstationary wind and temperature fields over complex terrain and comparison with observations. Beitr Phys Atmos 64:27–48
3. Belward AS, Estes JE, Kline KD (1999) The IGBP-DIS global 1-km land-cover data set DISCover: A project overview. Photogramm Eng Rem S 65(9):1013–1020
4. Blocken B, van der Hout A, Dekker J, Weiler O (2015) CFD simulation of wind flow over natural complex terrain: case study with validation by field measurements for Ria de Ferrol, Galicia, Spain. J Wind Eng Ind Aerod 147:43–57
5. Blumberg DG, Greeley R (1993) Field studies of aerodynamic roughness length. J Arid Environ 25:39–48
6. Bossard M, Feranec J, Otahel J (2000) CORINE land cover technical guide - Addendum 2000. Technical Report European Environment Agency, Copenhagen
7. Bosveld FC (1997) Derivation of fluxes from profiles over a moderately homogeneous forest. Bound-Lay Meteorol 84(2):289–327. https://doi.org/10.1023/a:1000453629876
8. Bottema M, Mestayer PG (1998) Urban roughness mapping—validation techniques and some first results. J Wind Eng Ind Aerod 74–76:163–173. https://doi.org/10.1016/s0167-6105(98)00014-2
9. Brooks FA (1959) An introduction to physical micrometeorology. University of California, Davis, California
10. Brutsaert W (1982) Evaporation into the atmosphere. D. Reidel Publish Company, Dordrecht, Holland
11. Campbell DR, Lavoie C, Rochefort L (2002) Wind erosion and surface stability in abandoned milled peatlands. Can J Soil Sci 82(1):85–95. https://doi.org/10.4141/s00-089
12. Chang Y, Tan J, Grimmond S, Tang Y (2015) Distribution of aerodynamic roughness based on land cover and DEM-A case study in Shanghai, China. In: ICUC9—9th international conference on urban climate jointly with 12th symposium on the urban environment
13. Claassen HC, Riggs AC (1993) An estimate of the roughness length and displacement height of Sonoran Desert vegetation, South-Central Arizona. U.S. Geological Survey. Technical Report Water-Resources Investigations Report 92-4065, Denver, Colorado
14. Coceal O, Belcher SE (2004) A canopy model of mean winds through urban areas. Q J Roy Meteorol Soc 130(599):1349–1372. https://doi.org/10.1256/qj.03.40
15. Colin J, Faivre R (2010) Aerodynamic roughness length estimation from very high-resolution imaging lidar observations over the heihe basin in china. Hydrol Earth Syst Sci 14:2661–2669. https://doi.org/10.5194/hess-14-2661-2010
16. Davenport AG (1967) The dependence of wind loads on meteorological parameters. In: Conference on wind loads on buildings. University of Toronto Press, Toronto. Paper 2
17. Deacon EL (1953) Vertical profiles of mean wind in the surface layers of the atmosphere, Meteorol Off Geoph Mem 91. Technical Report Meteorological Office, United Kingdom
18. DeBruin HAR, Moore CJ (1985) Zero-plane displacement and roughness length for tall vegetation, derived from a simple mass conservation hypothesis. Bound-Lay Meteorol 31(1):39–49
19. Dobos E (2006) Albedo. Encyclopedia of soil science, vol 2, pp 24–25
20. Dong Z, Gao S, Fryrear DW (2001) Drag coefficients, roughness length and zero-plane displacement height as disturbed by artificial standing vegetation. J Arid Environ 49(3):485–505. https://doi.org/10.1006/jare.2001.0807
21. Dorman JL, Sellers PJA (1989) Global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB). J Appl Meteorol 28(9):833–855
22. Drennan WM, Taylor PK, Yelland MJ (2005) Parameterizing the sea surface roughness. J. Phys. Oceanogr. 35(5):835–848. https://doi.org/10.1175/jpo2704.1

23. ESDU (1972) Characteristics of wind speed in the lower layers of the atmosphere near the ground: Strong winds (neutral atmosphere). Technical Report Engineering Sciences Data Unit, Regent Street, London, UK
24. Fry JA, Xian G, Jin S, Dewitz JA, Homer CG, Yang L, Barnes CA, Herold ND, Wickham JD (2011) Completion of the 2006 national land cover database for the conterminous United States. Photogramm Eng Rem S 77(9):858–864
25. Gao Z, Wang L, Bi X, Song Q, Gao Y (2012) A simple extension of "an alternative approach to sea surface aerodynamic roughness" by Zhiqiu Gao, Qing Wang, and Shouping Wang. J Geophys Res 117(D16110):1–8. https://doi.org/10.1029/2012jd017478
26. Garratt JR (1980) Surface influence upon vertical profiles in the atmospheric near-surface layer. Q J Roy Meteorol Soc 106(450):803–819. https://doi.org/10.1002/qj.49710645011
27. Giambelluca TW, Hölscher D, Bastos TX, Frazão RR, Nullet MA, Ziegler AD (1997) Observations of albedo and radiation balance over postforest land surfaces in the eastern amazon basin. J Clim 10(5):919–928
28. Graefe J (2004) Roughness layer corrections with emphasis on SVAT model applications. Agr Forest Meteorol 124(3–4):237–251. https://doi.org/10.1016/j.agrformet.2004.01.003
29. Graf A, van de Boer A, Moene A, Vereecken H (2014) Intercomparison of methods for the simultaneous estimation of zero-plane displacement and aerodynamic roughness length from single-level eddy-covariance data. Bound-Lay Meteorol 151(2):373–387. https://doi.org/10.1007/s10546-013-9905-z
30. Grimmond CSB, Oke TR (1999) Aerodynamic properties of urban areas derived from analysis of surface form. J Appl Meteorol 38(9):1262–1292
31. Hammond DS, Chapman L, Thornes JE (2012) Roughness length estimation along road transects using airborne LIDAR data. Meteorol Appl 19(4):420–426. https://doi.org/10.1002/met.273
32. Hanna SR, Chang JC (1992) Boundary layer parameterizations for applied dispersion modeling over urban areas. Bound-Lay Meteorol 58(3):229–259. https://doi.org/10.1007/BF02033826
33. Hansen FV (1993) Albedos. Technical Report ARL-TR-57, U.S. Army Research Laboratory, AMSRL-BE, White Sands Missile Range, NM 88002-5501
34. Hansen FV (1993) Surface roughness lengths. Technical Report ARL-TR-61, U.S. Army Research Laboratory, AMSRL-BE, White Sands Missile Range, NM 88002-5501
35. Hazeu GW, Schuiling C, Dorland GJ, Roerink GJ, Naeff HSD, Smidt RA (2014) Landelijk Grondgebruiksbestand Nederland versie 7 (LGN7). vervaardiging, nauwkeurigheid en gebruik. Technical Report Alterra Wageningen UR, Wageningen. In Dutch
36. Hicks BB, Hyson P, Moore CJ (1975) A study of eddy fluxed over a forest. J Appl Meteorol 14(1):58–66
37. Holland DE, Berglund JA, Spruce JP, McKellip RD (2008) Derivation of effective aerodynamic surface roughness in urban areas from airborne lidar terrain data. J Appl Meteorol Clim 47(10):2614–2626. https://doi.org/10.1175/2008JAMC1751.1
38. Holt T, Pullen J (2007) Urban canopy modeling of the New York city metropolitan area: a comparison and validation of single- and multilayer parameterizations. Mon Weather Rev 135(5):1906–1930. https://doi.org/10.1175/mwr3372.1
39. Hsu SA (1974) A dynamic roughness equation and its application to wind stress determination at the air-sea interface. J Phys Oceanogr 4(1):116–120. https://doi.org/10.1175/1520-0485(1974)004<0116:ADREAI>2.0.CO;2
40. van den Hurk BJJM (1995) Sparse canopy parameterization for meteorological models. PhD thesis, Department of Meteorology, WAU, Wageningen, The Netherlands
41. Hurtalová T, Matejka F (1999) Surface characteristics and energy fluxes above different plant canopies. Agr Forest Meteorol 98–99:491–500. https://doi.org/10.1016/S0168-1923(99)00118-5
42. Jacobs AFG, Boxel JH (1988) Changes of the zero-plane displacement and aerodynamic roughness length of maize during the growing season. Agr Forest Meteorol 42:53–62. https://doi.org/10.1016/0168-1923(88)90066-4

43. Kalma JD, Fuchs M (1976) Citrus orchards. In: Monteith JL (ed) Vegetation and the atmosphere. Academic Press, London
44. Kanda M, Kanega M, Kawai T, Moriwaki R, Sugawara H (2007) Roughness lengths for momentum and heat derived from outdoor urban scale models. J Appl Meteorol Clim 46(7):1067–1079. https://doi.org/10.1175/jam2500.1
45. Kanda M, Moriwaki R, Roth M, Oke T (2002) Area-averaged sensible heat flux and a new method to determine zero-plane displacement length over an urban surface using scintillometry. Bound-Lay Meteorol 105(1):177–193. https://doi.org/10.1023/a:1019668424982
46. Kimura R, Kondo J (1998) Heat balance model over a vegetated area and its application to a paddy field. J Meteorol Soc Jpn Ser II 76(6):937–953
47. Kondo J, Yamazawa H (1986) Aerodynamic roughness over an inhomogeneous ground surface. Bound-Lay Meteorol 35(4):331–348. https://doi.org/10.1007/bf00118563
48. Kustas WP, Choudhury BJ, Kunkel KE, Gay LW (1989) Estimate of the aerodynamic roughness parameters over an incomplete canopy cover of cotton. Agr Forest Meteorol 46(1–2):91–105. https://doi.org/10.1016/0168-1923(89)90114-7
49. Landsberg J, Powell D, Butler D (1973) Microclimate in an apple orchard. J Appl Ecol 881–896
50. Larson DW, Matthes U, Kelly PE (2005) Cliff ecology: pattern and process in cliff ecosystems. Cambridge University Press
51. Lettau HH (1969) Note on the aerodynamic roughness-parameter estimation on the basis of roughness-element description. J Appl Meteorol 8(5):828–832
52. Li X, Feng G, Sharratt B, Zheng Z (2015) Aerodynamic properties of agricultural and natural surfaces in northwestern Tarim Basin. Agr Forest Meteorol 204:37–45. https://doi.org/10.1016/j.agrformet.2015.01.005
53. MacArthur CD, Haines PA (1982) The roughness lengths associated with regions of heterogeneous vegetation and elevation. Technical Report Contract DAAD7-8-D-0206, University of Dayton Research Institute, Dayton, OH 45469
54. Mahrt L, Vickers D, Edson J, Wilczak JM, Hare J, Højstrup J (2001) Vertical structure of turbulence in offshore flow during rasex. Bound-Lay Meteorol 100(1):47–61
55. Medeiros SC, Hagen SC, Weishampel JF (2012) Comparison of floodplain surface roughness parameters derived from land cover data and field measurements. J Hydrol 452–453:139–149. https://doi.org/10.1016/j.jhydrol.2012.05.043
56. Millward-Hopkins JT, Tomlin AS, Ma L, Ingham D, Pourkashanian M (2011) Estimating aerodynamic parameters of urban-like surfaces with heterogeneous building heights. Bound-Lay Meteorol 141(3):443–465. https://doi.org/10.1007/s10546-011-9640-2
57. Mochida A, Murakami S, Ojima T, Kim S, Ooka R, Sugiyama H (1997) CFD analysis of mesoscale climate in the greater tokyo area. J Wind Eng Ind Aerod 67-68:459–477. https://doi.org/10.1016/s0167-6105(97)00060-3
58. Molero-Paredes T, Matos A (2008) Efectos de la inducción artificial de la poliploidia en plantas de aloe vera(l.). Boletín del Centro de Investigaciones Biológicas 42(1):111–133. (In Spanish)
59. Monteith JL, Unsworth MH (1990) Principles of environmental physics. Edward Arnold, LTD, New York, 291 pp
60. Montero G, Rodríguez E, Oliver A, Calvo J, Escobar JM, Montenegro R (2018) Optimisation technique for improving wind downscaling results by estimating roughness parameters. J Wind Eng Ind Aerod 174:411–423
61. Moore PA, Pypker TG, Waddington JM (2013) Effect of long-term water table manipulation on peatland evapotranspiration. Agric Forest Meteorol 178–179:106–119. https://doi.org/10.1016/j.agrformet.2013.04.013
62. Nakai T, Sumida A, Daikoku K, Matsumoto K, van der Molen MK, Kodama Y, Kononov AV, Maximov TC, Dolman AJ, Yabuki H, Hara T, Ohta T (2008) Parameterisation of aerodynamic roughness over boreal, cool- and warm-temperate forests. Agric Forest Meteorol 148(12):1916–1925. https://doi.org/10.1016/j.agrformet.2008.03.009

63. Nakai T, Sumida A, Matsumoto K, Daikoku K, Iida S, Park H, Miyahara M, Kodama Y, Kononov AV, Maximov TC, Yabuki H, Hara T, Ohta T (2008) Aerodynamic scaling for estimating the mean height of dense canopies. Bound-Lay Meteorol 128(3):423–443. https://doi.org/10.1007/s10546-008-9299-5
64. National Technique Team SIOSE (2011) Documento Técnico SIOSE2005—Versión 2.2. Technical Report, D.G. Instituto Geográfico Nacional, Madrid. (In Spanish)
65. National Technique Team SIOSE (2011) Manual de Fotointerpretación SIOSE—Versión 2. Technical Report D.G. Instituto Geográfico Nacional, Madrid. (In Spanish)
66. Oke TR (2002) Boundary layer climates. Routledge
67. Parlange MB, Brutsaert W (1989) Regional roughness of the landes forest and surface shear stress under neutral conditions. Bound-Lay Meteorol 48(1–2):69–81. https://doi.org/10.1007/BF00121783
68. Pires LBM, Fisch G, Gielow R, Souza LF, Avelar AC, De-Paula IB, Girardi RDM (2015) A study of the internal boundary layer generated at the alcantara space center. Am J Environ Eng 5(1A):82–164. https://doi.org/10.5923/s.ajee.201501.08
69. Qualls RJ, Brutsaert W (1996) Effect of vegetation density on the parameterization of scalar roughness to estimate spatially distributed sensible heat fluxes. Water Resour Res 32(3):645–652
70. Randall JM (1969) Wind profiles in an orchard plantation. Agr Forest Meteorol 6(6):439–452
71. Ratto C (1996) The aiolos and wlnds codes. In: Modelling of atmospheric flow fields, World Scientific, pp 421–431
72. Riou C, Pieri P, Valancogne C (1987) Variation de la vitesse du vent à l'intérieur et au-dessus d'une vigne. Agric Forest Meteorol 39:143–154 In French
73. Roballo ST, Fisch G (2008) Escoamento atmosfétrico no Centro de Lançamento de Alcântara (CLA): Parte I - Aspectos observcionais. Rev Bras Meteorol 23(4):510–519. https://doi.org/10.1590/S0102-77862008000400010. (In Portuguese)
74. Rotach MW (1994) Determination of the zero plane displacement in an urban environment. Bound-Lay Meteorol 67(1–2):187–193. https://doi.org/10.1007/BF00705513
75. Sacré C, Moisselin JM, Sabre M, Flori JP, Dubuisson B (2007) A new statistical approach to extreme wind speeds in france. J Wind Eng Ind Aerod 95(9–11):1415–1423. https://doi.org/10.1016/j.jweia.2007.02.013
76. Schmid HP, Bunzli B (1995) The influence of surface texture on the effective roughness length. Q J Roy Meteorol Soc 121(521):1–21. https://doi.org/10.1002/qj.49712152102
77. Sempreviva, A.: Roughness changes: response of neutral boundary layers. In: Modelling of atmospheric flow fields, World Scientific, pp 213–245
78. Shimoyama K, Hiyama T, Fukushima Y, Inoue G (2004) Controls on evapotranspiration in a west siberian bog. J Geophys Res Atmos 109(D8):1–12. https://doi.org/10.1029/2003JD004114
79. Stanhill G (1969) A simple instrument for the field measurement of a turbulent diffusion flux. J Appl Meteorol 8(4):509–513
80. Steyaert LT, Knox RG (2008) Reconstructed historical land cover and biophysical parameters for studies of land-atmosphere interactions within the eastern united states. J Geophys Res Atmos 113(D2)
81. Su Z (2006) An introduction to the surface energy balance system (SEBS). Lecture notes, ESA TIGER capacity building facility 1st training course on advanced optical remote sensing
82. Su Z, Schmugge T, Kustas WP, Massman WJ (2001) An evaluation of two models for estimation of the roughness height for heat transfer between the land surface and the atmosphere. J Appl Meteorol 40(11):1933–1951
83. Takagi K, Miyata A, Harazono Y, Ota N, Komine M, Yoshimoto M (2003) An alternative approach to determining zero-plane displacement, and its application to a lotus paddy field. Agric Forest Meteorol 115:173–181
84. Tanaka S, Sugawara H, Narita K, Yokoyama H, Misaka I, Matsushima D (2011) Zero-Plane displacement height in a highly built-up area of Tokyo. SOLA 7:93–96. https://doi.org/10.2151/sola.2011-024

85. Taylor PA (1987) Comments and further analysis on effective roughness lengths for use in numerical three-dimensional models. Bound-Lay Meteorol 39:403–418
86. Thom AS (1971) Momentum absorption by vegetation. Q J Roy Meteorol Soc 97(414):414–428
87. Thom AS (1972) Momentum, mass and heat exchange of vegetation. Q J Roy Meteorol Soc 98(415):124–134
88. Tian X, Li ZY, van der Tol C, Su Z, Li X, He QS, Bao YF, Chen EX, Li LH (2011) Estimating zero-plane displacement height and aerodynamic roughness length using synthesis of lidar and spot-5 data. Remote Sens Environ 115(9):2330–2341. https://doi.org/10.1016/j.rse.2011.04.033
89. Tieleman HW (2003) Roughness estimation for wind-load simulation experiments. J Wind Eng Ind Aerod 91(9):1163–1173. https://doi.org/10.1016/S0167-6105(03)00058-8
90. Troen I, Petersen EL (1989) European wind atlas, 300 pp. Technical Report Risø National Laboratory, Roskilde, Denmark
91. van Wijk BM (2011) Predicting the rooftop wind climate for urban wind energy in the Rotterdam - Delft - Zoetermeer region. new approaches for implementing urban height data in the wind atlas method. Technical Report Eindhoven University of Technology, Eindhoven. 114 pp
92. Verhoef A, McNaughton KG, Jacobs AFG (1997) A parameterization of momentum roughness length and displacement height for a wide range of canopy densities. Hydrol Earth Syst Sci 1(1):81–91. https://doi.org/10.5194/hess-1-81-1997
93. Villalobos FJ, Orgaz F, Testi L, Fereres E (2000) Measurement and modeling of evapotranspiration of olive (Olea europaea L.) orchards. Eur J Agron 13(2–3):155–163. https://doi.org/10.1016/s1161-0301(00)00071-x
94. Warner TT (2009) Desert meteorology. Cambridge University Press
95. Weiss A, Allen LH (1976) Vertical and horizontal air flow above rows of a vineyard. Agr Meteorol 17(6):433–452. https://doi.org/10.1016/0002-1571(76)90021-2
96. Wieringa J (1992) Updating the davenport roughness classification. J Wind Eng Ind Aerod 41(1–3):357–368. https://doi.org/10.1016/0167-6105(92)90434-C
97. WindSim AS (2014) Wind Resource Assessment—Annual Energy Production. Project: Hundhammer_WS_Express. Technical Report WindSim AS, Fjordgaten 15, N- 3125 Tønsberg, Norway. 29 pp
98. de Wit AJW, Van der Heijden TGC, Thunnissen HAM (1999) Vervaardiging en nauwkeurigheid van het LGN3-grondgebruiksbestand. Technical Report DLO-Staring Centrum, Wageningen In Dutch
99. Yang K, Koike T, Ishikawa H, Kim J, Li X, Liu H, Liu S, Ma Y, Wang J (2008) Turbulent flux transfer over bare-soil surfaces: Characteristics and parameterization. J Appl Meteorol Clim 47(1):276–290. https://doi.org/10.1175/2007jamc1547.1
100. Yang R, Friedl MA (2003) Determination of roughness lengths for heat and momentum over boreal forests. Bound-Lay Meteorol 107(3):581–603. https://doi.org/10.1023/a:1022880530523
101. Yoon JJ, Shim JS, Park KS, Lee JC (2014) Numerical experiments of storm winds, surges, and waves on the southern coast of Korea during Typhoon Sanba: the role of revising wind force. Nat Hazards Earth Sys Sci 14(12):3279–3295. https://doi.org/10.5194/nhess-14-3279-2014
102. Zannetti P (1990) Air pollution modeling. Comp. Mech. Publications, Southampton (UK)
103. Zilitinkevich SS, Esau IN (2002) On integral measures of the neutral barotropic planetary boundary layer. Bound-Lay Meteorol 104:371–379
104. Zilitinkevich SS, Tyuryakov SA, Troitskaya YI, Mareev EA (2012) Theoretical models of the height of the atmospheric boundary layer and turbulent entrainment at its upper boundary. Atmos Ocean Phys 48(1):150–160. https://doi.org/10.1134/S0001433812010148

Discretization of the Region of Interest

J. Manuel Cascón, José María Escobar and Rafael Montenegro

Abstract The meccano method was recently introduced to construct simultaneously tetrahedral meshes and volumetric parameterizations of solids. The method requires the information of the solid geometry that is defined by its surface, a meccano, i.e., an outline of the solid defined by connected polyhedral pieces, and a tolerance that fixes the desired approximation of the solid surface. The method builds an adaptive tetrahedral mesh of the solid (physical domain) as a deformation of an appropriate tetrahedral mesh of the meccano (parametric domain). The main stages of the procedure involve an admissible mapping between the meccano and the solid boundaries, the nested Kossaczký's refinement, and our simultaneous untangling and smoothing algorithm. In this chapter, we focus on the application of the method to build tetrahedral meshes over complex terrain, that is interesting for simulation of environmental processes. A digital elevation map of the terrain, the height of the domain, and the required orography approximation are given as input data. In addition, the geometry of buildings or stacks can be considered. In these applications, we have considered a simple cuboid as meccano.

J. M. Cascón (✉)
Deparment of Economics and Economic History, Faculty of Economic and Business, University of Salamanca. Edificio FES, Campus Miguel de Unamuno, 37008 Salamanca, Spain
e-mail: casbar@usal.es

J. M. Escobar · R. Montenegro
University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain
e-mail: josem.escobar@ulpgc.es

R. Montenegro
e-mail: rafael.montenegro@ulpgc.es

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_3

1 Introduction

Mesh generation is one of the most time-consuming processes in numerical simulations of engineering problems based on partial differential equations. On one hand, an accurate discretization of the physical domain is fundamental to obtain a realistic simulation. On the other hand, it is well known that the quality of the mesh plays a crucial role in the accuracy and stability of the numerical computation.

Many authors have devoted a great effort to solving the automatic mesh generation problem in different ways [3, 12, 13, 24]. Along the past, the main objective has been to achieve high-quality adaptive meshes of complex solids with minimal user intervention and low computational cost. At present, it is well known that most mesh generators are based on Delaunay triangulation and advancing front technique, but problems, related to mesh quality or mesh conformity with the solid boundary, can still appear for complex geometries. In addition, an appropriate definition of element sizes is demanded for obtaining good-quality elements and mesh adaption. Particularly, local-adaptive refinement strategies have been employed to mainly adapt the mesh to singularities of numerical solution. These adaptive methods usually involve remeshing or nested refinement [4, 14, 16, 17, 23].

In this direction, we have introduced the meccano technique in [5, 6, 18–20] for constructing adaptive tetrahedral meshes of solids. This algorithm requires a coarse computational domain, then builds a surface parameterization and combines refinement and mesh optimization to produce an adaptive tetrahedral mesh of the input domain. As a result, a piecewise linear volumetric parameterization is obtained. The name of the method is due that the process starts from a coarse approximation of the solid, i.e., a meccano composed by connected polyhedral pieces.

The rest of the chapter is structured as follows. In Sect. 2, we overview the meccano method for a general solid. In Sect. 3, we detail the algorithm to the construction of the tetrahedral meshes over complex terrain. Finally, in Sect. 4, we present several examples that illustrate the capabilities of the method.

2 The Meccano Method

The meccano method is a tetrahedral mesh generator [6, 18–20]. The method requires a surface triangulation of the solid boundaries and a computational domain that coarsely approximates the solid. This computational domain is called meccano. The procedure builds an adaptive tetrahedral mesh in the meccano and deforms it to match the physical domain. For this purpose, the method combines several procedures: an automatic mapping from the boundary of the meccano to the boundary of the solid, a 3-D local refinement algorithm, and a simultaneous mesh untangling and smoothing. It is important to point out that this method also provides a continuous element-wise linear volumetric parameterization from the computational domain to the solid.

Algorithm 1 Meccano tetrahedral mesh generation.

1: **function** MECCANOMESHER(Solid Ω, Real ε)
2: Meccano $\mathscr{M} \leftarrow$ GETMECCANO(Ω)
3: Mapping $\boldsymbol{\Pi} \leftarrow$ GETBOUNDARYMAPPING($\mathscr{M}$, Ω)
4: Mesh $\mathscr{T} \leftarrow$ GETINITIALMESH($\mathscr{M}$, $\boldsymbol{\Pi}$)
5: **while** DISTANCE($\partial\mathscr{T}$, $\partial\Omega$) $> \varepsilon$ **do**
6: TriangleList $\mathscr{R} \leftarrow$ GETTRIANGLESTOREFINE($\mathscr{T}$, Ω, ε)
7: TriangleList $\widehat{\mathscr{R}} \leftarrow$REFINETRIANGLES($\mathscr{R}$)
8: PROJECTNEWNODESTOBOUNDARY($\widehat{\mathscr{R}}$, $\boldsymbol{\Pi}$)
9: **end while**
10: QUALITYOPTIMIZATION($\mathscr{T}$)
11: **end function**

The construction of the meccano (the computational domain) is not automatic for complex solids (genus bigger than zero) and requires user intervention. However, in this chapter, we focus on the generation of tetrahedral meshes over complex terrain, that are interesting for simulation of environmental processes, and whose boundary is genus zero. Therefore, the meccano algorithm is a fully automatic procedure.

The main steps of the meccano tetrahedral mesh generation algorithm are summarized in Algorithm 1. The input data is a solid, Ω, defined by its boundary representation (surface triangulation or CAD model), and a given precision to approximate its boundary, ε.

The first step of the procedure, Line 2, is to construct a meccano, $\mathscr{M}$, that approximates the solid, by connecting polyhedral pieces. Then, in Line 3, a discrete mapping, $\boldsymbol{\Pi}$, between the boundary of the meccano and the boundary of the solid is computed using a procedure based on the mean value parametrization proposed by Floater in [10, 11]. Note that this parameterization is a continuous and element-wise linear mapping. In Line 4, an initial coarse mesh of the meccano is generated, and the boundary nodes are located on the solid boundary using the mapping $\boldsymbol{\Pi}$. In Lines 5–9, we obtain a mesh that approximates the solid boundary with the given tolerance ε. Specifically, in Line 6, we get the list of triangles that do not correctly approximate the boundary of the solid, and then, in Line 7, we refine those triangles by dividing their adjacent tetrahedra using the Kossaczký method [16]. In Line 8, we project the new nodes onto the boundary of the solid using the mapping $\boldsymbol{\Pi}$. We iterate this process until there are no triangles to refine. Note that when the nodes are mapped onto the solid boundary, low-quality and inverted elements may appear. Thus, a simultaneous untangling and smoothing procedure [8] is applied in order to obtain a valid and high-quality tetrahedral mesh. As commented before, the meccano method automatically provides a volumetric parameterization from the computational domain to the solid. We show in Fig. 1 an example, where different steps of the meccano algorithm are sumarized.

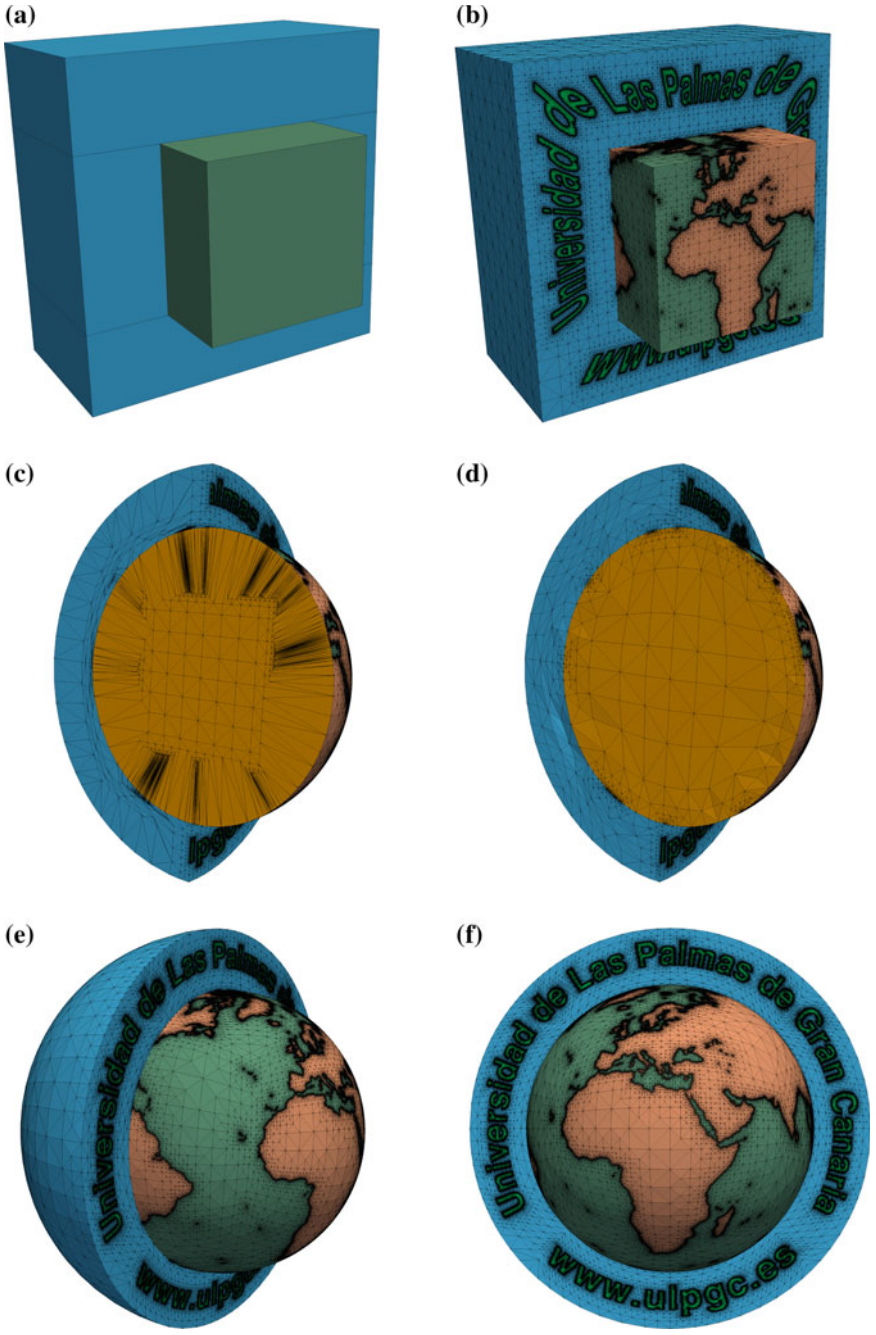


Fig. 1 The different meccano steps. **a** Parametric space, **b** adaptive triangulation of the meccano boundary, **c** tangled interior elements after surface mapping, **d** optimized mesh, **e** resulting surface mesh, **f** frontal view

3 Meccano Method for the Construction of 3D Meshes over Complex Topography

In this section, we describe, in detail, the meccano method for the discretization of three-dimensional domain that is limited in its lower part by a complex terrain, and in its upper part by a rectangular horizontal plane region placed at a given height. The lateral walls are formed by four vertical planes. The generated tetrahedral mesh is appropriated for finite element simulations of environmental processes, such as weather forecasting, wind field (see Chap. 4), fire propagation [22], or atmospheric pollution [21].

In this case, the meccano consists of a simple cuboid. A digital elevation map of the terrain and the required orography approximation are given as input data. The mapping between the meccano and domain boundaries is obtained by applying an automatic Floater's parameterization [10, 11].

3.1 Meccano

As we commented before, a simple cuboid, $\mathscr{M}$, is defined as meccano. Its upper face coincides with the upper boundary of the domain, and its lower face is placed at the minimum terrain height of the rectangular region.

We now divide the surface of the domain into six trivial patches (lower and upper bound and lateral walls), and associate then with counterparts in the meccano boundary. This correspondence must be compatible, in the sense, that if two patches of the domain have nonempty intersection, their corresponding images on meccano boundary also satisfy this property.

3.2 Parametrization of the Domain Boundary

Once the meccano is fixed, we have to determine a mapping between the cuboid faces and the domain boundary. For that purpose, a discrete mapping from each surface patch to the corresponding cuboid face is built using the mean value parameterization proposed in [11]. In order to get an admisible mapping, we note that the discrete mappings have to coincide on the intersection of their associate patches.

We now describe how the parametrization Π_b of the lower bound $\partial\Omega_b$ of the physical space, can be obtained. The other ones (upper bound and lateral walls) are simpler, and could be generated in the same way. In fact, in this particular case, the parametrization of the upper bound is the identity. The method was introduced by Floater [11], and provides a parametrization of a simply connected surface triangulation.

We assume that the digital elevation map, that captures the orography, is given by a triangular mesh embedded in 3D, that we denote $\mathscr{T}_b$. The bottom cuboid face $\partial\mathscr{M}_b$ will be the parametric space. Then, we find a mapping

$$\Pi_b : \partial\mathscr{M}_b \to \partial\Omega_b$$

continuous and piecewise linear, where $\tau_b = (\Pi_b)^{-1}(\mathscr{T}_b)$ will be the planar triangulation of $\partial\mathscr{M}_b$ associated to $\mathscr{T}_b$. Note that the construction of the mapping is equivalent to find an *admissible* localization of nodes of the planar triangulation τ_b. The Floater solution first fixes the boundary nodes of τ_b and then the position of the inner ones is given by the solution of a linear system based on convex combinations.

Formally, let $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be the inner nodes and $\{\mathbf{x}_{n+1}, \ldots, \mathbf{x}_N\}$ be the boundary nodes of $\mathscr{T}_b$, respectively, where N denotes the total number of nodes of $\mathscr{T}_b$. Fixed the position of boundary nodes $\{\mathbf{y}_{n+1}, \ldots, \mathbf{y}_N\}$ of τ_b, the position of its inner nodes $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ is given by the solution of the system:

$$\mathbf{y}_k = \sum_{l=1}^{N} \lambda_{kl}\mathbf{y}_l, \qquad k = 1, \ldots, n.$$

The values of $\{\lambda_{kl}\}_{k=1,\ldots,n}^{l=1,\ldots,N}$ are the weights of the convex combinations, such that

$$\begin{aligned} \lambda_{kl} &= 0, && \text{if } \mathbf{x}_k \text{ and } \mathbf{x}_l \text{ are not connected} \\ \lambda_{kl} &> 0, && \text{if } \mathbf{x}_k \text{ and } \mathbf{x}_l \text{ are connected} \\ \sum_{l=1}^{N} \lambda_{kl} &= 1, && \text{for } k = 1, \ldots n. \end{aligned}$$

The quality parametrization depend on the weights λ_{kl}. In [10] three alternatives are analyzed: uniform parametrization, weighted least squares of edge lengths and shape preserving parametrization. Another choice, called mean value coordinate, is presented in [11]. The goal is to obtain an approximation of a conformal mapping.

Finally, we remark that in order to get an admissible mapping between meccano boundary, $\partial\mathscr{M}$, and domain boundary, $\partial\Omega$, the boundary nodes that the Floater algorithm fixes for the parameterization of each face, must coincide on their common cuboid edges.

In Fig. 2 we show a surface triangulation, that approximates the orography of an area of Comunidad de Madrid (Spain), and its corresponding parameterization space obtained by the procedure described in this section.

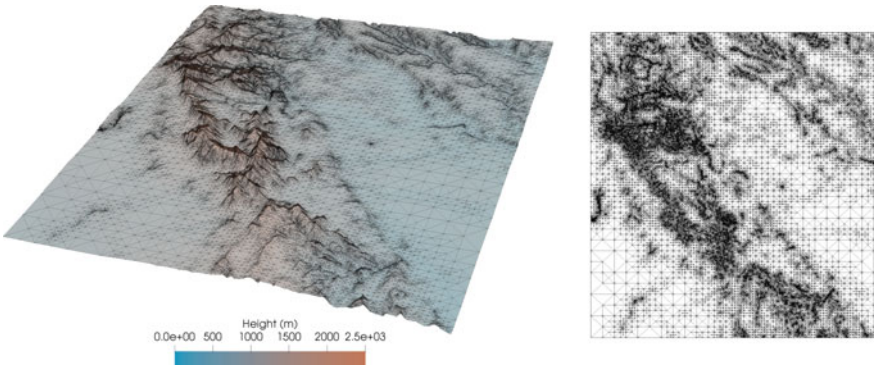


Fig. 2 Orography of an area of Comunidad de Madrid (Spain) and its corresponding parameterization space

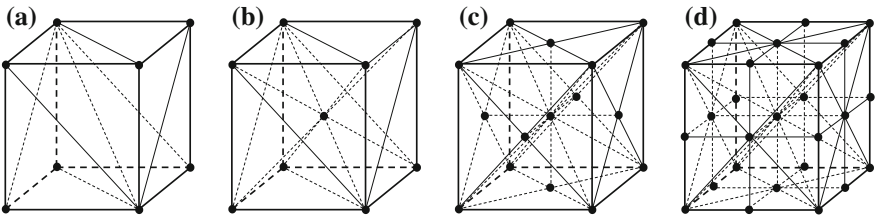


Fig. 3 Refinement of a cube by using Kossaczký's algorithm: **a** cube subdivision into six tetrahedra, **b** bisection of all tetrahedra by inserting a new node in the cube main diagonal, **c** new nodes in diagonals of cube faces and **d** global refinement with new nodes in cube edges

3.3 Coarse Tetrahedral Mesh of the Meccano

We build a coarse and high-quality tetrahedral mesh, $\mathscr{T}_0(\mathscr{M})$ by splitting the cuboid in cubes, and each cube is subdivided into six tetrahedra [16]. For this purpose, it is necessary to define a main diagonal on the cube and corresponding diagonal on its faces, see Fig. 3a. The resulting mesh can be recursively and globally bisected [16] for fixing a uniform element size in the whole mesh. Three consecutive global bisections for a cube are presented in Figs. 3b–d. The resulting mesh of Fig. 3d contains 8 cubes similar to the one shown in Fig. 3a. Therefore, the recursive refinement of the cube mesh produces similar tetrahedra to the initial ones.

3.4 Approximation of the Orography: Refinement and Distance Evaluation

The next step of the meccano algorithm is to approximate the orography of the terrain with a prescribed tolerance ϵ. In fact, according to the choice of the meccano, see Sect. 3.1, we only need to impose the approximation criteria on the bottom face of the meccano, $\partial\Omega_b$.

In order to obtain this approximation, the initial mesh $\mathscr{T}_0(\mathscr{M})$, is recursively refined. Specifically, we select the set of triangles on the bottom of the meccano, $T \in \partial\mathscr{T}(\mathscr{M}_b)$, that do not correctly approximate the terrain of the domain, that is

$$T \equiv \langle v_0, v_1, v_2\rangle \in \partial\mathscr{T}(\mathscr{M}_b), \quad d(\langle \Pi(v_0), \Pi(v_1), \Pi(v_2)\rangle, \partial\Omega_b) > \epsilon \tag{1}$$

where, $\langle \Pi(v_0), \Pi(v_1), \Pi(v_2)\rangle$ denotes the triangle that has as vertices $\{\Pi(v_i)\}_{i=0}^2$. Then we refine those triangles by dividing their adjacent tetrahedra using the Kossaczký method [16]. This procedure is iterated until the prescribed tolerance is reached.

We use the following strategy for a practical computation of the distance. Given $T \in \partial\mathscr{T}(\mathscr{M}_b)$, let $\{q_i\}_i^k$ be a set of Gauss quadrature points, then $d(\Pi(T), \partial\Omega_b)$ is estimated as

$$d(\langle \Pi(v_0), \Pi(v_1), \Pi(v_2)\rangle, \partial\Omega_b) \approx \max_{i=1,\dots,k} d(\langle \Pi(v_0), \Pi(v_1), \Pi(v_2)\rangle, \Pi(q_i)). \tag{2}$$

Once the final tetrahedral mesh $\mathscr{T}(\mathscr{M})$ is generated in the cuboid, their boundary nodes are mapped by Π onto the boundary of the domain. After this step, low-quality and inverted elements may appear.

3.5 Relocation of Inner Nodes

There would be several strategies for defining a reasonable position for each inner node of the domain: Laplacian smoothing, Coons patches, radial basis functions, etc.

Another effective possibility hinges on the volumetric mapping that produces the meccano method (see Sect. 3.7). However, this information is not known a priori. In fact, we will reach this piecewise linear volume mapping at the end of the mesh generation.

In practice, a good strategy is: we start meshing the solid by using a high value of ε (a coarse tetrahedral mesh of the solid is obtained) and we continue decreasing it gradually. In the first step of this strategy, no relocation is applied. In this case, the number of nodes of the resulting mesh is low and the mesh optimization algorithm, that we describe below, is fast. In the following steps a relocation of inner nodes is applied by using the mapping (volumetric parameterization) that is defined by the previous iteration.

3.6 Simultaneous Untangling and Smoothing

Since the current mesh could be not valid (it could contain inverted element), it is necessary to optimize it. The process that we describe in this section [8, 9], must be able to smooth and untangle the mesh and is crucial in the proposed mesh generator.

Usual techniques to improve the quality of a *valid* mesh, that is, one that does not have inverted elements, are based upon local smoothing. In short, these techniques consist of finding the new positions that the mesh nodes must hold, in such a way that they optimize an objective function. Such a function is based on a certain measurement of the quality of the *local submesh* $N(q)$, formed by the set of tetrahedra connected to the *free node* q. Usually, objective functions are appropriate to improve the quality of a valid mesh, but they do not work properly when there are inverted elements. This is because they present singularities (barriers) when any tetrahedron of $N(q)$ changes the sign of its Jacobian.

Most of what is stated below are taken from [8], where we developed a procedure for untangling and smoothing tetrahedral meshes simultaneously. For that purpose, we use a suitable modification of the objective function such that it is regular all over $\mathscr{R}^3$. When a feasible region (subset of $\mathscr{R}^3$ where q could be placed, being $N(q)$ a valid submesh) exists, the minima of both the original and the modified objective functions are very close, and when this region does not exist, the minimum of the modified objective function is located in such a way that it tends to untangle $N(q)$. The latter occurs, for example, when the fixed boundary of $N(q)$ is tangled. With this approach, we can use any standard and efficient unconstrained optimization method to find the minimum of the modified objective function, see for example [1].

3.6.1 Objective Functions

Several tetrahedron shape measures could be used to construct an objective function. Nevertheless, those obtained by algebraic operations [15] are specially indicated for our purpose because they can be computed very efficiently and they allow us to choose the shape of the tetrahedra to optimize. Our objective is to relocate the nodes of $\mathscr{T}$ in positions where not only the mesh gets untangled, but also the distortion introduced by the parameterization is minimized.

Let T be a tetrahedral element of $\mathscr{T}$ whose vertices are $\mathbf{x}_k = (x_k, y_k, z_k)^T \in \mathscr{R}^3$, $k = 0, 1, 2, 3$ and T_R be the reference tetrahedron with vertices $\mathbf{u}_0 = (0, 0, 0)^T$, $\mathbf{u}_1 = (1, 0, 0)^T$, $\mathbf{u}_2 = (0, 1, 0)^T$ and $\mathbf{u}_3 = (0, 0, 1)^T$. If we choose $\mathbf{x}_0$ as the translation vector, the affine map that takes T_R to T is $\mathbf{x} = A\mathbf{u} + \mathbf{x}_0$, where A is the Jacobian matrix of the affine map referenced to node $\mathbf{x}_0$, and expressed as $A = (\mathbf{x}_1 - \mathbf{x}_0, \mathbf{x}_2 - \mathbf{x}_0, \mathbf{x}_3 - \mathbf{x}_0)$.

Let us consider that T_I is our ideal or target tetrahedron whose vertices are $\mathbf{v}_0$, $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$. If we take $\mathbf{v}_0 = (0, 0, 0)^T$ the linear map that takes T_R to T_I is $\mathbf{v} = W\mathbf{u}$, where $W = (\mathbf{v}_1 - \mathbf{v}_0, \mathbf{v}_2 - \mathbf{v}_0, \mathbf{v}_3 - \mathbf{v}_0)$ is its Jacobian matrix. As the parametric and real meshes are topologically identical, each tetrahedron of $\mathscr{T}$ has its counterpart in

$\mathscr{C}_K$. Thus, in order to reduce the distortion in the volumetric parameterization we will fix the target tetrahedra of $N(q)$ as their counterparts of the local mesh in the parametric space.

The affine map that takes T_I to T is $\mathbf{x} = AW^{-1}\mathbf{v} + \mathbf{x}_0$, and its Jacobian matrix is $S = AW^{-1}$. Note that this weighted matrix S depends on the node chosen as reference, so this node must be the same for T and T_I. We can use matrix norms, determinant or trace of S to construct algebraic quality metrics of T. For example, the *mean ratio*, $Q = \frac{3\sigma^{\frac{2}{3}}}{|S|^2}$, is an easily computable algebraic quality metric of T, where $\sigma = \det(S)$ and $|S|$ is the Frobenius norm of S. The maximum value of Q is the unity, and it is reached when $A = \mu RW$, where μ is a scalar and R is a rotation matrix. In other words, Q is maximum if and only if T and T_I are similar. Besides, any flat tetrahedron has quality measure zero. We can derive an optimization function from this quality metric. Thus, let $\mathbf{x} = (x, y, z)^T$ be the position of the free node, and let S_m be the weighted Jacobian matrix of the m-th tetrahedron of $N(q)$. We define the distortion, η_m, of the m-th tetrahedron as the inverse of its quality:

$$\eta_m = \frac{|S_m|^2}{3\sigma_m^{\frac{2}{3}}} \tag{3}$$

Then, the corresponding objective function for $N(q)$ is constructed by using the p-norm of $(\eta_1, \eta_2, \ldots, \eta_M)$ as

$$\left|K_\eta\right|_p(\mathbf{x}) = \left[\sum_{m=1}^{M} \eta_m^p(\mathbf{x})\right]^{\frac{1}{p}} \tag{4}$$

where M is the number of tetrahedra in $N(q)$. We obtain the optimal position of the free node by minimizing (4).

Although this optimization function is smooth in those points where $N(q)$ is a valid submesh, it becomes discontinuous when the volume of any tetrahedron of $N(q)$ goes to zero. It is due to the fact that η_m approaches infinity when σ_m tends to zero and its numerator is bounded below. In fact, it is possible to prove that $|S_m|$ reaches its minimum, with strictly positive value, when q is placed in the geometric center of the fixed face of the m-th tetrahedron. The positions where q must be located to get $N(q)$ to be valid, i.e., the feasible region, is the interior of the polyhedral set P defined as $P = \bigcap_{m=1}^{M} H_m$, where H_m are the half-spaces defined by $\sigma_m(\mathbf{x}) \geq 0$. This set can occasionally be empty, for example, when the fixed boundary of $N(q)$ is tangled. In this situation, function $\left|K_\eta\right|_p$ stops being useful as an optimization function. Moreover, when the feasible region exists, that is $int\ P \neq \emptyset$, the objective function tends to infinity as q approaches the boundary of P. Due to these singularities, it is formed a barrier which avoids reaching the appropriate minimum when using gradient-based algorithms, and when these start from a free node outside the feasible

region. In other words, with these algorithms we cannot optimize a tangled mesh $N(q)$ with the above objective function.

3.6.2 Modified Objective Functions

We proposed in [8] a regularization in the previous objective function (4), so that the barrier associated with its singularities will be eliminated and the new function will be smooth all over $\mathscr{R}^3$. An essential requirement is that the minima of the original and modified functions are nearly identical when $int\ P \neq \emptyset$. Our modification consists of substituting σ in (3) by the positive and increasing function

$$h(\sigma) = \frac{1}{2}(\sigma + \sqrt{\sigma^2 + 4\delta^2}) \tag{5}$$

being the parameter $\delta = h(0)$. Thus, the new objective function here proposed is given by

$$\left|K_\eta^*\right|_p(\mathbf{x}) = \left[\sum_{m=1}^{M}\left(\eta_m^*\right)^p(\mathbf{x})\right]^{\frac{1}{p}} \tag{6}$$

where

$$\eta_m^* = \frac{|S_m|^2}{3h^{\frac{2}{3}}(\sigma_m)} \tag{7}$$

is the modified objective function for the m-th tetrahedron. With this modification, we can untangle the mesh and, at the same time, improve its quality. An implementation of the simultaneous untangling and smoothing procedure for an equilateral reference tetrahedron is freely available in [9].

3.7 Volumetric Parameterization

One of the consequences of the meccano method is that it produces automatically a volumetric parametrization of the domain. This mapping $\boldsymbol{\Pi}$ is an extension of the surface parameterization Π built at Sect. 3.2:

$$\boldsymbol{\Pi} : \mathscr{M} \rightarrow \mathscr{T} \approx \Omega \tag{8}$$

where a point p included in a tetrahedron of $\mathscr{M}$ is mapped, preserving barycentric coordinates, into a point q belonging to the transformed tetrahedron of $\mathscr{T}$.

3.8 Meccano as Surface Mesher

Numerical simulation of some engineering/environmental problem, such as solar radiation (see Chap. 10), only requires a surface mesh. The meccano algorithm, that we describe along this section, can be particularized to generate triangular meshes of complex terrains. In this case, the meccano is a rectangle, that plays the role of parametric space of the surface defined by the elevation map of the terrain. Then, Floater's strategy (Sect. 3.2) allows to build a mapping between the meccano and the terrain surface. Finally, an analogous approximation strategy to Sect. 3.4 provides the final surface mesh.

Since the Floater's parameterization is a quasi-conformal mapping, the smoothing step can be avoided in general. Nevertheless, the surface smoothing [7] could be used to improve the mesh quality.

4 Numerical Examples

The performance of our new mesh generator is shown in the following applications. The first corresponds to a domain placed at La Palma (Canary Islands, Spain), of 10×30 km. In the second example, we discretize a bigger domain: Gran Canaria (Canary Island, Spain), 60×70 km, and we analyze several approximations for different values of ε parameter. Finally, in the third example, we include a chimney in a region of Gran Canaria Island and build the associated mesh. In all cases, the topography is given by a digitalization of the area where heights are defined over a uniform grid with a spacing step of 25 m in directions x and y (raster data). The computations were done on a MacBook Pro with two processors, 3.5 GHz, Intel Core, and 16 Gb RAM memory.

4.1 La Palma

We first consider a rectangular area in *Isla de La Palma* (Canary Islands, Spain) of 10×30 km. The upper boundary of the domain has been placed at 3 km. To define the topography we use a digitalization of the area where heights are defined over a uniform grid with a spacing step of 25 m in directions x and y. Therefore, the associated meccano is a cuboid of dimension $10 \times 30 \times 3$ km. After computing a boundary mapping, Π between the boundary of the meccano and the boundary of the domain (see Sect. 3.2 for details), we divided it into $5 \times 15 \times 1$ cuboids. Each cuboid is subdivided into six tetrahedra by using the subdivision proposed in [16], see Fig. 3. Then, we fix $\varepsilon = 5$ m, and begin the approximation procedure described at Sect. 3.4.

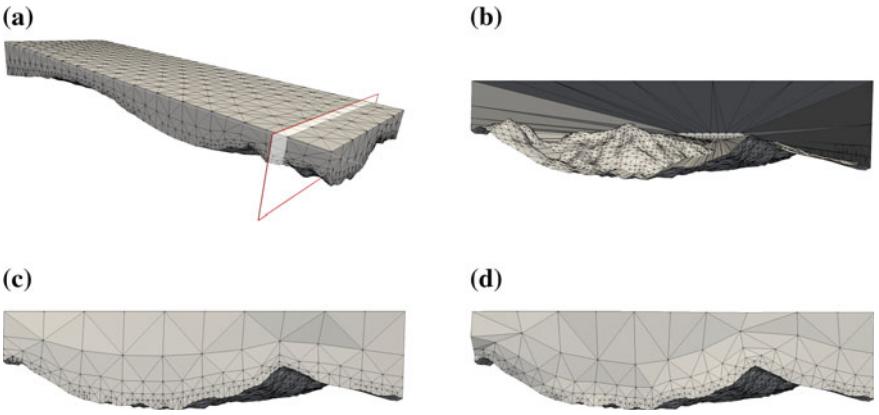


Fig. 4 **a** Clip of the mesh along the algorithm: **b** After node boundary projection, **c** After inner node relocation, **d** After smoothing (final mesh)

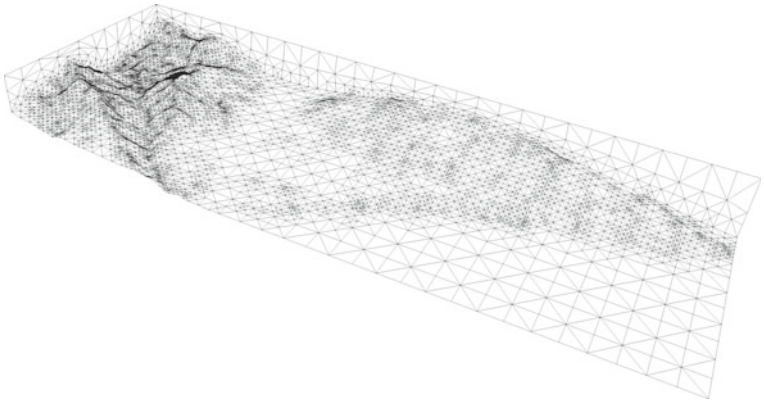


Fig. 5 Detail of *Isla de La Palma* (Canary Island): Surface of the final mesh

The resulting mesh has 75320 tetrahedra and 17664 nodes and it nears the terrain surface with an error less than $\varepsilon = 5$ m. The adaptive procedure performs 18 refinement steps (bisection). Besides, we relocate the inner nodes using the volumetric parameterization produced by a coarse ϵ ($\varepsilon = 100$ m.), and reduce the number of inverted tetrahedra from 23612 to 5266. Finally, we apply the optimization process of Sect. 3.6. The node distribution is hardly modified after 12 steps, resulting in a valid and high- quality mesh (the mesh is untangled after 8 iterations). We remark that we have not relocated those nodes placed on the terrain during this optimization process. In Fig. 4 we show several clips of the tetrahedral mesh along the generation procedure and in Fig. 5 the surface of the final mesh is presented.

The evolution of the mesh quality during the optimization process is shown in Fig. 6. These curves are obtained by sorting the elements in increasing order of

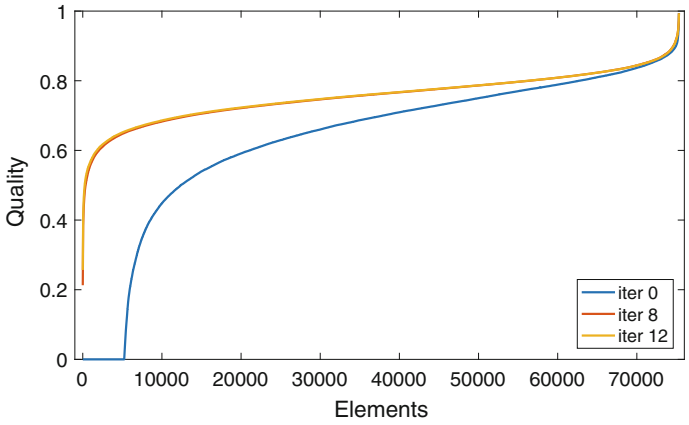


Fig. 6 Quality curves for the initial (blue line) and optimized meshes after 8th (red line) and 12th (orange line) iterations for the domain of La Palma (Canary Island)

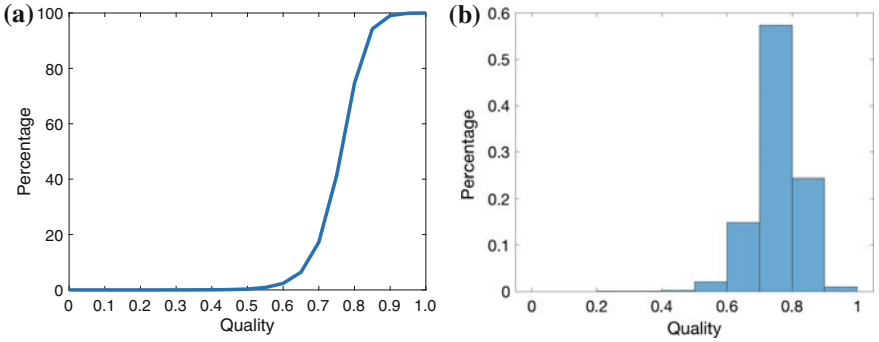


Fig. 7 **a** Cumulative frequency polygon of the final mesh; for a given quality value $x \in (0, 1)$ the line represents the percentage of elements that have a quality $\leq x$. **b** Quality histogram of the final mesh. The height of each bar is the relative number of tetrahedra with quality associate to the bar

its quality. This measure tends to stagnate quickly. Note that the quality curves corresponding to the eighth and twelfth optimization steps are very close. The average quality measure increases to $\overline{q}_\kappa = 0.75$. After this optimization process, the worst quality measure of the optimized mesh tetrahedra is 0.25. In addition, we present in Fig. 7a the 'cumulative frequency polygon' of the generated mesh, for a given value of $x \in (0, 1)$ the line represents the percentage of elements that have a quality $\leq x$ and in Fig. 7b the quality histogram of the final mesh: where the height of each bar is the relative number of tetrahedra with quality associate to the bar. We note that the 80% of tetrahedra have a quality bigger than 0.70.

The total CPU time for the mesh generation is less than 20 s. In particular, the computational cost of the simultaneous untangling and smoothing procedure is about 8 s.

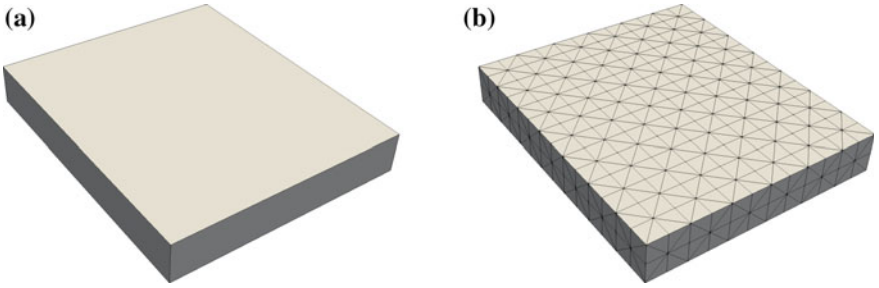


Fig. 8 Initial meccano of Gran Canaria Island (**a**) and its corresponding intial mesh (**b**)

Table 1 Main features of meshes of Gran Canaria Island for several values of the discretization parameter ε

ε (m)	Tetrahedra	Nodes	Ref. level	SUS iter. (untang.+smooth.)	$q_{\min}$	$q_{\max}$	CPU Time (s.)
200	8760	2127	12	$3+6$	0.45	0.77	$\approx$13
100	39800	9191	17	$3+6$	0.30	0.76	$\approx$22
75	70572	16231	18	$2+5$	0.31	0.76	$\approx$33
50	153252	35092	21	$2+5$	0.31	0.76	$\approx$71
35	278674	63740	24	$3+6$	0.32	0.76	$\approx$145
25	478420	109315	24	$3+6$	0.26	0.75	$\approx$212

4.2 Gran Canaria

We now approximate the orography of Gran Canaria (Canary Islands, Spain). We consider a rectangular area of 60×70 km, and fix the upper bound of the domain to 11 km. As in the previous case, the topography is defined by a rasted data, with 25 m as discretization step. In Fig. 8 we show the corresponding meccano and its initial triangulation, respectively.

We generate several meshes of Gran Canaria, for several values of the discretization parameter ε. Table 1 reports their main features. The volumetric parameterization that induces a tetrahedral mesh is used to relocate the inner nodes of the following finer mesh. Note that neither the number of optimization steps nor the minimum quality depend drastically on ε. The CPU time includes surface parameterization, approximation, and smoothing.

In Fig. 9 we show the surface of the final mesh for $\varepsilon = 25$ m and some details of the discretization of Gran Canaria topography.

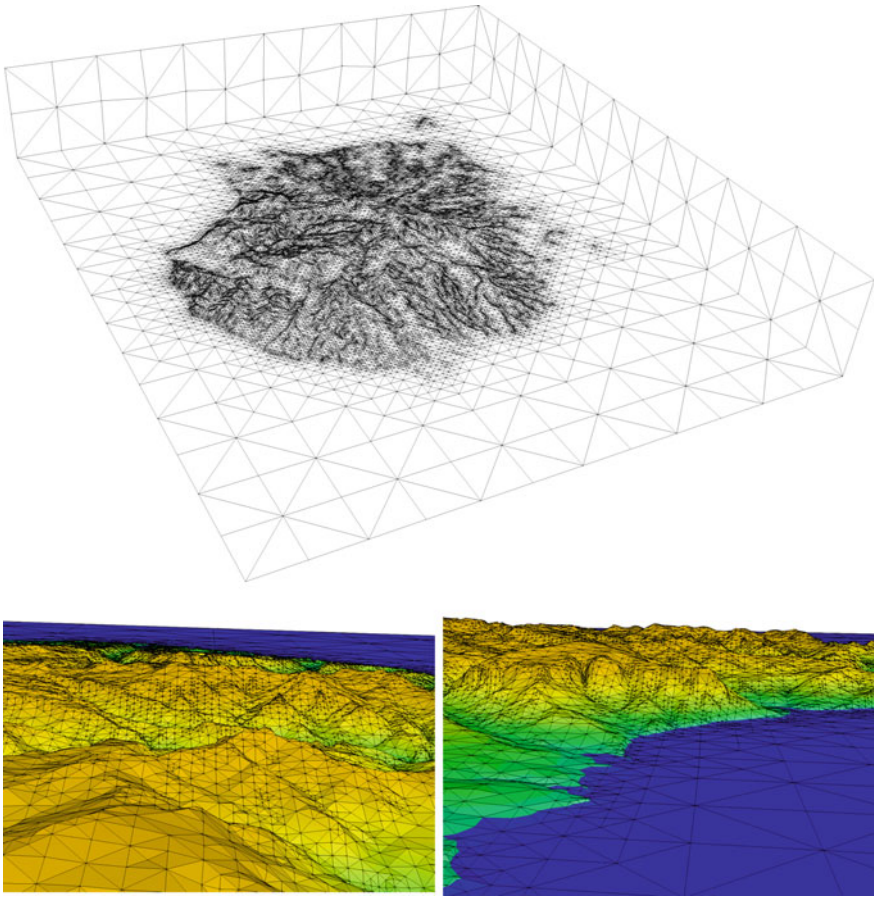


Fig. 9 Surface of the final mesh of Gran Canaria Island, for $\varepsilon = 50\,\text{m.}$, and several details

4.3 Industrial Chimney

In this last example, we show that meccano algorithm allows to include the geometry of building or stacks. We now consider an area located at north east of Gran Canaria, Spain (see Fig. 10a), where we introduce a chimney of a power plant. The dimension of the domain is $15 \times 25 \times 10$ km, and the chimney has 5 m. of radius and a height of 100 m. In order to get a good approximation of the chimney we define the tolerance ϵ as:

$$\varepsilon(\mathbf{x}) = \begin{cases} 0.5\,m. & \text{if the distance from } \mathbf{x} \text{ to the chimney is less than 100 m.} \\ 5\,m. & \text{otherwise} \end{cases}$$

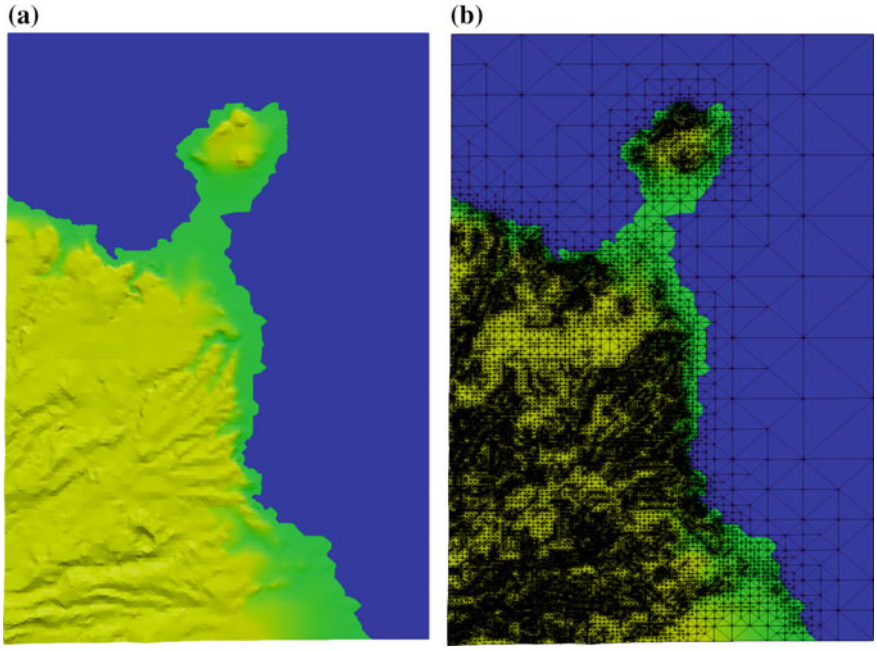


Fig. 10 **a** Map of the Gran Canaria's north east, and 2D mesh of the surface

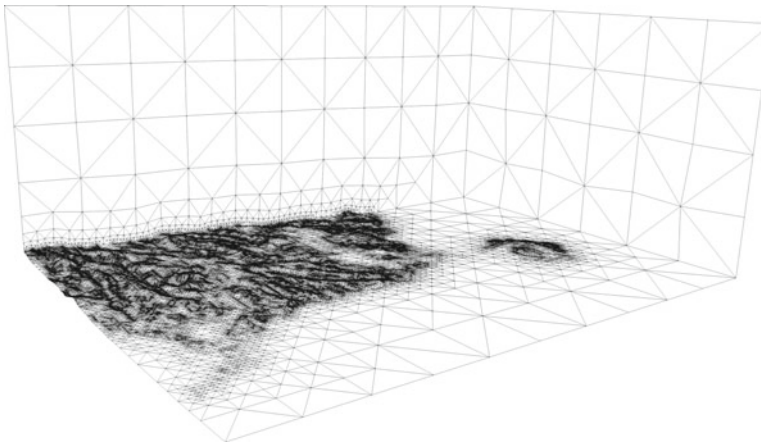


Fig. 11 Surface of the final mesh of Gran Canaria's north east

The adaptive procedure performs 96 refinement step (bisection) to reach the prescribed tolerance. The resulting mesh has 695678 tetrahedra and 159166 nodes and its surface is shown in Fig. 11.

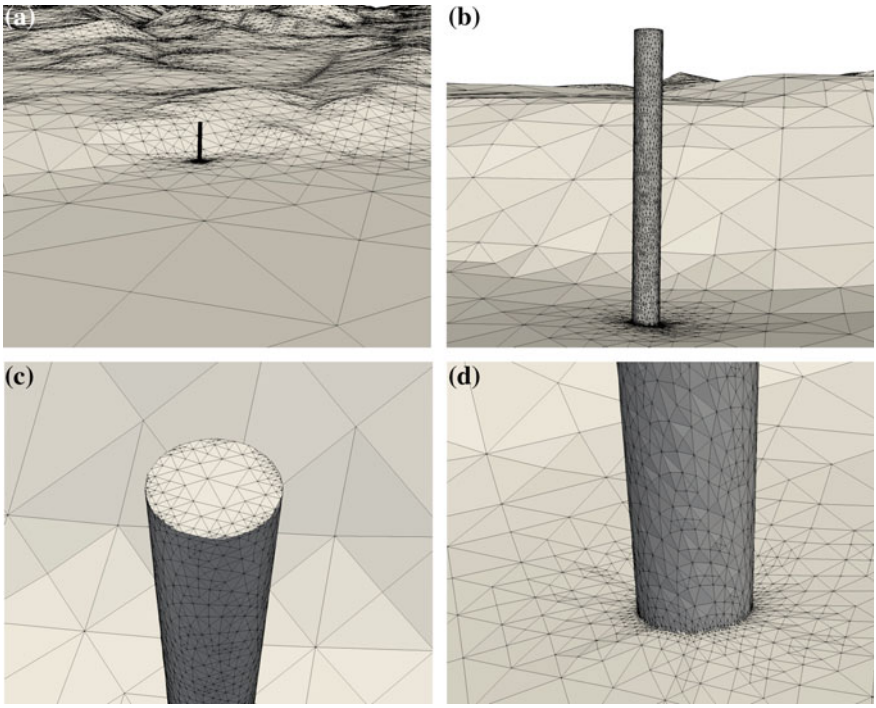


Fig. 12 Details of the chimney discretization

In Fig. 12 some details of the chimney discretization are presented. Note that the features of the structure are property approximated by our algorithm.

After relocating the inner nodes, using the volumetric parameterization produced by a coarse ε, the optimization process only required 16 steps to untangle and smooth the mesh. The average quality measure of the final mesh is $\overline{q}_{\kappa} = 0.70$. We present in Fig. 13a the 'cumulative frequency polygon' of the generated mesh, for a given value of $x \in (0, 1)$ the line represents the percentage of elements that have a quality $\leq x$ and in Fig. 13b the quality histogram of the final mesh: where the height of each bar is the relative number of tetrahedra with quality associate to the bar. We note that the 90% of tetrahedra has a quality bigger than 0.7. The CPU time to generate the final mesh is about 600 s. We remark that this time can be drastically reduced by using a parallel algorithm of the simultaneous untangling and smoothing procedure [2].

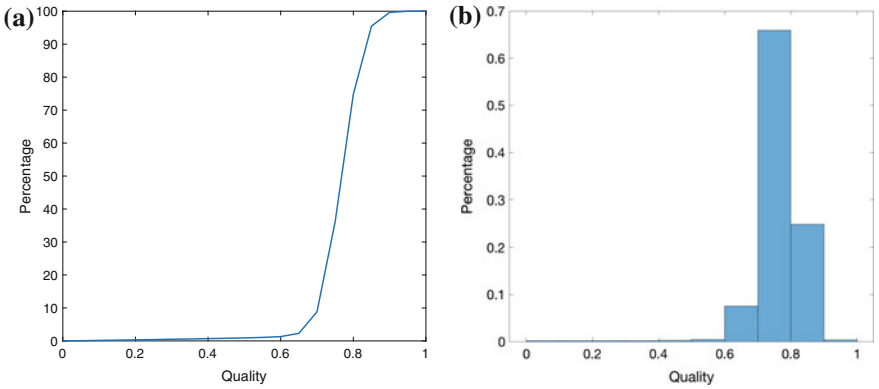


Fig. 13 **a** Cumulative frequency polygon of the final mesh; for a given quality value $x \in (0, 1)$ the line represents the percentage of elements that have a quality $\leq x$. **b** Quality histogram of the final mesh. The height of each bar is the relative number of tetrahedra with quality associate to the bar

5 Conclusions

We have established the main aspects to generate a tetrahedral mesh able to adapt to the topography of a rectangular area with a minimum intervention of users. Our three-dimensional domain is limited on its lower part by the terrain and on its upper part by a horizontal plane placed at a height where the magnitudes under study may be considered steady. The lateral walls are formed by four vertical planes. The main input data consist on a digital elevation map of the terrain and its desired approximation. The adaptive mesh is efficiently built by using the Meccano Method, that was previously introduced by the authors. The node distribution is obtained automatically in the domain under study, able to get the irregular topographic information of the terrain, and with a decreasing density as altitude increases in relation to the terrain. In addition, the meshes can be generated on a laptop with a low CPU time. Our procedure constructs high-quality meshes that are appropriated to simulate environmental problems over complex terrain, in particular, those analyzed in this book: Wind simulation and solar radiation. In this last case, a surface adaptive triangulation of the terrain is only necessary in the numerical model. The mesh generation procedure can also consider the inclusion of stacks for air pollution simulation.

References

1. Bazaraa MS, Sherali HD, Shetty CM (1993) Nonlinear programing: theory and algorithms. Wiley, New York
2. Benítez D, Rodríguez E, Escobar JM, Montenegro R (2014) Performance evaluation of a parallel algorithm for simultaneous untangling and smoothing of tetrahedral meshes. In: Proceedings of the 22nd international meshing roundtable. Springer Science Business Media, pp 579–598. https://doi.org/10.1007/978-3-319-02335-9_32

3. Carey GF (1997) Computational grids: generation adaptation and solution strategies. Taylor, Washington
4. Carey GF (2006) A perspective on adaptive modeling and meshing (AM&M). Comput Methods Appl Mech Eng 195(4–6):214–235. https://doi.org/10.1016/j.cma.2004.11.027. 1st ADMOS Conference 2003, Goteborg, SWEDEN, Sep, 2003
5. Cascón JM, Montenegro R, Escobar JM, Rodríguez E, Montero G (2007) A new meccano technique for adaptive 3-d triangulations. In: Proceedings of the 16th international meshing roundtable. Springer, Berlin, Germany, pp 103–120. https://doi.org/10.1007/978-3-540-75103-8_6
6. Cascón JM, Rodríguez E, Escobar JM, Montenegro R (2015) Comparison of the meccano method with standard mesh generation techniques. Eng Comput 31(1):161–174. https://doi.org/10.1007/s00366-013-0338-6
7. Escobar JM, Montero G, Montenegro R, Rodríguez E (2006) An algebraic method for smoothing surface triangulations on a local parametric space. Int J Numer Meth Eng 66(4):740–760. https://doi.org/10.1002/nme.1584
8. Escobar JM, Rodríguez E, Montenegro R, Montero G, González-Yuste JM (2003) Simultaneous untangling and smoothing of tetrahedral meshes. Comput Methods Appl Mech Eng 192(25):2775–2787. https://doi.org/10.1016/S0045-7825(03)00299-8
9. Escobar JM, Rodríguez E, Montenegro R, Montero G, González-Yuste JM (2010) SUS code: simultaneous mesh untangling and smoothing code. http://www.dca.iusiani.ulpgc.es/SUScode
10. Floater MS (1997) Parametrization and smooth approximation of surface triangulations. Comput Aided Geom Des 14(3):231–250. https://doi.org/10.1016/S0167-8396(96)00031-3
11. Floater MS (2003) Mean value coordinates. Comput Aided Geom Des 20(1):19–27. https://doi.org/10.1016/S0167-8396(03)00002-5
12. Frey PJ, George PL (2000) Mesh generation. Hermes Science Publishing, Oxford
13. George PL, Borouchaki H (1998) Delaunay triangulation and meshing: application to finite elements. Editions Hermes, Paris
14. González-Yuste JM, Montenegro R, Escobar JM, Montero G, Rodríguez E (2004) Local refinement of 3-d triangulations using object-oriented methods. Adv Eng Softw 35(10):693–702. https://doi.org/10.1016/j.advengsoft.2003.07.003. Engineering Computational Technology
15. Knupp PM (2001) Algebraic mesh quality metrics. SIAM J Sci Comput 23(1):193–218. https://doi.org/10.1137/s1064827500371499
16. Kossaczký I (1994) A recursive approach to local mesh refinement in two and three dimensions. J Comput Appl Math 55(3):275–288. https://doi.org/10.1016/0377-0427(94)90034-5
17. Löhner R, Baum JD (1992) Adaptive h-refinement on 3-D unstructured grids for transient problems. Int J Numer Meth Fluids 14:1407–1419. https://doi.org/10.1002/fld.1650141204
18. Montenegro R, Cascón JM, Escobar JM, Rodríguez E, Montero G (2009) An automatic strategy for adaptive tetrahedral mesh generation. Appl Numer Math 59(9):2203–2217. https://doi.org/10.1016/j.apnum.2008.12.010
19. Montenegro R, Cascón JM, Escobar JM, Rodríguez E, Montero G (2010) The meccano method for simultaneous volume parametrization and mesh generation of complex solids. IOP Conf Ser Mater Sci Eng 10(1):012–018. https://doi.org/10.1088/1757-899X/10/1/012018
20. Montenegro R, Cascón JM, Rodríguez E, Escobar JM, Montero G (2010) The meccano method for automatic three-dimensional triangulation and volume parametrization of complex solids. In: Topping B, Adam J, Pallarés F, Bru R, Romero M (eds) Developments and applications in engineering computational technology, seventh edn., Chap. 2. Saxe-Coburg Publications, Stirlingshire, pp 19–48. https://doi.org/10.4203/csets.26.2
21. Oliver A, Montero G, Montenegro R, Rodríguez E, Escobar JM, Pérez-Foguet A (2013) Adaptive finite element simulation of stack pollutant emissions over complex terrains. Energy 49:47–60. https://doi.org/10.1016/j.energy.2012.10.051

22. Prieto D, Asensio MI, Ferragut L, Cascón JM, Morillo A (2017) A gis-based fire spread simulator integrating a simplified physical wildland fire model and a wind field model. Int J Geogr Inf Sci 31(11):2142–2163. https://doi.org/10.1080/13658816.2017.1334889
23. Rivara M (1987) A grid generator bases on 4-triangles conforming mesh-refinement algorithms. Int J Numer Meth Eng 24(7):1343–1354. https://doi.org/10.1002/nme.1620240710
24. Thompson JF, Soni B, Weatherill N (1999) Handbook of grid generation. CRC Press, London

Part II
Wind Field Diagnostic and Forecasting

Wind fields knowledge is required for renewable energy modeling and forecasting. The electrical power generation using wind energy is affected by the natural resource variability. This fluctuations could be estimated and forecasted using different deterministic and probabilistic models. The knowledge of wind fields behaviour may lead grid operators to increase renewable energy integration.

In the following three Chapters this book will focus in different diagnostic and forecasting models. Chapter 4 describes Wind3D governing equations, the solver of the corresponding linear system and a numerical experiment over complex terrain. WIND3D is a mass-consistent diagnostic model with an updated vertical wind profile and atmospheric parameterization. While Chaps. 5 and 6 focus in wind fields deterministic and probabilistic forecasting. First one describes Regional Numerical Weather Prediction models and offers a snapshot of how HARMONIE-AROME model deals with NWP issues to derive a formulation for the 10 m wind. Last chapter describes the need for ensemble forecasting, the different techniques used to generate the different initial conditions, and the operational ensemble models that are used nowadays in meteorological agencies.

Wind Field Diagnostic Model

Eduardo Rodríguez, Gustavo Montero and Albert Oliver

Abstract This chapter describes Wind3D, a mass-consistent diagnostic model with an updated vertical wind profile and atmospheric parameterization. First, a description of Wind3D is provided, along with their governing equations. Next, the finite element formulation of the model and the description of the solver of the corresponding linear system are presented. The model requires an initial wind field, interpolated from data obtained in a few points of the domain. It is constructed using a logarithmic wind profile that considers the effect of both stable boundary layer (SBL) and the convective boundary layer (CBL). One important aspect of mass-consistent models is that they are quite sensitive to the values of some of their parameters. To deal with this problem, a strategy for parameter estimation based on a memetic algorithm is presented. Finally, a numerical experiment over complex terrain is presented along with some concluding remarks.

1 Mass-Consistent Model

Diagnostic models apply conservation of mass, momentum, and energy singularly or fully, considering the terrain effects on an initial flow field. Although these models are used to obtain wind fields at a given time, the results usually represent winds of a time-averaged period. Diagnostic models are limited in comparison with prognostic models because they do not take into account the transient and thermal effects so

E. Rodríguez (✉) · G. Montero · A. Oliver
University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain
e-mail: eduardo.rodriguez@ulpgc.es

G. Montero
e-mail: gustavo.montero@ulpgc.es

A. Oliver
e-mail: albert.oliver@ulpgc.es

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_4

they cannot simulate the evolution of the boundary layer; however, the computational requirements of the former are much lower than the latter.

Diagnostic models can be classified into three different categories according to the conservation laws applied. The first category comprises the diagnostic models that are based only on the conservation of mass; see, e.g., [1–3]. These models obtain a divergence-free flow that minimizes the differences with an initial known wind field. Mass-consistent models have been applied to the dynamical-downscaling of NWP models for local and regional scale wind forecasting, e.g., the WindNinja model [4]. The second category considers a linearized momentum equation [5, 6]. Nonlinear momentum effects in steep terrain are not represented by these models [7]. Compared to the mass-consistent models, computational cost is comparable providing similar results [8, 9]. Nevertheless, mass-consistent models are better suited than linearized models for some atmospheric dispersion problems where a fast response is required [10]. The third type of diagnostic model applies conservation of both mass and momentum to some form of turbulence closure [7, 11–13], and even conservation of energy [14]. The RANS RNG $k-\varepsilon$ turbulence model has handled nonlinear flow effects better than mass-consistent models [7] but it is computationally more expensive.

Wind3D uses the logarithmic wind profile to construct the initial wind field. Under this profile, the value of z_0 is the height above ground level where the wind speed follows the logarithmic law and below that height the wind speed is considered zero. The value of d preserves the logarithmic law above tall obstacles [15]. Both z_0 and d determine the effect of the land cover to the near-surface airflow [16]. Therefore, the values of these parameters are directly related to the vegetation and topographical characteristics of the terrain, which can be defined by the land coverage of the terrain, as explained in Chap. 2.

1.1 Governing Equation

We consider a mass-consistent model [2, 3, 17, 18] to compute a wind field u in a domain Ω with a boundary $\Gamma = \Gamma_a \cup \Gamma_b$, which satisfies the mass continuity equation in Ω, for an incompressible flow, and the impermeability condition on the terrain Γ_a:

$$\nabla \cdot \mathbf{u} = 0 \ \text{ in } \ \Omega \tag{1}$$

$$\mathbf{n} \cdot \mathbf{u} = 0 \ \text{ in } \ \Gamma_a \tag{2}$$

where n is the outward-pointing normal unit vector and Γ_b the free boundary. The model formulates a least-squares problem in the domain Ω to find a wind field $\mathbf{u}(\widetilde{u}, \widetilde{v}, \widetilde{w})$ such that it is adjusted as much as possible to an interpolated wind field $\mathbf{v}_0(u_0, v_0, w_0)$. The adjusting functional for a field $\mathbf{u}(\widetilde{u}, \widetilde{v}, \widetilde{w})$ is defined as

$$E(\widetilde{u},\widetilde{v},\widetilde{w}) = \int_{\Omega} [\alpha_1^2\Big((\widetilde{u}-u_0)^2 + (\widetilde{v}-v_0)^2\Big) + \alpha_2^2(\widetilde{w}-w_0)^2]\ d\Omega \tag{3}$$

being α_1 and α_2 the Gauss Precision moduli, considered equal for the horizontal direction. Mass-consistent models are very sensitive to the values chosen for α_1 and α_2, so special care must be taken in its selection. Dividing Eq. 3 by α_2^2 leads to the so-called stability parameter α,

$$\alpha = \frac{\alpha_1}{\alpha_2} \tag{4}$$

Note that coefficients α_1 and α_2 are the adjusting weights for the horizontal and vertical components of wind velocity. For $\alpha \gg 1$, vertical wind component has more weight, so wind tends to pass over terrain barriers; with $\alpha \ll 1$ wind tends to surround such barriers. In particular, there is pure vertical adjustment for $\alpha \to \infty$, while $\alpha \to 0$ means pure horizontal adjustment.

In order to find the wind field $\mathbf{v}(u, v, w)$ the following problem must be solved: "*Find* $\mathbf{v} \in K$ *such that,*

$$E(\mathbf{v}) = \min_{\mathbf{u}\in K} E(\mathbf{u}), \quad K = \Big\{\mathbf{u};\ \nabla\cdot\mathbf{u} = 0,\ \mathbf{n}\cdot\mathbf{u}|_{\Gamma_b} = 0\Big\}" \tag{5}$$

This problem is equivalent to finding the saddle point $(\mathbf{v}, \phi)$ of Lagrangian [19],

$$L(\mathbf{u},\lambda) = E(\mathbf{u}) + \int_{\Omega} \lambda\nabla\cdot\mathbf{u}\, d\Omega \tag{6}$$

The Lagrange multiplier technique can be used to obtain the saddle point of Eq. (6), $L(\mathbf{v},\lambda) \le L(\mathbf{v},\phi) \le L(\mathbf{u},\phi)$, such that the solution field $\mathbf{v}$ can be obtained from Euler–Lagrange equations,

$$\mathbf{v} = \mathbf{v}_0 + T\nabla\phi \tag{7}$$

being ϕ Lagrange multiplier and $T = (T_h, T_h, T_v)$ the diagonal transmissivity tensor

$$T_h = \frac{1}{2\alpha_1^2}\ ,\ \ T_v = \frac{1}{2\alpha_2^2}\ \text{ y }\ \frac{T_v}{T_h} = \alpha^2 \tag{8}$$

If α_1 and α_2 are considered constant in the whole domain, variational formulation leads to an elliptic equation defined in ϕ. Substituting Eq. (7) in (1) results in

$$-\nabla\cdot(T\nabla\phi) = \nabla\cdot\mathbf{v}_0 \tag{9}$$

which can be completed with null Dirichlet condition in the permeable boundaries of the domain (vertical boundaries)

$$\phi = 0 \text{ in } \quad \Gamma_a \tag{10}$$

and a Neumann condition in the non-permeable boundaries (terrain and upper boundary)

$$\mathbf{n} \cdot T \nabla \phi = -\mathbf{n} \cdot \mathbf{v}_0 \text{ in } \quad \Gamma_b \tag{11}$$

Taking into account that the initial wind field $\mathbf{v}_0$ is horizontal in the upper boundary, condition (11) becomes

$$\mathbf{n} \cdot T \nabla \phi = 0 \tag{12}$$

Considering T_h and T_v as constants, Eq. (9) becomes

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \alpha^2 \frac{\partial^2 \phi}{\partial z^2} = -\frac{1}{T_h}\left(\frac{\partial u_0}{\partial x} + \frac{\partial v_0}{\partial y} + \frac{\partial w_0}{\partial z}\right) \tag{13}$$

1.2 Finite Element Formulation

The classic formulation given in Eqs. (9), (10) and (11) is solved using the finite element method (FEM) and tetrahedral meshes (see Chap. 3).

Note that in the variational formulation of the problem, integrals in the boundary with Neumann condition are canceled using Eq. (11), while those corresponding to Dirichlet conditions are eliminated canceling out the corresponding test function.

This leads to a set of elemental matrices of dimension 4×4 associated to the element Ω_e, being $\hat{\psi}_i$ the shape function corresponding to the i-th node, $i = 1, 2, 3, 4$, defined in the reference element $\hat{\Omega}_e$ and $|\mathbf{J}|$ the Jacobian of the transformation of Ω_e into $\hat{\Omega}_e$,

$$\begin{aligned}\left\{\mathbf{A}^e\right\}_{ij} = \int_{\hat{\Omega}_e} \Bigg\{ &\left(\frac{\partial \hat{\psi}_i}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial \hat{\psi}_i}{\partial \eta}\frac{\partial \eta}{\partial x} + \frac{\partial \hat{\psi}_i}{\partial \varphi}\frac{\partial \varphi}{\partial x}\right)\left(\frac{\partial \hat{\psi}_j}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial \hat{\psi}_j}{\partial \eta}\frac{\partial \eta}{\partial x} + \frac{\partial \hat{\psi}_j}{\partial \varphi}\frac{\partial \varphi}{\partial x}\right) \\ &+ \left(\frac{\partial \hat{\psi}_i}{\partial \xi}\frac{\partial \xi}{\partial y} + \frac{\partial \hat{\psi}_i}{\partial \eta}\frac{\partial \eta}{\partial y} + \frac{\partial \hat{\psi}_i}{\partial \varphi}\frac{\partial \varphi}{\partial y}\right)\left(\frac{\partial \hat{\psi}_j}{\partial \xi}\frac{\partial \xi}{\partial y} + \frac{\partial \hat{\psi}_j}{\partial \eta}\frac{\partial \eta}{\partial y} + \frac{\partial \hat{\psi}_j}{\partial \varphi}\frac{\partial \varphi}{\partial y}\right) + \\ &+ \frac{T_v}{T_h}\left(\frac{\partial \hat{\psi}_i}{\partial \xi}\frac{\partial \xi}{\partial z} + \frac{\partial \hat{\psi}_i}{\partial \eta}\frac{\partial \eta}{\partial z} + \frac{\partial \hat{\psi}_i}{\partial \varphi}\frac{\partial \varphi}{\partial z}\right)\left(\frac{\partial \hat{\psi}_j}{\partial \xi}\frac{\partial \xi}{\partial z} + \frac{\partial \hat{\psi}_j}{\partial \eta}\frac{\partial \eta}{\partial z} + \frac{\partial \hat{\psi}_j}{\partial \varphi}\frac{\partial \varphi}{\partial z}\right)\Bigg\} \cdot |\mathbf{J}| \; d\xi \, d\eta \, d\varphi\end{aligned} \tag{14}$$

and elemental vectors of dimension 4×1,

$$\left\{\mathbf{b}^e\right\}_i = \int_{\hat{\Omega}_e} -\frac{1}{T_h}\left\{u_0\left(\frac{\partial\hat{\psi}_i}{\partial\xi}\frac{\partial\xi}{\partial x}+\frac{\partial\hat{\psi}_i}{\partial\eta}\frac{\partial\eta}{\partial x}+\frac{\partial\hat{\psi}_i}{\partial\varphi}\frac{\partial\varphi}{\partial x}\right)+\right.$$

$$+\,v_0\left(\frac{\partial\hat{\psi}_i}{\partial\xi}\frac{\partial\xi}{\partial y}+\frac{\partial\hat{\psi}_i}{\partial\eta}\frac{\partial\eta}{\partial y}+\frac{\partial\hat{\psi}_i}{\partial\varphi}\frac{\partial\varphi}{\partial y}\right)+ \tag{15}$$

$$\left.+\,w_0\left(\frac{\partial\hat{\psi}_i}{\partial\xi}\frac{\partial\xi}{\partial z}+\frac{\partial\hat{\psi}_i}{\partial\eta}\frac{\partial\eta}{\partial z}+\frac{\partial\hat{\psi}_i}{\partial\varphi}\frac{\partial\varphi}{\partial z}\right)\right\}\cdot|\mathbf{J}|\;d\xi\,d\eta\,d\varphi$$

1.3 Linear System Resolution

The application of finite element method to these problems leads to the resolution of large and symmetric linear systems of equations with a sparse matrix of coefficients (stiffness matrix):

$$\mathbf{Ax} = \mathbf{b} \tag{16}$$

Using iterative solvers is an appropriate strategy to solve such linear systems. In particular, the conjugated gradient method [20] is the most efficient Krylov's subspace method for solving symmetric linear systems.

In the case of a sparse matrix, major memory requirement reductions can be achieved by storing only the nonzero entries in the computer memory. In particular, the compressed storage row technique uses three one-dimensional arrays to represent the stiffness matrix that respectively contain nonzero values, the extents of rows and column indices. An storage of order $3 \times n$ (being n the matrix dimension) is needed, in contrast to n^2 needed for the whole representation (including zero coefficients). The trade-off is that accessing the individual elements becomes more complex and additional structures are needed to be able to recover the original matrix unambiguously.

The solution, ϕ, of the linear system is used to obtain the wind field with Eq. 7.

The rate of convergence of methods based on Krylov subspaces, and conjugated gradient in particular, can be improved with the use of preconditioning techniques. In general, they consist of replacing the original system of equations (16) by another one with identical solution, in such a way that the condition of the matrix of the new system is lower than that of $\mathbf{A}$. In general, a preconditioning matrix $\mathbf{M}^{-1}$ is considered, being $\mathbf{M}$ an approximation of $\mathbf{A}$,

$$\mathbf{M}^{-1}\mathbf{Ax} = \mathbf{M}^{-1}\mathbf{b} \tag{17}$$

such that, $\kappa\left(\mathbf{M^{-1}A}\right) < \kappa\left(\mathbf{A}\right)$.

The lowest value corresponds to ideal case $\mathbf{M} = \mathbf{A}$, $\kappa\left(\mathbf{A}^{-1}\mathbf{A}\right) = 1$, with the system converging in one iteration, but the computational cost of obtaining $\mathbf{A}^{-1}$ would

be equivalent to solve the system by means of a direct method. The objective is to calculate a matrix $\mathbf{M}$ as close to $\mathbf{A}$ as possible with low computational cost.

Matrix $\mathbf{M}$ should also be easily invertible in order to have a reasonable computational cost in $\mathbf{M}^{-1}$–vector products in the preconditioned algorithms.

The preconditioning may be carried out in three different ways,

$$\begin{array}{ll} \mathbf{M}^{-1}\mathbf{A}\mathbf{x} = \mathbf{M}^{-1}\mathbf{b} & \text{(Left preconditioning)} \\ \mathbf{A}\mathbf{M}^{-1}\mathbf{M}\mathbf{x} = \mathbf{b} & \text{(Right preconditioning)} \\ \mathbf{M}_1^{-1}\mathbf{A}\mathbf{M}_2^{-1}\mathbf{M}_2\mathbf{x} = \mathbf{M}_1^{-1}\mathbf{b} & \text{(Both sides preconditioning)} \end{array} \tag{18}$$

if $\mathbf{M}$ can be factorized as $\mathbf{M} = \mathbf{M}_1\mathbf{M}_2$. Preconditioned Conjugated Gradient method is shown in Algorithm 1.

A number of preconditions have been developed and widely used in several application fields; see, e.g., [21]. Despite, we have achieved good results in our simulations with Wind3D using the Jacobi precondition, i.e., $\mathbf{M} = diag(\mathbf{A})$, in [22] can be found a specific precondition for this wind problem.

Algorithm 1 Preconditioned Conjugate Gradient (PCG).

1: Initial approximation $\mathbf{x}_0$. $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
2: Solve $\mathbf{M}\mathbf{z}_0 = \mathbf{r}_0$, $\mathbf{p}_0 = \mathbf{z}_0$;
3: **while** $\| \mathbf{r}_j \| / \| \mathbf{r}_0 \| \geq \varepsilon \quad (j = 0, 1, 2, 3, \ldots)$ **do**
4: $\quad \alpha_j = \dfrac{\langle \mathbf{r}_j, \mathbf{z}_j \rangle}{\langle \mathbf{A}\mathbf{p}_j, \mathbf{p}_j \rangle}$;
5: $\quad \mathbf{x}_{j+1} = \mathbf{x}_j + \alpha_j \mathbf{p}_j$;
6: $\quad \mathbf{r}_{j+1} = \mathbf{r}_j - \alpha_j \mathbf{A}\mathbf{p}_j$;
7: $\quad$ Solve $\mathbf{M}\mathbf{z}_{j+1} = \mathbf{r}_{j+1}$;
8: $\quad \beta_j = \dfrac{\langle \mathbf{r}_{j+1}, \mathbf{z}_{j+1} \rangle}{\langle \mathbf{r}_j, \mathbf{z}_j \rangle}$;
9: $\quad \mathbf{p}_{j+1} = \mathbf{z}_{j+1} + \beta_j \mathbf{p}_j$;
10: **end while**

1.4 Construction of the Interpolated Wind Field

The first step of Wind3D is to create an initial wind field using wind data available only in a few locations on the domain. Data is typically obtained from wind measurement stations or from large-scale numerical weather models with coarse grids. With this data, a suitable interpolation is performed in order to construct a wind field on the whole domain, using a logarithmic vertical profile.

At this stage, we want to interpolate the available wind data at any point located at a height z_m over $z_t + d$, where z_t is the terrain surface height. If we have a set of dispersed data, a simple technique for this interpolation is formulated as a weighted sum of the inverse-square law and the height difference interpolation [2]:

$$u_0(z_m) = \xi \frac{\sum_{i=1}^{n_h} \frac{u_i^h}{\delta_i^2}}{\sum_{i=1}^{n_h} \frac{1}{\delta_i^2}} + (1-\xi) \frac{\sum_{i=1}^{n_h} \frac{u_i^h}{|\Delta h_i|}}{\sum_{i=1}^{n_h} \frac{1}{|\Delta h_i|}}, \tag{19}$$

where the value of u_i^h is the wind velocity at the point i; n_h is the number of available points; δ_i is the horizontal distance between point i and the point of interest; $|\Delta h_i|$ is their height difference, and ξ is a weighting parameter ($0 \leq \xi \leq 1$) that determines to what degree the focus is put on the inverse-square law or the height difference interpolation.

When a grid of wind data is available, e.g., downscaling HARMONIE-AROME or ECMWF wind results, it is preferable to interpolate the wind field with a simple bilinear Lagrange interpolation in the cell containing the studied point.

The next step is the construction of the vertical wind profile. We have implemented the log wind profile proposed in [23] for stable/neutral conditions (see Chap. 2):

$$u = \frac{u_*}{k}\left(\ln\frac{\zeta}{\zeta_0} + b_1(\zeta-\zeta_0) + b_2(\zeta-\zeta_0)^2 + b_3(\zeta-\zeta_0)^3\right), \tag{20a}$$

$$v = -\frac{u_*}{k}\delta\left(-(\zeta-\zeta_0)\ln(\zeta-\zeta_0) + a_1(\zeta-\zeta_0) + a_2(\zeta-\zeta_0)^2 + a_3(\zeta-\zeta_0)^3\right), \tag{20b}$$

for $z > d + z_0$, where u and v are the components of the horizontal wind velocity along the x and y axis of a right-hand Cartesian coordinate system with the x-axis along the surface stress; $\zeta = (z-d)/h$ and $\zeta_0 = z_0/h$ are dimensionless heights; $k \approx 0.4$ is the von Kármán constant; $\delta = fh/(ku_*) = \gamma/k$ is the dimensionless rotation rate parameter; $a_1 = 4/\delta^2 + \overline{\Pi}, a_2 = -\frac{3}{2}\overline{\Pi}, a_3 = \frac{1}{3}\left(1 - 4/\delta^2 + 2\overline{\Pi}\right), b_1 = \Pi - 3$, $b_2 = -\frac{3}{2}\Pi$, $b_3 = \frac{2}{3}(\Pi+1)$, with $\Pi = C_R\delta^2 + C_L\frac{kh}{L} + C_N\frac{Nh}{u_*}$ and $\overline{\Pi} = \overline{C}_R\delta^2 + \overline{C}_L\frac{kh}{L} + \overline{C}_N\frac{Nh}{u_*}$; and $L = -u_*^3/(kB_s)$ is the Monin–Obukhov length. This model uses the values of estimates of the dimensionless constants obtained by [23] on the basis of empirical and numerical (LES) data: $C_R = 7, C_L = 4.5, C_N = 0.4, \overline{C}_R = 0, \overline{C}_L = -7, \overline{C}_N = -1$. In addition, the original expression in [23] has been slightly modified to verify $u(z) = 0$ and $v(z) = 0$ at $z = d + z_0$. The vertical component w of the wind velocity is assumed to be zero. From (20a), the surface friction velocity may be computed using the horizontal wind velocity interpolated at $z = z_m$, $u_0(z_m)$:

$$\ln\frac{z_m - d}{z_0}u_*^3 - \left[ku_0(z_m) - \left(C_R\frac{\gamma}{k^2} + \frac{C_N N_{2h-h}}{f} - \frac{3}{\gamma}\right) f\,(z_m - (d+z_0))\right]u_*^2 - C_L k^2 B_s(z_m - (d+z_0)) = 0, \tag{21}$$

where the squared and cubic terms of the wind profile were neglected at $z = z_m$.

The estimation of the CBL height is provided by the mesoscale model estimations (see Chap. 2). If the ratio between the mechanical velocity scale $V_* = (2Uu_*)^{1/3}$ and the convective velocity scale $W_* = (B_s h)^{1/3}$ is negligible, i.e., $V_*/W_* \ll 1$, we have

a Purely Convective Layer (PCL). Otherwise, it is a Mechanically Convective Layer (MCL) [24]. The log wind profile in the CBL was given in [25] as follows:

$$|u| = \begin{cases} \dfrac{u_*}{k}\ln\dfrac{z-d}{z_0} & z_0 + d < z < \frac{\zeta_u |L|}{k} + d, \quad (22a) \\ \dfrac{u_*}{k}\left[a_u + C_u\left(\dfrac{k(z-d)}{L}\right)^{-\frac{1}{3}} + \ln\dfrac{-L}{kz_0}\right] & \frac{\zeta_u |L|}{k} + d \le z \le h, \quad (22b) \end{cases}$$

where $\zeta_u \approx 0.1$, $a_u \approx 0.7$ and $C_u \approx 1.4$ are dimensionless constants (see [1]). The angle of wind turn in the boundary layer is given by the expression:

$$\sin\alpha = \sin(\alpha_s - \alpha_{h-0}) = \frac{a_\alpha}{k}\left(\frac{hk}{|L|}\right)^{-\frac{1}{3}} \frac{u_*}{|\overline{u}|}\,\text{sign}\, f, \tag{23}$$

where α_s and α_{h-0} are the angles between the wind direction and x-axis at the terrain and $z = h$, respectively. The estimation of $a_\alpha = 3$ was proposed in [25]. The mean wind velocity $|\overline{u}|$ in the CBL is obtained from

$$|\overline{u}| = |u|_{h-0} = \frac{u_*}{k}\left[a_u + \ln\frac{-L}{kz_0}\right], \tag{24}$$

if we consider $h \gg |L|$. In practice, we assume that the wind turn angle varies linearly with height and reaches zero at the top of the CBL Finally, note that in the CBL wind profile, the calculation of the surface friction velocity u_* from the horizontal wind velocity interpolated at $z = z_m$, $u_0(z_m)$, is generally straightforward using (22a):

$$u_* = \frac{k\ |u_0(z_m)|}{\ln\dfrac{z_m - d}{z_0}}. \tag{25}$$

In the case that $z_m \ge \dfrac{\zeta_u\,|L|}{k} + d$, we have to use (22b).

2 Parameter Estimation

The results of the mass-consistent modeling have proved to be very sensitive to the values of α, ξ, z_0, and d. Thus, an accurate definition of these parameters is critical to obtain a reliable downscaling wind field. We have to estimate a value of α and ξ for the whole domain [18], and a value of z_0 and d for each land cover class. This means that the number of unknowns depends on the number of the different land covers in the region of interest.

The objective of the optimization is to find the values of the parameters such that the wind computed with the model is the most similar to a known wind at some control points. The wind values at the control points can be known from a NWP model (i.e., HARMONIE-AROME) or from measurement stations. To measure the error between the model and the known data, we use the RMSE, i.e.,

$$RMSE = \sqrt{\frac{1}{n_c}\sum_{i=1}^{n_c}(u_{xi} - u_{xi}^c)^2 + (u_{yi} - u_{yi}^c)^2 + (u_{zi} - u_{zi}^c)^2}, \tag{26}$$

where n_c is the number of control points, (u_{xi}, u_{yi}, u_{zi}) and $(u_{xi}^c, u_{yi}^c, u_{zi}^c)$ are, respectively, the wind velocity obtained with the mass-consistent model and the known wind at the ith control point. So, the parameter estimation consists of the minimization of the RMSE. Note that for each evaluation of the fitness function, the wind model has to be executed.

Evolutionary algorithms are a family of heuristic optimization methods using techniques inspired in biology to find out optimal configuration for a specific system within given constraints. For this reason, they can be used to estimate the parameters stated above. We have successfully used a classic genetic algorithm [26, 27] algorithm to estimate parameters of a simplified version of Wind3D that does not consider any land cover parameters. In the next section, we present a memetic algorithm to optimize the fitness function (26).

Note that the values of z_0 and d obtained in a particular numerical experiment are not representative values for a given land cover; they only represent the optimal solution compared to the available wind measurements for the land covers in the domain of interest. However, the general methodology can be applied to any combination of regions, databases, and downscaling wind models. So, the final aim of the proposed strategy is to improve the results of a downscaling wind model by estimating the optimum aerodynamic parameters values.

2.1 Memetic Algorithm

As stated before, an evolutionary algorithm is a suitable technique to find the optimal values of the parameters of the wind field model. A population of individual representing different values of the parameters is allowed to evolve during a number of iterations (generations). In each one, the selected individuals, according to the fitness function, are combined to create the next population. During the process, some of the individuals can go under mutation. Finally, the fittest individual is chosen as solution of the optimization problem.

The generic denomination of Memetic Algorithms (MAs) is used to encompass a broad class of metaheuristics (i.e., general purpose methods aimed to guide an underlying heuristic). In this case, we propose a memetic method composed of three tools: the differential evolution algorithm (DE) [28], a Rebirth Operator (RBO) [29],

and the L-BFGS-B algorithm [30]. DE is an evolutionary algorithm that utilizes a population composed of a fixed number n_v of D-dimensional parameter vectors $p_{i,g}$ for each generation g; $g = 1, \ldots, n_g$. The initial population, which must cover the parameter searching space, is chosen randomly. The mutation procedure modifies an existing vector by adding to itself a weighted difference between two other vectors. In the crossover step, these mutated vectors are mixed with another target vector to obtain the so-called trial vector. If the trial vector yields a lower fitness function value than the target vector, the target vector is replaced by the trial vector (selection). Each population vector has to serve as target vector at least once, so n_v competitions will take place per generation.

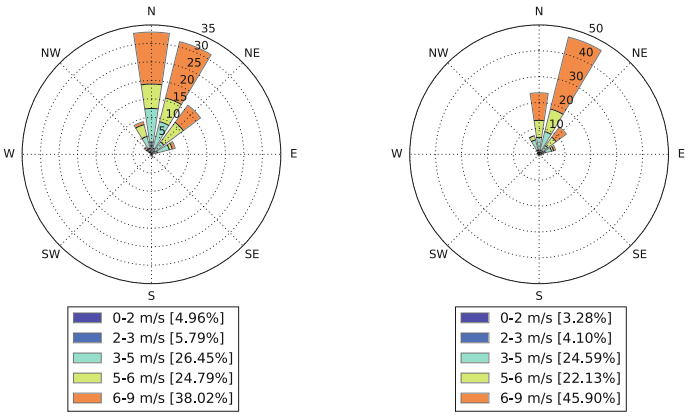
The accuracy of the results obtained using DE may be insufficient. To increase it, we have run n_e DE experiments and have performed a statistic analysis of the results obtained for each one. This analysis will allow us to reduce the search interval. Let $p^j_{i,n_g}(j = 1, \ldots, n_e;\ i = 1, \ldots, n_u)$ be the estimation of the n_u unknown parameters obtained in each of the n_e experiments. We can compute its average $\overline{p}_{i,n_g}$, and standard deviation σ_{i,n_g}. Then, the search interval can be reduced to the confidence interval of each variable, i.e., $\overline{p}_{i,n_g} \pm \frac{\sigma_{i,n_g}}{\sqrt{n_e}} T_{n_e-1,\frac{\tau}{2}}$, where $1 - \tau$ is the confidence coefficient and T, the Student's t-distribution. If one extreme of the new interval exceeds the old extreme, the latter is preserved. This allows the rebirth of a new population to restart DE. This procedure may be repeated as many times as required. Note that the n_e DE experiments can be run in parallel.

When the last generation of the last reborn population is evaluated, the best parameter vector among all the DE experiments is selected to be the starting point of the L-BFGS-B algorithm. This algorithm is a procedure for solving large nonlinear optimization problems with simple bounds. It is based on the gradient projection method and uses a limited memory BFGS matrix to approximate the Hessian of the fitness function. The results of this final minimization will be the estimated parameters.

3 Numerical Experiment

In this section, we characterize the roughness length and displacement height in the island of Gran Canaria (N27°58′ W15°36′), Spain. To this end, a map of z_0 and d is constructed. Many authors have concluded that the roughness length and height displacement values depend on the wind speed and direction, as well as on the atmospheric stability class. Therefore, the characterization is performed for different typical meteorological episodes. For this reason, we have studied the characteristics of the wind in Gran Canaria for the summer months of 2015, i.e., from June to September.

Figure 1 represents the wind roses of Gran Canaria in the selected months. For convenience, we have separated daytime and nighttime. Based on this classification in speeds and directions, we have selected eight characteristic winds that represent about 71.90% of daytime and 77.88% of nighttime.



(a) Daytime (b) Night-time

Fig. 1 Wind Rose of Gran Canaria at 10 m relating to the period from June 1 to September 30 of the year 2015

Table 1 displays the chosen episodes. The AROME/HARMONIE forecast wind speeds and directions are taken at a representative point located in the sea in the NE of the island with a height of 10 m above the sea. The Pasquill stability class has been obtained from the daytime incoming solar radiation and the nighttime cloud amount [31].

So, we are going to estimate the parameters α, ξ, as well as z_0 and d for each of the 30 basic land covers of Gran Canaria for each of the eight episodes. The land use of Gran Canaria is obtained from the SIOSE database. According to it, in the island there are 5237 different zones with its particular linear combinations of these 30 basic classes; see Fig. 2.

In addition, wind measures at 7 stations of the State Meteorological Agency of Spain (AEMET) network are available. Its UTM coordinates and heights above sea level are given in Table 2 and shown in Fig. 3.

The application is solved in a region of $76 \times 85\,\text{km}^2$ that contains the island of Gran Canaria and is limited by an upper plane at a height of 4 km. We have generated a tetrahedral mesh adapted to the terrain with local refinement around the measurement stations and the shoreline; see a detail of the terrain triangulation in Fig. 4. The mesh contains 1492804 tetrahedra and 326101 nodes. To compute the interpolated wind field, we have used the AROME/HARMONIE predictions of wind velocities at 10 m above the terrain. Finally, we have estimated the parameters using the memetic method described in Sect. 2.1. We have chosen 11 control points: the measurements from the seven stations and four AROME/HARMONIE wind forecast values located at 10 m over the sea in the corners of the mesh. Remember that for each evaluation of the fitness function (RMSE) the mass-consistent model has to be applied.

The results obtained for each wind episode are presented in Tables 3 and 4 for daytime and nighttime, respectively. In all of the eight episodes, using the

Table 1 Most frequent wind speeds and directions in the island of Gran Canaria during the summer months

Cases		Daytime			
Surface wind direction	Surface wind speed range (m/s)	AROME/HARMONIE 10 m wind speed	AROME/HARMONIE 10 m wind direction	Incoming solar radiation	Pasquill stability class
NNE	>6	7.61	32.89	530.85	D
NE	>6	9.01	37.29	583.39	D
N	>6	7.11	349.53	494.75	D
NE	5–6	5.57	40.50	665.60	C
N	3–5	4.16	4.26	797.14	B
NNE	5–6	5.88	12.76	771.82	C
NNE	3–5	3.54	12.70	586.43	C
N	5–6	5.39	350.28	695.26	C
Cases		Nighttime			
Surface wind direction	Surface wind speed range (m/s)	AROME/HARMONIE 10 m wind speed	AROME/HARMONIE 10 m wind direction	Cloud amount (oktas)	Pasquill stability class
NNE	>6	9.87	25.75	4.59	D
NE	>6	8.07	34.16	3.39	D
N	>6	6.82	355.35	7.04	D
NE	5–6	5.10	42.58	2.24	D
N	3–5	4.99	359.27	6.45	D
NNE	5–6	5.13	19.94	0.61	D
NNE	3–5	4.90	12.56	0.00	E
N	5–6	5.62	353.72	1.86	D

estimated parameters, the wind predicted by the mass-consistent model has reduced the AROME/HARMONIE error. This reduction has gone from 4.97 to 56.21% in the daytime, and from 22.49 to 58.93% in the nighttime. However, this approach is strongly dependent on the forecast values of the mesoscale model, which determines the quality of the predicted wind.

The literature points to a relationship between these parameters and the atmospheric state (wind speed, direction, and its stability). However, from the results presented here it is difficult to obtain a correlation between them. Particular conclusions may be drawn for some specific parameters. For example, z_0^{ARR} decreases when the wind direction goes from N to NE. Another outcome is that the range of variation of many of the parameters is considerably shorter than that given in Table 2.1. This is the case, e.g., for the values of z_0 corresponding to HMA, HSM, LAA, LVI, SNE and ZQM, and d in CNF, HMA, HSM, LFC, LFN, LOC, LVI, MTR, PST, RMB, SDN, SNE, ZEV, and ZQM. What is clear from these experiments is that the optimal

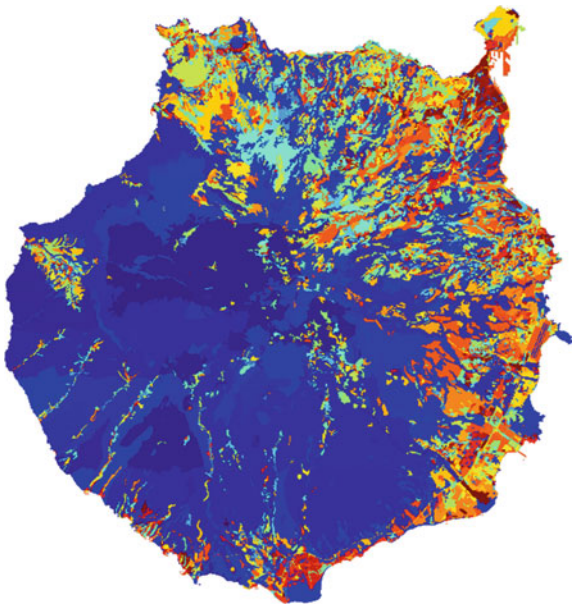


Fig. 2 SIOSE land cover polygons in the Island of Gran Canaria

Table 2 Location in UTM zone 28 N coordinates and height above the sea level of the anemometers used in the numerical application in Gran Canaria

Code	Name	$x(m)$	$y(m)$	$z(m)$
C619X	Agaete	429982	3108624	15
C629Q	Mogán, Puerto Rico	429927	3073056	20
C648N	Telde, Centro Forestal Doramas	454970	3095890	354
C649R	Telde, Melenara	462854	3095804	19
C656V	Teror	446227	3105674	693
C659M	Plaza de la Feria	458627	3109809	25
C669B	Arucas	450225	3113015	96
C689E	Maspalomas	441057	3068075	35

values of z_0 and d are different from the nominal values found in the literature. So, it is convenient to correct them for the region of the Canary Islands.

Concerning the values of α, there is no direct conclusion from the results. It seems that α is normally lower at nighttime than daytime, where it is more irregular, but further experiments should be carried out. Regarding ξ, it often reached values near 1.

Another important aspect to be considered is a sensibility study of all the involved parameters in the resulting wind field. This analysis may lead to a reduction of the number of unknowns by not considering those parameters with a negligible effect on the final result. One of the limitations of this approach is that one basic coverage may affect different regions with very different characteristics. However, in the current

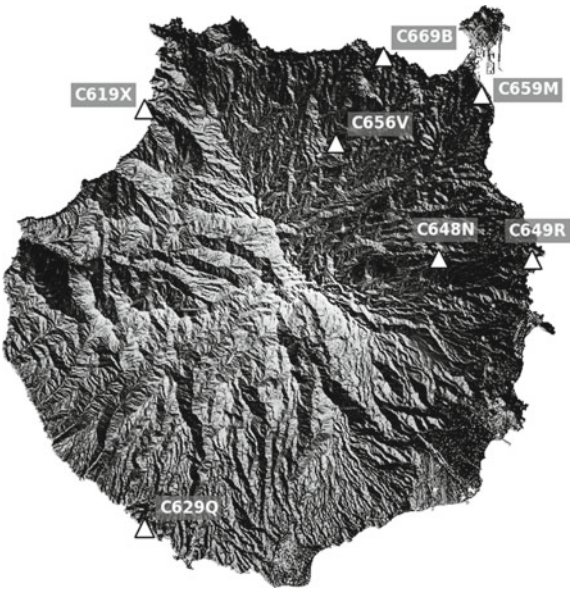


Fig. 3 Location of the wind measurement stations

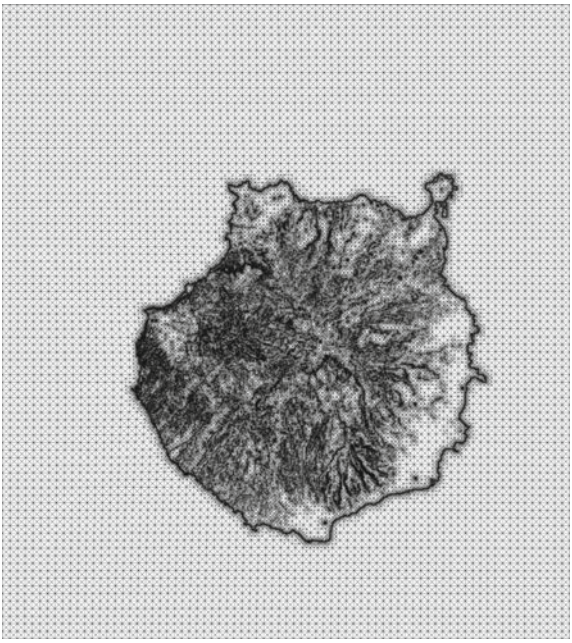


Fig. 4 Detail of the terrain of the adaptive mesh of Gran Canaria island

Table 3 Results of the daytime experiments in Gran Canaria (z_0 and d in m)

Wind direction Wind speed (ms^{-1}) Pasquill stability	N $v > 6$ D	NNE $v > 6$ D	NE $v > 6$ D	N $5 < v \leq 6$ C	NNE $5 < v \leq 6$ C	NE $5 < v \leq 6$ C	N $3 < v \leq 5$ B	NNE $3 < v \leq 5$ C
$RMSE(AROME/HARMONIE)$	2.19	2.31	2.81	1.58	1.81	1.81	1.05	1.78
$RMSE(Wind3D)$	1.25	1.21	2.11	1.30	1.08	0.79	0.81	1.70
α	9.98	0.34	0.19	0.72	0.28	0.35	0.48	9.84
ξ	0.99	0.99	0.92	0.99	0.63	0.99	0.83	0.98
z_0^{ACM}	0.15	0.07	0.09	0.07	0.14	0.08	0.05	0.11
z_0^{AEM}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{ALC}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{ALG}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{AMO}	0.02	0.00	0.00	0.02	0.01	0.00	0.02	0.02
z_0^{ARR}	0.04	0.00	0.00	0.17	0.16	0.06	0.17	0.14
z_0^{CLC}	0.00	0.02	0.02	0.01	0.04	0.06	0.02	0.04
z_0^{CNF}	1.03	0.31	0.54	1.64	0.32	0.30	1.85	1.44
z_0^{CHL}	0.69	0.01	0.19	0.21	0.01	0.07	0.07	0.28
z_0^{EDF}	3.69	0.70	0.74	0.71	3.69	3.68	0.70	1.39
z_0^{FDC}	0.75	1.35	1.09	0.32	0.54	1.15	0.99	0.20
z_0^{FDP}	1.64	2.60	2.64	1.84	1.85	2.59	1.10	0.64
z_0^{HMA}	0.04	0.05	0.05	0.01	0.04	0.04	0.10	0.07
z_0^{HSM}	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.03
z_0^{LAA}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{LFC}	0.35	0.03	0.03	0.13	0.30	0.03	0.03	0.07
z_0^{LFN}	0.92	0.89	0.17	0.11	0.94	0.90	0.07	0.09
z_0^{LOC}	0.06	0.07	0.07	0.03	0.06	0.06	0.04	0.07

(continued)

Table 3 (continued)

Wind direction Wind speed (ms^{-1}) Pasquill stability	N $v > 6$ D	NNE $v > 6$ D	NE $v > 6$ D	N $5 < v \leq 6$ C	NNE $5 < v \leq 6$ C	NE $5 < v \leq 6$ C	N $3 < v \leq 5$ B	NNE $3 < v \leq 5$ C
z_0^{LVI}	0.27	0.40	0.33	0.36	0.38	0.23	0.33	0.38
z_0^{MTR}	0.01	0.34	0.01	0.01	0.99	0.99	0.01	0.01
z_0^{OCT}	0.06	0.86	0.47	0.08	0.57	0.30	0.20	0.06
z_0^{PDA}	0.03	0.00	0.01	0.03	0.02	0.03	0.00	0.03
z_0^{PST}	0.07	0.13	0.13	0.08	0.00	0.05	0.12	0.13
z_0^{RMB}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{SDN}	0.03	0.00	0.00	0.01	0.01	0.01	0.03	0.00
z_0^{SNE}	0.02	0.00	0.01	0.01	0.03	0.00	0.00	0.02
z_0^{VAP}	0.00	0.08	0.28	0.05	0.00	0.37	0.00	0.10
z_0^{ZAU}	0.03	1.12	0.71	0.04	0.03	1.22	0.04	0.04
z_0^{ZEV}	0.14	0.00	0.00	0.00	0.15	0.13	0.09	0.08
z_0^{ZQM}	0.55	0.39	0.79	0.44	0.44	0.73	0.86	0.81
d^{ACM}	37.5	14.5	56.8	16.0	6.68	76.2	36.6	7.91
d^{ARR}	0.38	0.53	0.80	0.45	0.53	0.54	0.48	0.18
d^{CLC}	0.21	0.09	0.17	0.19	0.12	0.32	0.23	0.24
d^{CNF}	10.4	12.0	17.8	7.43	12.8	12.8	8.89	8.83
d^{CHL}	2.71	2.03	0.75	0.23	2.51	2.41	1.00	0.94
d^{EDF}	17.2	17.4	15.4	16.5	7.05	10.7	7.15	19.6

(continued)

Table 3 (continued)

Wind direction Wind speed (ms^{-1}) Pasquill stability	N $v > 6$ D	NNE $v > 6$ D	NE $v > 6$ D	N $5 < v \leq 6$ C	NNE $5 < v \leq 6$ C	NE $5 < v \leq 6$ C	N $3 < v \leq 5$ B	NNE $3 < v \leq 5$ C
d^{FDC}	18.3	8.68	14.7	16.1	6.81	19.5	12.3	16.5
d^{FDP}	14.6	24.6	3.99	24.3	29.2	30.4	9.20	10.8
d^{HMA}	0.63	0.46	0.45	0.55	0.62	0.45	0.60	0.52
d^{HSM}	0.11	0.12	0.04	0.03	0.11	0.10	0.03	0.13
d^{LFC}	3.79	0.04	1.17	1.45	3.87	3.32	1.04	3.75
d^{LFN}	1.64	3.60	1.06	1.93	2.77	3.24	1.42	2.69
d^{LOC}	0.31	0.39	0.30	0.44	0.27	0.28	0.31	0.32
d^{LVI}	0.81	1.18	0.70	1.06	0.71	0.51	1.06	1.18
d^{MTR}	6.94	5.26	5.88	2.46	1.13	4.51	1.35	6.61
d^{OCT}	2.34	4.91	3.09	2.43	12.8	11.2	3.52	2.21
d^{PDA}	0.18	0.15	0.14	0.16	0.21	0.06	0.09	0.31
d^{PST}	0.25	0.60	0.07	0.30	0.33	0.60	0.49	0.57
d^{RMB}	0.00	0.00	0.01	0.01	0.01	0.00	0.02	0.01
d^{SDN}	0.16	0.11	0.07	0.09	0.11	0.12	0.17	0.13
d^{SNE}	0.12	0.06	0.08	0.12	0.16	0.15	0.03	0.15
d^{VAP}	2.24	1.71	1.76	0.54	0.15	1.51	0.53	0.07
d^{ZAU}	13.3	10.8	12.5	13.8	11.6	4.05	11.5	13.7
d^{ZEV}	0.16	0.37	0.22	0.16	0.25	0.69	0.23	0.33
d^{ZQM}	2.69	5.35	3.79	4.23	4.41	3.95	1.42	4.58

Table 4 Results of the nighttime experiments in Gran Canaria (z_0 and d in m)

Wind direction Wind speed (ms^{-1}) Pasquill stability	N $v > 6$ D	NNE $v > 6$ D	NE $v > 6$ D	N $5 < v \leq 6$ D	NNE $5 < v \leq 6$ D	NE $5 < v \leq 6$ D	N $3 < v \leq 5$ D	NNE $3 < v \leq 5$ E
$RMSE(AROME/HARMONIE)$	3.47	3.92	3.03	2.53	2.23	1.54	2.19	2.79
$RMSE(Wind3D)$	1.47	1.61	1.34	1.28	1.73	1.10	1.19	1.43
α	0.10	0.10	0.10	0.10	9.48	0.10	0.22	0.10
ξ	0.90	0.33	0.65	0.99	0.99	0.99	0.99	0.99
z_0^{ACM}	0.10	0.11	0.06	0.10	0.17	0.13	0.11	0.07
z_0^{AEM}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{ALC}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{ALG}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{AMO}	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.01
z_0^{ARR}	0.11	0.02	0.00	0.15	0.02	0.01	0.16	0.03
z_0^{CLC}	0.04	0.04	0.01	0.02	0.05	0.03	0.05	0.03
z_0^{CNF}	0.73	0.25	0.29	0.36	1.35	0.60	0.33	0.30
z_0^{CHL}	0.13	0.00	0.36	0.57	0.70	0.71	0.19	0.01
z_0^{EDF}	1.52	0.77	0.71	2.70	2.82	0.70	3.06	0.71
z_0^{FDC}	0.67	1.25	1.24	1.24	0.23	1.15	0.53	1.24
z_0^{FDP}	2.08	2.57	2.49	0.90	0.69	2.62	0.63	2.61
z_0^{HMA}	0.03	0.10	0.05	0.07	0.11	0.09	0.11	0.12
z_0^{HSM}	0.00	0.01	0.02	0.03	0.01	0.01	0.01	0.02
z_0^{LAA}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{LFC}	0.05	0.03	0.03	0.35	0.17	0.13	0.38	0.03

(continued)

Table 4 (continued)

Wind direction Wind speed (ms^{-1}) Pasquill stability	N $v > 6$ D	NNE $v > 6$ D	NE $v > 6$ D	N $5 < v \leq 6$ D	NNE $5 < v \leq 6$ D	NE $5 < v \leq 6$ D	N $3 < v \leq 5$ D	NNE $3 < v \leq 5$ E
z_0^{LFN}	0.07	0.06	0.07	0.44	0.07	0.24	0.10	0.92
z_0^{LOC}	0.05	0.04	0.06	0.08	0.05	0.05	0.04	0.04
z_0^{LVI}	0.13	0.32	0.30	0.33	0.30	0.29	0.30	0.34
z_0^{MTR}	0.08	0.96	0.01	0.02	0.02	0.01	0.95	0.80
z_0^{OCT}	0.92	0.91	0.18	0.52	0.47	0.67	0.16	0.12
z_0^{PDA}	0.04	0.04	0.04	0.04	0.04	0.05	0.02	0.03
z_0^{PST}	0.12	0.07	0.13	0.14	0.13	0.13	0.00	0.14
z_0^{RMB}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
z_0^{SDN}	0.02	0.02	0.00	0.02	0.03	0.00	0.02	0.03
z_0^{SNE}	0.03	0.02	0.02	0.02	0.02	0.03	0.03	0.03
z_0^{VAP}	0.39	0.05	0.09	0.48	0.49	0.43	0.02	0.01
z_0^{ZAU}	1.22	0.05	0.84	0.95	0.20	0.04	0.05	0.93
z_0^{ZEV}	0.14	0.12	0.12	0.15	0.04	0.14	0.16	0.15
z_0^{ZQM}	0.20	0.74	0.58	0.35	0.47	0.49	0.38	0.76
d^{ACM}	4.80	82.2	77.3	55.9	38.4	75.6	80.8	84.1
d^{ARR}	0.52	0.79	0.72	0.07	0.66	0.64	0.57	0.51
d^{CLC}	0.16	0.13	0.37	0.17	0.07	0.35	0.31	0.17
d^{CNF}	19.7	13.3	6.27	17.5	8.59	10.4	20.4	13.3
d^{CHL}	2.20	0.23	0.48	2.01	2.97	0.55	2.95	1.06

(continued)

Table 4 (continued)

Wind direction Wind speed (ms^{-1}) Pasquill stability	N $v > 6$ D	NNE $v > 6$ D	NE $v > 6$ D	N $5 < v \leq 6$ D	NNE $5 < v \leq 6$ D	NE $5 < v \leq 6$ D	N $3 < v \leq 5$ D	NNE $3 < v \leq 5$ E
d^{EDF}	10.1	19.1	19.6	10.9	17.4	19.3	7.02	12.8
d^{FDC}	21.4	7.43	10.5	19.4	16.9	6.25	19.8	3.23
d^{FDP}	27.8	3.93	10.7	30.7	18.7	8.10	29.3	4.26
d^{HMA}	0.14	0.54	0.38	0.15	0.76	0.20	0.56	0.42
d^{HSM}	0.06	0.09	0.03	0.10	0.16	0.06	0.04	0.06
d^{LFC}	1.77	0.68	0.32	3.82	2.34	1.25	3.82	0.30
d^{LFN}	1.29	0.50	0.33	1.39	0.60	2.31	2.40	1.58
d^{LOC}	0.34	0.27	0.37	0.23	0.41	0.20	0.34	0.38
d^{LVI}	1.03	0.79	0.76	0.88	0.81	0.75	0.90	1.04
d^{MTR}	2.53	1.07	5.08	6.01	7.06	6.95	6.89	6.84
d^{OCT}	13.9	13.1	6.07	10.3	11.9	13.9	13.4	11.6
d^{PDA}	0.20	0.15	0.12	0.20	0.31	0.19	0.03	0.15
d^{PST}	0.55	0.49	0.13	0.47	0.56	0.40	0.30	0.37
d^{RMB}	0.01	0.02	0.01	0.00	0.02	0.01	0.02	0.00
d^{SDN}	0.09	0.04	0.13	0.06	0.20	0.15	0.04	0.07
d^{SNE}	0.09	0.02	0.09	0.07	0.09	0.05	0.03	0.12
d^{VAP}	2.50	0.32	2.07	2.36	2.47	2.31	0.65	2.08
d^{ZAU}	9.30	5.23	5.81	13.8	13.7	8.64	10.5	10.7
d^{ZEV}	0.17	0.69	0.37	0.70	0.25	0.20	0.73	0.60
d^{ZQM}	3.61	5.40	3.04	3.68	2.10	4.47	2.50	3.14

implementation each parameter is represented by a single value, but their optimal values could be different in each region. For this reason, if the sensibility analysis determines that the results of the model are highly dependent on those parameters, one can split them and consider different unknowns for each region.

Finally, this methodology can be applied to construct a reduced basis for its use in wind forecasting, but a larger period must be studied (at least one year) so that we can analyze all the wind condition types (speed, direction, stability) occurring in Gran Canaria.

4 Conclusions

This chapter presents Wind3D, a mass-consistent diagnostic model with an updated vertical wind profile and atmospheric parameterization which uses an initial wind field with a logarithmic wind profile that consider the effect of both stable boundary layer (SBL) and the convective boundary layer (CBL).

The strategy adopted to deal with the sensitivity of the models to the value of some of its parameters is to estimate their values using a memetic algorithm.

A numerical experiment over Gran Canaria island shows that this model is a suitable tool to study wind fields over complex terrains.

References

1. Burlando M, Georgieva E, Ratto CF (2007) Parameterisation of the planetary boundary layer for diagnostic wind models. Bound-Lay Meteorol 125:389–397. https://doi.org/10.1007/s10546-007-9220-7
2. Montero G, Montenegro R, Escobar JM (1998) A 3-D diagnostic model for wind field adjustment. J Wind Eng Ind Aerod 74–76:249–261. https://doi.org/10.1016/S0167-6105(98)00022-1
3. Sherman CA (1978) A mass-consistent wind model for wind fields over complex terrain. J Appl Meteorol 17(3):312–319
4. Wagenbrenner NS, Forthofer JM, Lamb BK, Shannon KS, Butler BW (2016) Downscaling surface wind predictions from numerical weather prediction models in complex terrain with windninja. Atmos Chem Phys 16(8):5229–5241. https://doi.org/10.5194/acp-16-5229-2016, www.atmos-chem-phys.net/16/5229/2016/
5. Mortensen NG, Landberg L, Troen I, Petersen EL (1993) Wind atlas analysis and application program (WAsP), vol 2: Users guide. Riso National Laboratory, Roskilde, Denmark
6. Walmsley JL, Taylor PA, Keith T (1986) A simple model of neutrally stratified boundary layer flow over complex terrain with surface roughness modulations (ms3djh/3r). Bound-Lay Meteorol 36(1–2):157–186. https://doi.org/10.1007/BF00117466
7. Lopes AMG (2003) Windstation a software for the simulation of atmospheric flows over complex topography. Environ Model Softl 18(1):81–96. https://doi.org/10.1016/S1364-8152(02)00024-5
8. Barnard JC (1991) An evaluation of three models designed for siting wind turbines in areas of complex terrain. Sol Energy 46(3):283–294. https://doi.org/10.1016/0038-092X(91)90096-F

9. Walmsley JL, Troen IB, Demetrius P, Lalas DP, Mason PJ (1990) Surface-layer flow in complex terrain: comparison of models and full-scale observations. Bound-Lay Meteorol 52(3):259–281. https://doi.org/10.1007/BF00122090
10. Homicz GF (2002) Three-dimensional wind field modeling: a review. SAND Report 2597. Technical Report, Sandia National Laboratories, Albuquerque, NM
11. Apsley DD, Castro IP (1997) Flow and dispersion over hills: comparison between numerical predictions and experimental data. J Wind Eng Ind Aerod 67–68:375–386. https://doi.org/10.1016/S0167-6105(97)00087-1
12. Maurizi A, Palma JMLM, Castro FA (1998) Numerical simulation of the atmospheric flow in a mountainous region of the north of portugal. J Wind Eng Ind Aerod 74–76:219–228. https://doi.org/10.1016/S0167-6105(98)00019-1
13. Uchida T, Ohya Y (1999) Numerical simulation of atmospheric flow over complex terrain. J Wind Eng Ind Aerod 81(1–3):283–293. https://doi.org/10.1016/S0167-6105(99)00024-0
14. Montavon C (1998) Validation of a non-hydrostatic numerical model to simulate stratified wind fields over complex topography. J Wind Eng Ind Aerod 74–76:273–282
15. Dong Z, Gao S, Fryrear DW (2001) Drag coefficients, roughness length and zero-plane displacement height as disturbed by artificial standing vegetation. J Arid Environ 49(3):485–505. https://doi.org/10.1006/jare.2001.0807
16. Abtew W, Gregory JM, Borrelli J (1989) Wind profile: estimation of displacement height and aerodynamic roughness. Trans ASAE 32(2):0521–0527. https://doi.org/10.13031/2013.31034
17. Ferragut L, Montenegro R, Montero G, Rodríguez E, Asensio MI, Escobar JM (2010) Comparison between 2.5-D and 3-D realistic models for wind field adjustment. J Wind Eng Ind Aerod 98(10–11):548–558 (2010). https://doi.org/10.1016/j.jweia.2010.04.004
18. Montero G, Rodríguez E, Montenegro R, Escobar JM, González-Yuste JM (2005) Genetic algorithms for an improved parameter estimation with local refinement of tetrahedral meshes in a wind model. Adv Eng Softw 36(1):3–10. https://doi.org/10.1016/j.advengsoft.2004.03.011
19. Winter G, Montero G, Ferragut L, Montenegro R (1995) Adaptive strategies using standard and mixed finite elements for wind field adjustment. Sol Energy 54(1):49–56. http://www.sciencedirect.com/science/article/pii/0038092X9400100R
20. Hestenes MR, Stiefel E (1952) Methods of conjugate gradients for solving linear systems. J Res Natl Bur Stand 49(6):409–436
21. Montero G, Montenegro R, Escobar J, Rodríguez E (2004) Resolution of sparse linear systems of equations: the rpk strategy. In: Topping B, Soares CM (eds) Progress in engineering computational technology, Chap 5. Saxe-Coburg Publications, Stirlingshire, UK, pp 81–109
22. Suárez A, Sarmiento H, Flórez E, García M, Montero G (2011) Updating incomplete factorization preconditioners for shifted linear systems arising in a wind model. J Comput Appl Math 235(8):2640–2646. https://doi.org/10.1016/j.cam.2010.11.015, http://www.sciencedirect.com/science/article/pii/S0377042710006321
23. Zilitinkevich SS, Johansson PE, Mironov DV, Baklanov A (1998) A similarity-theory model for wind profile and resistance law in stably stratified planetary boundary layers. J Wind Eng Ind Aerod 74–76:209–218
24. Zilitinkevich SS, Tyuryakov SA, Troitskaya YI, Mareev EA (2012) Theoretical models of the height of the atmospheric boundary layer and turbulent entrainment at its upper boundary. Atmos Ocean Phys 48(1):150–160. https://doi.org/10.1134/S0001433812010148
25. Zilitinkevich SS, Fedorovich EE, Shabalova MV (1992) Numerical model of a non-steady atmospheric planetary boundary layer, based on similarity theory. Bound-Lay Meteorol 59:387–411
26. Montero G, Rodríguez E, Montenegro R, Escobar J, González-Yuste J (2005) Genetic algorithms for an improved parameter estimation with local refinement of tetrahedral meshes in a wind model. Adv Eng Softw 36(1):3–10. https://doi.org/10.1016/j.advengsoft.2004.03.011, http://www.sciencedirect.com/science/article/pii/S0965997804000742. Evolutionary Optimization of Engineering Problems
27. Oliver A, Rodríguez E, Escobar JM, Montero G, Hortal M, Calvo J, Cascón JM, Montenegro R (2015) Wind forecasting based on the harmonie model and adaptive finite elements. Pure Appl Geophys 172(1):109–120. https://doi.org/10.1007/s00024-014-0913-9

28. Storn R, Price K: Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11(4):341–359. https://doi.org/10.1023/a:1008202821328
29. Greiner D, Emperador JM, Winter G (2004) Single and multiobjective frame optimization by evolutionary algorithms and the auto-adaptive rebirth operator. Comput Methods Appl Mech Engrg 193(33–35):3711–3743. https://doi.org/10.1016/j.cma.2004.02.001
30. Byrd RH, Lu P, Nocedal J, Zhu C (1995) A limited memory algorithm for bound constrained optimization. SIAM J Sci Comput 16(5):1190–1208. https://doi.org/10.1137/0916069
31. Mohan M, Siddiqui TA (1998) Analysis of various schemes for the estimation of atmospheric stability classification. Atmos Environ 32(21):3775–3781. https://doi.org/10.1016/s1352-2310(98)00109-5

Wind Field Deterministic Forecasting

Javier Calvo Sánchez and Gema Morales Martín

Abstract Regional Numerical Weather Prediction (NWP) models are nowadays integrated at resolutions between 1 and 3 km. They are non-hydrostatic models, generally run with explicit deep convection. These models have achieved a significant improvement on high-impact weather simulation comparing with synoptic scale models. Modeling at these scales needs big computer resources. Wind simulations are very sensitive to different features of the model: space resolution, orography representation, surface physiography, and flux exchanges between the surface and the atmosphere. Different formulations and parameterizations are followed to take into account all these topics depending on the stability and the surface properties. This chapter offers a snapshot of how HARMONIE-AROME model deals with these issues to derive a formulation for the 10 m wind.

1 Introduction

Numerical Weather Prediction (NWP) models have improved significantly over the last decades (see Sect. 3). For wind prediction, model resolution is a key aspect. Currently, Limited Area Models (LAM) are run operationally at horizontal resolutions around 1–3 km, but these resolutions may not be enough to represent local wind with complex terrain. There are several methods to further enhance the NWP output but they rely on the quality of the mesoscale model: evolution of the pressure systems, stability of the atmosphere, representation of regional winds and local circulations, etc. When convection takes place, the uncertainty of the model predictions increases and it is recommended to use ensemble methods to estimate the predictability of the forecasts.

J. Calvo Sánchez (✉) · G. Morales Martín
NWP Group of the State Meteorological Agency (AEMET), c/ Leonardo Prieto Castro, 28040 Madrid, Spain
e-mail: fcalvos@aemet.es

G. Morales Martín
e-mail: gmoralesm@aemet.es

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_5

2 HARMONIE-AROME Model

The HARMONIE-AROME non-hydrostatic convection-permitting model is a particular configuration of the ALADIN-HIRLAM shared system resulting from the collaboration between ALADIN and HIRLAM Consortia. This configuration described in [4] is based on the AROME-France model [34].

The model performance is very sensitive to the initial state that is estimated by its assimilation system, based on the 3D-Var scheme developed in ALADIN [8], which shares most of the code with the ECMWF and ARPEGE models. A summary of its main features can be found in [17]. A 4DVar system, under construction, will allow to account for flow-dependent forecast errors, improving the use of observations and diminishing model spin up. Moreover, ensemble assimilation techniques are under development. Currently, the analysis of screen level variables is done using a statistical interpolation algorithm [36]. In the near future, assimilation of other soil parameters as soil moisture and leaf area index will be included using an extended Kalman filter approach.

The spectral dynamical core uses a two-time level semi-implicit semi-Lagrangian discretization based on SETTLS approach [19] which allows long time steps (75 s for a 2.5 km resolution). In order to enhance stability, an upper level nesting is applied using Davies relaxation. The non-hydrostatic component is based on ALADIN dynamics [3, 10].

The physics is adapted from Meso-NH research model [7] as it is described in [34]. Surface processes are treated within an externalized surface model called SURFEX [23] (Surface Externalisée, in French), developed by Météo-France in cooperation with the scientific community. Turbulence scheme follows a turbulent kinetic energy approach [13] and convection in the boundary layer uses the EDMF-M scheme which combines eddy diffusivity and mass flux scheme for shallow convection [24, 33]. Deep convection processes are treated explicitly so the microphysics package plays a very important role in the model performance. The package known as ICE3 is a one-moment bulk scheme which uses a three-class ice parametrization [22, 32].

3 Parameterization of Surface Processes. Wind Representation

The surface fluxes which are input to the atmospheric turbulence and radiation schemes are computed within SURFEX [23, 29], which represents surface heterogeneity dividing each grid box in four surfaces (tiles): nature, water (lake), urban areas, and sea. The fraction of each surface is extracted from a global data base named ECOCLIMAP [28]. The fluxes passed to the atmosphere are the averaged fluxes for each subtype weighted by their relative fraction in the grid cell. All the tiles experiment the same forcing by the mean atmospheric variables and radiative fluxes (Fig. 1).

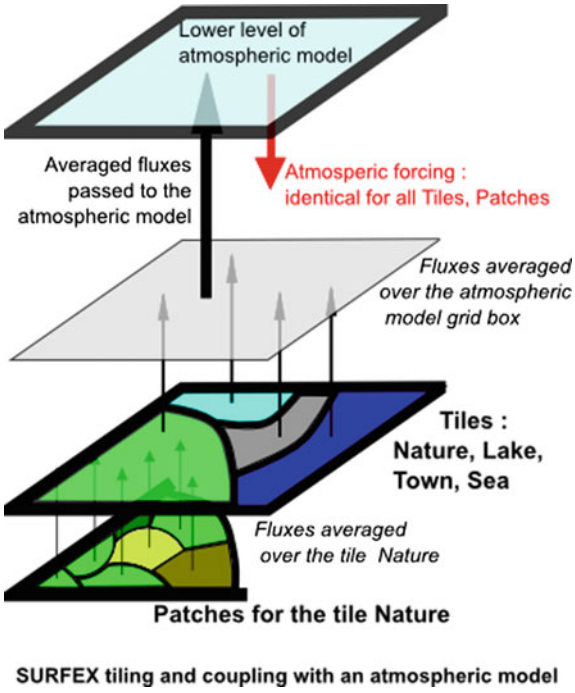


Fig. 1 Tiling approach used in the parameterization of surface processes where fluxes are computed independently in each tile. The grid is divided into 4 tiles and the nature tile is subdivided into 12 patches. The atmosphere feels the averaged fluxes in the grid cell. On the other hand, atmospheric variables and radiative fluxes are sent to the surface where all the tiles receive the same forcing. *Source* CNRM: http://www.umr-cnrm.fr/surfex, [29]

3.1 Soil and Vegetation (ISBA Scheme)

The prognostic equations for surface and soil temperatures and humidities are based on the force-restore method. The soil is divided into several layers including a root zone from which vegetation can extract humidity. Soil freezing effects may play an important role in the energy and humidity fluxes. Vegetation leaves may retain water from precipitation or dew deposition that could be evaporated later. Generally, a one layer snow scheme is used [14].

Following [23], the surface momentum fluxes can be expressed using drag coefficients:

$$\left(\overline{u'w'}\right)_s = -C_D \overline{u}\,|\mathbf{V}| \tag{1}$$

$$\left(\overline{v'w'}\right)_s = -C_D \overline{v}\,|\mathbf{V}| \tag{2}$$

where u, v, w are the wind components, $|\mathbf{V}| = \sqrt{\overline{u}^2 + \overline{v}^2}$ the horizontal wind speed evaluated at first model level, subindex s means evaluated at the surface and C_D a drag coefficient based on Louis formulation [25], modified to consider different roughness lengths for heat z_{0h} and momentum z_0 [26]

$$C_D = C_{DN} F_m \tag{3}$$

where the neutral drag coefficient is

$$C_{DN} = \frac{k^2}{[\ln (z/z_0)]^2} \tag{4}$$

being k the *Von Karmann* constant and the stability function F_m is computed as

$$F_m = 1 - \frac{10 R_i}{1 + C_m \sqrt{|R_i|}} \qquad if\ R_i \leq 0 \tag{5}$$

$$F_m = \frac{1}{1 + \frac{10 R_i}{\sqrt{1+5R_i}}} \qquad if\ R_i > 0 \tag{6}$$

which are the function of the gradient Richardson number R_i. The coefficient C_m of the unstable case is computed using

$$C_m = 10\, C_m^* C_{DN}\, (z/z_0)^{p_m} \tag{7}$$

$$C_m^* = 6.8741 + 2.6933 \times \mu - 0.3601 \times \mu^2 + 0.0154 \times \mu^3 \tag{8}$$

$$p_m = 0.5233 + 0.0815 \times \mu - 0.0135 \times \mu^2 + 0.0010 \times \mu^3 \tag{9}$$

with

$$\mu = \ln (z/z_{0h}) \tag{10}$$

that depends on roughness lengths for momentum and heat.

Vegetation diversity is represented using 12 vegetation types in three categories:

- Bare soil, rocks, permanent snow and ice (bare soil types).
- C3 crops, C4 crops, irrigated crops, natural herbaceous temperate, natural herbaceous tropics, wetland herbaceous, and irrigated grass (herbaceous types).
- needleleaf trees, evergreen broadleaf trees, and deciduous broadleaf trees (woody trees).

Each vegetation type cover has defined parameters obtained from ECOCLIMAP data base [28]. A summary of the different roughness lengths can be found in Table 1.

Table 1 Roughness lengths for different surface and vegetation types [28]. LAI is the *Leaf Area Index* derived from satellite data and having an annual cycle, h is the typical tree height which is 2 m for bushes and ranges from 15–30 m for forests

Surface/Vegetation type	Roughness length (m)
Sea	$0.015(u_*^2/g)$
Ice/snow	0.0013
Bare soil	0.013
Rocks	0.13
C3 crops	0.13 min $[1, e^{(LAI-3.5)/1.3}]$
C4 crops and irrigated crops	0.13 min $[2.5, e^{(LAI-3.5)/1.3}]$
Herbaceous veg	$0.13^{LAI/6}$
Forest	$0.13h$

3.2 Water Surfaces

For sea and lakes, all the prognostic variables are kept constant. The roughness length is given by Charnock's formula:

$$z_{0sea} = 0.015\frac{u_*^2}{g} \tag{11}$$

and with ice ($SST < -2\,^{\circ}\mathrm{C}$) the roughness length is the one used for snow

$$z_{0ice} = 10^{-3} \tag{12}$$

Momentum fluxes follow Louis approach [25] as described for the ISBA scheme.

3.3 Urban Surfaces

The *Town Energy Budget* (TEB) scheme [27] is based on the canyon approach where the energy budgets are computed for three components: roofs, roads, and walls. If snow is present, two additional budgets are considered for snow on roofs and roads. A spatial average of town characteristics is needed so the parameterization performance is quite sensitive to a proper description of the main town features. The parameters of the scheme depend on building shapes and construction materials.

The problem is that the roughness sublayer can be above the first model level (typically around 10 m). Anyway, the momentum fluxes are computed with the roughness length and the stability coefficients using [26]

$$z_{0town} = \frac{h}{10} \tag{13}$$

where h is the typical building height for the entire surface area with a maximum value of 5 m. There are several types of urban surfaces (dense urban, suburban, urban parks, etc.) each one with specific characteristic parameters.

3.4 Coupling Between the Different Surfaces and the Atmosphere. 10 m Wind

Simple interpolation between the lowest level and the surface
The interpolation is done using Monin Obukov diagnostic profile functions including the roughness length and the surface fluxes computed in the surface parameterization. These wind profile functions follow a logarithmic profile corrected for stability effects (Fig. 2, *left*). The lowest model level is supposed to be high enough to be in the inertial sublayer (constant flux layer). This method is appropriate over the ocean and for homogeneous and smooth surfaces.

Surface Boundary Layer scheme (CANOPY scheme)
Another approach for the surface atmosphere coupling consists on dividing the surface-1st model level layer into different sublayers and run a simplified one column model scheme in these layers [30] (Fig. 2, *right*). In this model, the momentum tendencies and the turbulent kinetic energy tendency have additional terms, function of the LAI, and the vegetation height to account for the vegetation drag.

The same method is also used for urban canopies [18]. This method achieves a finer description of the profiles of the mean variables and fluxes in the surface

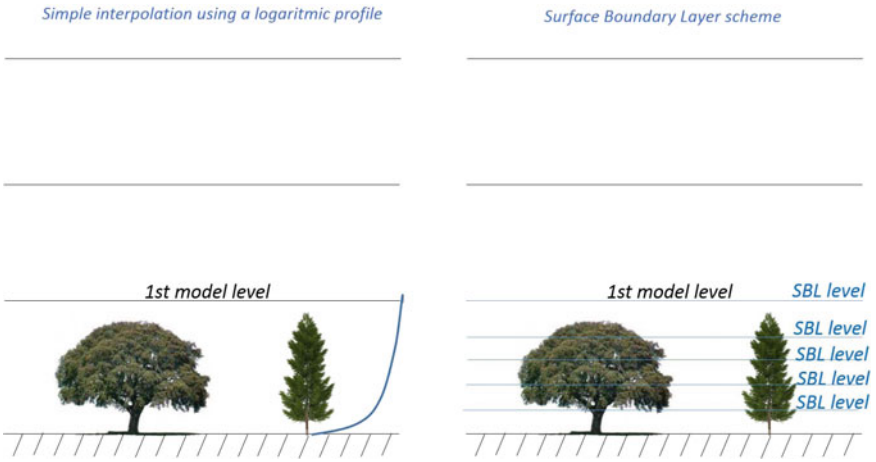


Fig. 2 Coupling between the surface and the first model level using a simple interpolation scheme using a logarithmic profile correcting for stability *(left)* or using the SBL scheme that divides the layer into several sublayers and runs a 1D turbulence scheme accounting for canopy or urban drag *right*

boundary layer that are function of the wind speed and the stability. The method retrieves the logarithmic profile in neutral conditions. In general, it allows a better representation of 2 m variables and 10 m wind. The major improvements are found in stable conditions and for mountainous regions.

3.5 Sub Grid Scale Orography (SSO) Parameterization

There are several options for the orographic drag parameterization in SURFEX. The effects of the small scale orography are parameterized as a momentum sink (drag). The larger scale effects such as mountain blocking and gravity wave breaking are supposed to be resolved at convection-permitting scales (1–3 km resolution).

Z01D
The orographic drag is function of the orographic roughness length z_0 (between 1–60 over orography) that does not depend on wind direction

$$\frac{\partial \left(\overline{u'w'}\right)_{z01D}}{\partial z} = \rho^2 \left[\frac{0.4}{\ln\left(\frac{z}{z_0}\right)}\right]^2 |\mathbf{V}| \tag{14}$$

with z the height of the atmospheric forcing level, $|\mathbf{V}|$ the horizontal wind speed and the roughness length has a maximum value of $z/2$.

Z04D
The same method as *Z01D* but with the roughness length function of the wind direction.

BE04
Following [2], the drag is not function of the roughness but of the sub-grid orography variance σ_{SO}^2

$$\frac{\partial \left(\overline{u'w'}\right)_{BE04}}{\partial z} = C\,\sigma_{SO}^2\, z^{-1.2} e^{-\left[\frac{z}{1500}\right]^{1.5}} |\mathbf{V}| \tag{15}$$

where z is the height, $|\mathbf{V}|$ the horizontal wind speed and the other parameters are constants [2].

Currently, there is no consensus about the benefits of activating the SSO parameterization in HARMONIE-AROME and some operational configurations activate it and others do not.

4 Verification of Operational Wind Forecast

The *State Meteorological Agency of Spain* (AEMET) runs HARMONIE-AROME at 2.5 km horizontal grid spacing over two domains (Iberian Peninsula-Balearic

Islands and the Canary Islands). The vertical discretization includes 65 levels with 15 levels below 1000 m and the model top at 10 hPa. The model analysis updates the atmospheric and surface variables every 3 h using a cutoff time of 1 h and 10 min for the observations, including convectional and aircraft data as well as GNSS zenith total delay and ATOVS satellite data. Other satellite observations and radar data will be included in the near future. The boundary conditions are provided by the ECMWF-IFS integrations corresponding to a cycle 6 h earlier than the Limited Area Model cycle. The HARMONIE-AROME 2.5 km model significantly improves local and extreme forecasts of coarser grid models like HIRLAM or ECMWF [31].

Verification of wind forecast against observations is a key aspect on model validation. In the traditional point verification, model output is interpolated to observation locations and different statistics are computed in order to assess forecast quality [15, 20, 37]. A new model version is only implemented when it is able to improve statistical scores. Comparison of models with different resolutions is a complex issue because double penalty problems take place [15]. A simple way to compare various models or several model versions is comparing the distribution of events in a forecast observation plot as it is done in Fig. 3 for HARMONIE-AROME (HARM) and ECMWF deterministic model. HARM shows a better distribution especially for the strongest winds where ECMWF has a clear tendency to underestimate these events.

Wind velocity shows a clear diurnal cycle as can be seen in Fig. 4 where mean values as function of the hour of the day are plotted for observed values and for several model versions. It should be taken into account that this plot is dominated by low winds that indeed are the ones that occur more frequently.

In order to take into account the uneven distribution of observations, it is very useful to split the verification into different categories corresponding to different intervals of observed wind speed and then compute categorical scores on these intervals. An example of this type of categorical verification is shown in Fig. 5 where Kuiper Skill Score is calculated. HARMONIE-AROME improves ECMWF forecasts for all the forecasting categories.

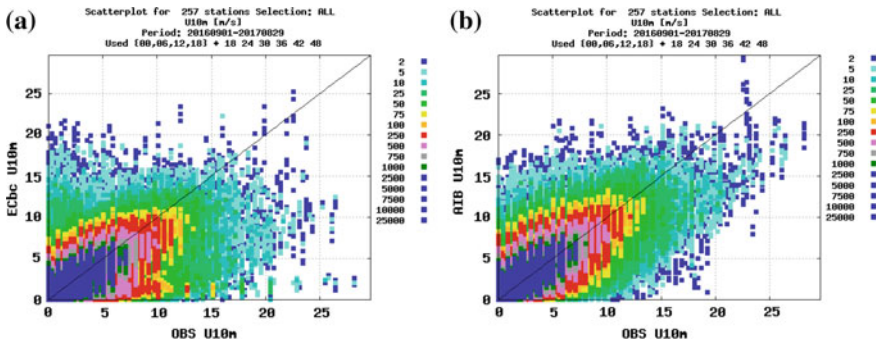


Fig. 3 Comparison of observation forecast events for ECMWF *(upper)* and HARMONIE-AROME *(below)* model for 1 year of forecasts. Narrower distribution and closer to the diagonal implies better forecasts. The biggest differences are found for strong wind cases

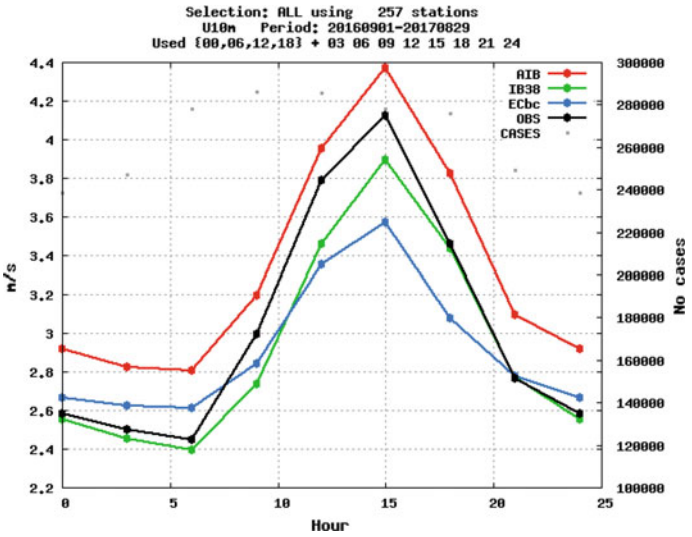


Fig. 4 Comparison of mean values of forecasts and observations plotted as function of the hour of the day showing strongest winds at 15 UTC. Red and green curves correspond to two HARM version whereas green curves correspond to ECMWF forecasts. Basically the models are able to reproduce the diurnal cycle

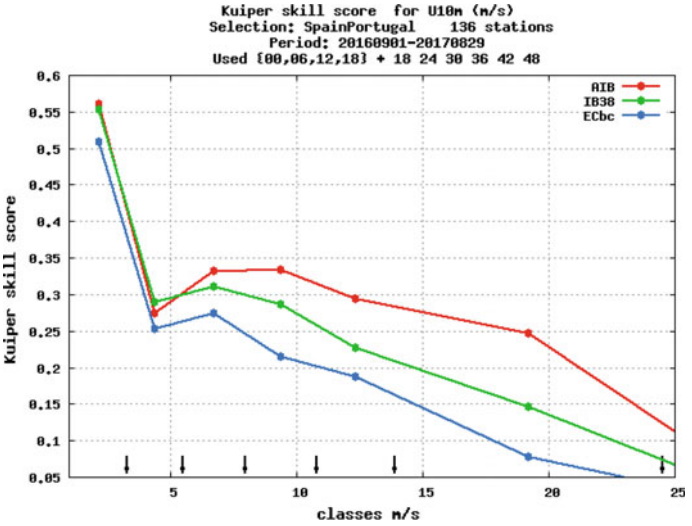


Fig. 5 Kuiper Skil Score comparing ECMWF forecasts *(blue)* with two versions of HARMONIE model (*red* and *green* curves) for 1 year of forecasts. Bigger scores mean better predictions. The differences for wind speeds below 5 m/s are small but the improvement is significant for bigger categories

5 Wind Forecast Case Studies

Mountain ranges exert a significant influence in the atmospheric flow affecting the dynamics of the synoptic systems, producing regional winds and a variety of local effects. Several observation field campaigns have been carried out to improve the understanding of orographic processes [5, 6, 21]. Generally, NWP models use a grid averaged orography what implies a smoothing of the real topographic height and an underestimation of the orographic obstacles. Coarse resolution models, as global models, include a parameterization of orographic processes (blocking effects and breaking of orographic waves) to overcome this limitation. Higher resolution models, as convection-permitting models, resolve better these processes and only include a parameterization for subgrid scale turbulence generated by the topography. Generally, synoptic models (above 10 km resolution) underestimate the orographic drag and the mountain effects [16]. Figure 6 shows a simulation with HIRLAM model at 10 km resolution, using an effective roughness length and no additional parameterization of the orographic processes. The blocking effect of the mountain and regional winds are underestimated in the simulations.

Another example of strong orographic effects took place during the passage of the Tropical Storm Delta over the Canary Islands. In order to simulate this large low pressure system, large modeling domains are necessary and also a good assimilation system, otherwise neither the trajectory nor the intensity can be reproduced. Finally, the major damages in the Islands took place due to downslope windstorms originated

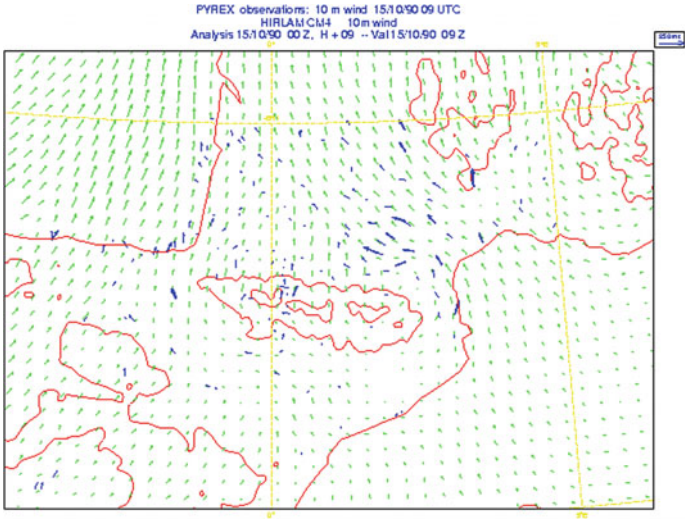


Fig. 6 Comparison of 10 m wind field simulated by HIRLAM model (*green* flags) at 10 km resolution with the observations (*blue*) from the PYREX field campaign [5]. *Red* curves indicate the 1000 and 2000 m topographic height as well as the cost lines. Blocking effects and regional winds are underestimated at this resolution

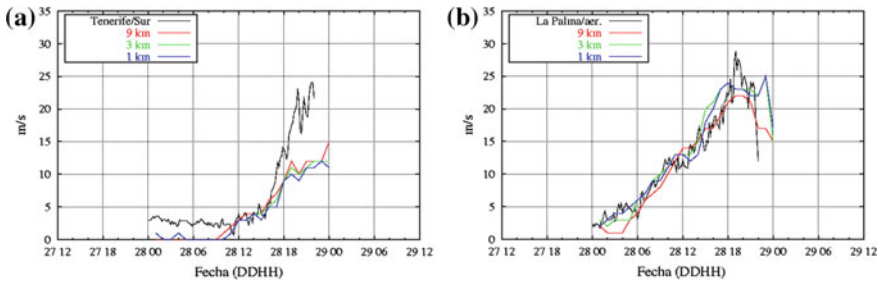


Fig. 7 Comparison of the evolution of the wind speed during 28-11-2005 for two leeward stations (*black*) compared with forecasts at different resolutions. Although the evolution of the wind is generally well represented, the peak in several locations as the Tenerife/Sur station was greatly underestimated even at 1 km resolution

when the flow crossed perpendicularly to the main mountain ranges. Indeed many infrastructures were destroyed and the observations were interrupted due to a generalized power cut. In order to simulate this phenomena, model resolution is a key aspect (Fig. 7):

6 Wind Gust Estimation

Whereas wind field components are forecast variables in the model, wind gust is generally diagnosed using model wind and information from the turbulence scheme [11, 12, 35]. Moreover, the processes leading to gust formation such as deep convection, boundary layer, and orographic processes are generally not well resolved by the models and tend to show a chaotic behavior. Figure 8 shows an example of different methods for estimating gusts associated with the pass of Storm Klaus trough the north of the Iberian Peninsula. The evolution of the gusts is well captured because the evolution of the storm was well reproduced.

However, the errors in the gusts estimation are bigger for purely convective events. Synoptic models use simple parametrization for convective gust but the skill of these estimations is small [1]. On the other hand, convection-permitting models resolve the convective circulations and have more chances of representing better the convective gusts. Nevertheless, at current operational horizontal resolutions, deep convection is not completely resolved what leads to too intense vertical circulations and a general overestimation of the convective gusts.

7 Kilometer and Sub-kilometer Resolutions

Currently, most AROME configurations are run at 2.5 km resolution although Meteo-France is already running the model operationally at 1.3 km including assimilation [9]. Nevertheless, it is important to take into account that the model effective resolution, which is the one of the processes actually resolved, may be six times bigger

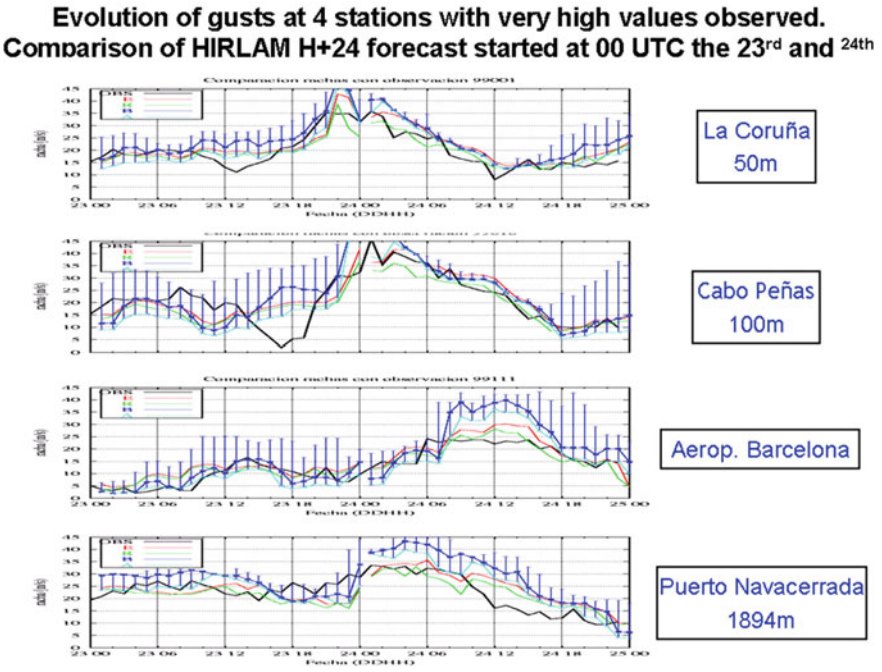


Fig. 8 Extremely intense winds and hurricane force gusts were reported in many places of the Iberian Peninsula when Klaus swept the Peninsula producing several casualties. The model estimate was relatively good because the synoptic evolution of the system was well captured

than the model grid spacing. The AROME system is also run at high resolution in nowcasting mode with very frequent assimilation cycles, going from 10 min to 1 h frequency and performing short forecast lengths (typically up to 6 h). Nowcasting applications normally need ensemble approaches as uncertainty at these scales is big.

Several implementations in the range of 1 km to 500 m are under construction but this is a big challenge because several processes need to be reformulated in the model. At these scales, shallow convection start being resolved by the model (gray scales for shallow convection) so it needs to be redesigned. Besides, to represent turbulence below 500 m, there is need to account for 3D fluxes which implies significant modifications in the current operational codes. Also, the resolution of the physiographic data needs to be enhanced. This includes soil and vegetation characteristics that currently have a resolution around 1 km.

Additionally, high-resolution modeling needs big computer resources, as doubling the model resolution typically implies to increase 8 times the computer cost. Currently, optimization of the models in the context of massive parallel systems is a key aspect and an active field of research (ESCAPE https://www.ecmwf.int/escape and SCALABILITY https://www.ecmwf.int/en/about/what-we-do/scalability programs).

8 Conclusions

Wind simulation has improved significantly using NWP models at convective scales. This is specially the case when orographic processes play an important role. These models generally represent deep convection explicitly having more chances to improve the circulations associated with convection but it should be taken into account that these resolutions are still too broad to resolve completely deep convection. When convection takes place, the predictability decreases and it is advised to follow ensemble approaches to estimate the prediction uncertainty.

The complexity of surface processes is large, in particular in the representation of the wind profile. Different surface covers are taken into account being particularly important the representation of vegetation effects and the presence of water, sea/ice, or urban surfaces. There is no general consensus about the need of an effective roughness length or a subgrid scale orographic parameterization to enhance the orographic effects at convective scale NWP modeling.

References

1. Bechtold P, Bidlot J (2009) Convective gusts. ECMWF, Newsletter No. 119, pp 15–18
2. Beljaars A, Brown R, Wood N (2004) A new parametrization of turbulent orographic form drag. Q J Roy Meteorol Soc 130:1327–1347
3. Bénard P, Vivoda J, Maek J, Smolíková P, Yessad K, Smith C, Broková R, Geleyn J (2010) Dynamical kernel of the Aladin-NH spectral limited-area model: revised formulation and sensitivity experiments. Q J Roy Meteorol Soc 136(646):155–169
4. Bengtsson L, Andrae U, Aspelien T, Batrak Y, Calvo J, de Rooy W, Gleeson E, Hansen-Sass B, Homleid M, Hortal M et al (2017) The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. Mon Weather Rev 145(5):1919–1935
5. Bougeault P, Benech B, Bessemoulin P, Carissimo B, Jansa A, Pelon J, Petitdidier M, Richard E (1997) PYREX: a summary of findings. Bull Am Meteorol Soc 78(4):637–650
6. Bougeault P, Binder P, Buzzi A, Dirks R, Kuettner J, Houze R, Smith R, Steinacker R, Volkert H (2001) The map special observing period. Bull Am Meteorol Soc 82(3):433–462
7. Bougeault P, Mascart P et al (2009) The Meso-NH atmospheric simulation system: scientific documentation part iii: Physics. Technical Report CNRM. http://mesonh.aero.obs-mip.fr/mesonh/dir_doc/book1_m48_19jan2009/scidoc_p3.pdf
8. Brousseau P, Berre L, Bouttier F, Desroziers G (2011) Background-error covariances for a convective-scale data-assimilation system: AROME-France 3D-Var. Q J Roy Meteorol Soc 137(655):409–422
9. Brousseau P, Seity Y, Ricard D, Lger J (2016) Improvement of the forecast of convective activity from the AROME-France system. Q J Roy Meteorol Soc 142(699):2231–2243
10. Bubnová R, Gwenaëlle H, Bénard P, Geleyn J (1995) Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. Mon Weather Rev 123(2):515–535
11. Calvo J, Lopez J, Martín F, Morales G, Pascual R (2009) In: 20th ALADIN workshop and HIRLAM all staff meeting. http://www.umr-cnrm.fr/aladin/IMG/pdf/poster_vientoklaus_final_asm10.pdf
12. Calvo J, Morales G (2009) Verification of wind gust forecasts. In: 19th ALADIN workshop and HIRLAM all staff meeting. http://www.umr-cnrm.fr/aladin/IMG/pdf/ME_JC.pdf

13. Cuxart J, Bougeault P, Redelsperger J (2000) A turbulence scheme allowing for mesoscale and large-eddy simulations. Q J Roy Meteorol Soc 126(562):1–30
14. Douville H, Royer J, Mahfouf J (1995) A new snow parameterization for the Météo-France climate model. Clim Dyn 12(1):21–35
15. Ebert EE (2012) Forecast verification: issues, methods and FAQ. Technical Report WWRP/WGNE Joint Working Group on Forecast Verification Research. http://www.cawcr.gov.au/projects/verification
16. Georgelin M, Bougeault P, Black T, Brzovic N, Buzzi A, Calvo J, Cassé V, Desgagné M, El-Khatib R, Geleyn J, Holt T, Hong SY, Kato T, Katzfey J, Kurihara K, Lacroix B, Lalaurette F, Lemaitre Y, Mailhot J, Majewski D, Malguzzi P, Masson V, McGregor J, Minguzzi E, Paccagnella T, Wilson C (2000) The second compare exercise: a model intercomparison using a case of a typical mesoscale orographic flow, the PYREX iop3. Q J Roy Meteorol Soc 126(564):991–1029
17. Gustafsson N, Janjić T, Schraff C, Leuenberger D, Weissman M, Reich H, Brousseau P, Montmerle T, Wattrelot E, Buánek A, Mile M, Hamdi R, Lindskog M, Barkmeijer J, Dahlbom M, Macpherson B, Ballard S, Inverarity G, Carley J, Alexander C, Dowell D, Liu S, Ikuta Y, Fujita T Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. Q J Roy Meteorol Soc
18. Hamdi R, Masson V (2008) Inclusion of a drag approach in the Town energy Balance (TEB) Scheme: offline 1D evaluation in a street canyon. J Appl Meteorol Climatol 47(10):2627–2644
19. Hortal M (2002) The development and testing of a new two-time-level semi-lagrangian scheme (SETTLS) in the ECMWF forecast model. Q J Roy Meteorol Soc 128(583):1671–1687
20. Jolliffe I, Stephenson D (2003) Forecast verification: a practitioner's guide in atmospheric science. Wiley
21. Kuettner J (1986) The aim and conduct of ALPEX. Scientific results of the Alpine Experiment. Technical Report GARP Publications Series No. 27, WMO/TD 108, 314
22. Lascaux F, Richard E, Pinty J (2006) Numerical simulations of three different MAP IOPs and the associated microphysical processes. Q J Roy Meteorol Soc 132(619):1907–1926
23. Le Moigne P (2012) SURFEX. scientific documentation. Technical Report CNRM. http://www.umr-cnrm.fr/surfex/IMG/pdf/surfex_scidoc_v2-2.pdf
24. Lenderink G, Holtslag AAM (2004) An updated length-scale formulation for turbulent mixing in clear and cloudy boundary layers. Q J Roy Meteorol Soc 130(604):3405–3427
25. Louis J (1979) A parametric model of vertical eddy fluxes in the atmosphere. Bound Layer Meteorol 17(2):187–202
26. Mascart P, Noilhan J, Hervé G (1995) A modified parameterization of flux-profile relationships in the surface layer using different roughness length values for heat and momentum. Bound Layer Meteorol 72(4):331–344
27. Masson V (2000) A physically-based scheme for the urban energy budget in atmospheric models. Bound Layer Meteorol 94(3):357–397
28. Masson V, Champeaux JL, Chauvin F, Meriguet C, Lacaze R (2003) A global database of land surface parameters at 1-km resolution in meteorological and climate models. J Clim 16(9):1261–1282
29. Masson V, Le Moigne P, Martín E, Faroux S, Alias A, Alkama R, Belamari S, Barbu A, Boone A, Bouyssel, F, Brousseau P, Brun E, Calvet JC, Carrer D, Decharme B, Delire C, Donier S, Essaouini K, Gibelin AL, Giordani H, Habets F, Jidane M, Kerdraon G, Kourzeneva E, Lafaysse M, Lafont S, Lebeaupin Brossier C, Lemonsu A, Mahfouf JF, Marguinaud P, Mokhtari M, Morin S, Pigeon G, Salgado R, Seity Y, Taillefer F, Tanguy G, Tulet P, Vincendon B, Vionnet V, Voldoire A (2013) The surfexv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geosci Model Dev 6(4):929–960. https://doi.org/10.5194/gmd-6-929-2013. https://www.geosci-model-dev.net/6/929/2013/
30. Masson V, Seity Y (2009) Including atmospheric layers in vegetation and urban offline surface schemes. J Appl Meteorol Climatol 48(7):1377–1397
31. Navascués B, Calvo J, Morales G et al (2013) Long-term verification of HIRLAM and ECMWF forecasts over Southern Europe: history and perspectives of numerical weather prediction at AEMET. Atmos Res 125–126(Supplement C), 20–33

32. Pinty J, Jabouille P (1998) A mixed–phase cloud parameterization for use in mesoscale non-hydrostatic model: simulations of a squall line and of orographic precipitations. In: Conference on cloud physics, pp 217–220. American Meteorological Society, Everett, WA (1998)
33. de Rooy WC, de Vries H (2017) Technical Report HIRLAM Technical Report
34. Seity Y, Brousseau P, Malardel S, Hello G, Bénard P, Bouttier F, Lac C, Masson V (2011) The AROME-France convective-scale operational model. Mon Weather Rev 139(3):976–991
35. Sheridan P (2011) Review of techniques and research for gust forecasting and parameterisation. Technical Report, Forecasting Research Technical Report 570, Met Office
36. Taillefer (2002) CANARI - Technical Documentation - Based on ARPEGE cycle CY25T1 (AL25T1 for ALADIN). In: CNRM/GMAP. http://www.umr-cnrm.fr/gmapdoc/spip.php?article3
37. Wilks D (2011) Statistical methods in the atmospheric sciences. Academic Press

Wind Field Probabilistic Forecasting

Albert Oliver, Eduardo Rodríguez and Luis Mazorra-Aguiar

Abstract Probabilistic wind forecasting is a methodology to deal with uncertainties in numerical weather prediction models (NWP). In this chapter, we describe the need for ensemble forecasting, the different techniques used to generate the different initial conditions, and the operational ensemble models that are used nowadays in meteorological agencies. Then, we develop an ensemble method designed for the downscaling wind model described in Chap. 4 coupled with the AROME–HARMONIE mesoscale model, a non-hydrostatic dynamic forecast model described in Chap. 5. As we have explained in Chap. 4, some parameters need to be estimated since we do not know its exact value. These parameters are, basically, the roughness length and the zero plane displacement (explained in Chap. 2), as well as the Gauss moduli parameter (α) used in the diagnostic wind model. This estimation is the main source of uncertainties in the model; therefore we will estimate some of these parameters using different forecast values of the AROME–HARMONIE. Finally, an example of the approach is applied in Gran Canaria island with a comparison of the ensemble results with experimental data from AEMET meteorological stations.

A. Oliver (✉) · E. Rodríguez · L. Mazorra-Aguiar
University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain
e-mail: albert.oliver@ulpgc.es

E. Rodríguez
e-mail: eduardo.rodriguez@ulpgc.es

L. Mazorra-Aguiar
e-mail: luis.mazorra@ulpgc.es

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_6

1 Probabilistic Forecasting

Up to this point, we have described deterministic weather models. These models are governed by the initial state, and the errors in this state grow as the model predicts the future, since the models are unstable systems characterized by nonperiodicity. So, the accuracy of the forecast depends on the initial state which is uncertain.

This relationship between the initial state and the deterministic prediction was discovered by Edward Lorenz and is discussed in his book "The Essence of Chaos" [24]. In 1962 [22], he simulated the evolution of the atmospheric state using the geostrophic form of the two-layer baroclinic model proposed in [21] consisting of 12 ordinary differential equations in 12 variables. It used a linear regression from the output of a model. When he ran the simulations, he found out that some solutions were drastically different. Analyzing the results, he found out that, in some experiments, he had truncated the model output to three digits accuracy while the original values had a precision of six digits. Just this small change lead to significant differences in the forecast results. These differences imply that observations need a precision up to the three decimal places to obtain a reliable forecast.

This result prompted the scientific community to determine a procedure to determine which is the best forecast of the atmosphere state according to the available data. Nowadays, there are several meteorological agencies worldwide running their numerical weather prediction models (NWP), each one different from the others. The results from these models are consistent with the observed data but they differ between them, so we cannot say which model is the "correct one". Instead, we can think of each forecast as a member of an ensemble of atmospheric states that are consistent with the observations.

With this idea, Epstein realized that the atmosphere is deterministic since it obeys the fundamental laws of hydrodynamics, but its state can only be known in a probabilistic way. Therefore, in [11], he proposed a "stochastic dynamic" (SD) approach consisting in using the continuity equation for probability [14] in the observations data. He compared the results of the SD model with the results of a deterministic model that used as the initial condition the ensemble mean from the Monte Carlo method.

The problem with SD is that it is expensive; the number of equations for SD prediction is equal to the number of spectral components raised to the power of the number of moments. Philip Thompson [34] proposed a more efficient model by using variances directly instead of covariances; this way the number of equations was reduced.

With the advent of parallel machines, researchers developed different approaches to deal with the uncertainty of the initial state. Murphy [26] ran an experiment using the hemispheric version of the Meteorological Office (UKMO) five-level general circulation model. Initial conditions were obtained by perturbing a given state. Seven individual perturbations were used, and the ensemble forecast consisted of their integration.

To obtain the perturbed initial state, Murphy considers two different methods: random perturbation and lagged-averaged forecast. The random perturbation generates the seven initial states by adding independent perturbations to the known initial state. These perturbations are consistent with some analysis errors. This random perturbation method is similar to the Monte Carlo method. The lagged-averaged forecast method uses past observations to generate each member of the ensemble.

The Monte Carlo (or random perturbation) approach has some limitation; for example, the perturbed parameters can lead to imbalances in the atmospheric state. Another issue is that the perturbations in the Monte Carlo approach were random, while the parameters should have certain preferred directions. With this ideas, different strategies for perturbing dynamical prediction models were studied. The two more used methods nowadays are the singular vector decomposition (SV) [1, 10, 28] used by the European Centre for Medium-Range Weather Forecasts (ECMWF) [19] and the Breeding Vector technique (BV) used by the National Centers of Environmental Prediction (NCEP) [36]. A comparison between the two methods using the ECMWF Integrated Forecast System is described in [25].

These advances, along with more powerful parallel machines, and improvements in deterministic forecasting [33], led to the birth of Ensemble Prediction Systems (EPSs). EPSs are operational systems that provide probabilistic forecasts based on ensemble members. The method to create these ensemble members is different between systems. The Meteorological Service of Canada (MSC) uses a Monte Carlo approach, and, as said previously, ECMWF uses SV, and NCEP uses BV [4].

More recently, a new approach to ensemble forecasting has been developed, the multimodel ensemble forecast. This approach uses forecasts from different models as ensemble members. The ensemble may be composed of deterministic forecasts or from ensemble prediction systems (called superensemble). The idea is to combine the strengths and weaknesses of each model and obtain a more reliable prediction [9, 16].

The THORPEX Interactive Grand Global Ensemble (TIGGE) [3] is a multimodel ensemble system that combines the predictions of the following models: ECMWF, UK Met Office (UKMO), National Centre for Medium Range Weather Forecasting—India (NCMRWF), CMA, Japan Meteorological Agency (JMA), National Centers for Environmental Prediction (NCEP-USA), Meteorological Service of Canada (CMC), Bureau of Meteorology Australia (BOM), Centro de Previsao Tempo e Estudos Climaticos Brazil (CPTEC), Korea Meteorological Administration (KMA), and MeteoFrance (MF) global models. Apart from this global initiative, there is the North American Ensemble Forecasting System (NAEFS) [7] that combines the systems from the Canadian Meteorological Centre (CMC) and the National Centers for Environmental Prediction (NCEP); and an European initiative: the Development of a European Multimodel Ensemble System for Seasonal to Interannual Prediction project (DEMETER) [27].

If the reader is interested in these developments, John M. Lewis [20] wrote a more thorough review of the history of ensemble models.

1.1 Initial State Perturbation Methods

In this subsection, we describe the two most used methods to perturbate the initial state. A simple way of converting a deterministic forecast into a probabilistic forecast would be to modify the deterministic result using a probability distribution constructed from previous forecast errors. This strategy would not work because the underpinning dynamical equations are nonlinear, then the errors at the initial state do not relate directly to the predicted result. So we need to perturbate the initial state. The Monte Carlo approach is to create a random perturbation of the initial state according to their known error characteristics. However, this leads to underdispersive forecast ensembles [5]. The reason for this is that there are many unrepresented sources of uncertainty not explicitly represented in a Monte Carlo forecast.

For this reason, new techniques were required to represent the nonlinearity of the dynamical equations in the ensemble predictions. Two of the most commons techniques will be discussed in this sections; the Singular Vector decomposition, and the Breeding Vector technique.

1.1.1 Singular Vector Decomposition

The main idea behind Singular Vector decomposition is the singular value decomposition of the forward tangent linear operator. This can be physically interpreted as the fastest growing perturbations. Therefore, SVs give information about the direction and dynamics of rapidly growing instabilities and perturbations.

The method was devised by Lacarra and Talagrand in [18] where they were interested in identifying the perturbations that lead to the maximum difference between the simulated state and a reference one. They defined $\mathbf{x}(0)$ as the vector containing the initial state information. The model is defined as $\mathbf{M} : \mathbb{R}^n \to \mathbb{R}^n$. Therefore the state at time t is defined as

$$\mathbf{x}(t) = \mathbf{M}(\mathbf{x}(0)) \tag{1}$$

Since they were interested in knowing the perturbations that differed more from a reference state they need to know how the state evolves. For this reason, they define the resolvent of $\mathbf{M}$ as

$$\mathbf{F}(\mathbf{x}) = \frac{d\mathbf{x}}{dt} \tag{2}$$

If the perturbed initial state is defined as $(\mathbf{x}(0) + \boldsymbol{\chi}(0))$, then the time evolution of the perturbed state can be written as

$$\frac{d}{dt}(\mathbf{x}(t) + \boldsymbol{\chi}(t)) = \mathbf{F}(\mathbf{x}(t) + \boldsymbol{\chi}(t)) = \mathbf{F}(\mathbf{x}(t)) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{\mathbf{x}(t)} \boldsymbol{\chi}(t) + \mathscr{O}\left(\boldsymbol{\chi}^2(t)\right) \tag{3}$$

the second-order term can be neglected, and the derivative of $\boldsymbol{\chi}$ is

$$\frac{d}{dt}(\mathbf{x}(t) + \boldsymbol{\chi}(t)) = \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{\mathbf{x}(t)} \boldsymbol{\chi}(t) \tag{4}$$

This linear system of equations is called the tangent linear system of $\mathbf{M}$ in the vicinity of the particular solution $\mathbf{x}(t)$. It describes the temporal evolution of the perturbation $\boldsymbol{\chi}(t)$, to first order concerning the initial perturbation $\boldsymbol{\chi}(0)$. We can rewrite Eq. (4) as

$$\chi(t) = \mathbf{L}(0, t)\boldsymbol{\chi}(0) \tag{5}$$

where the operator $\mathbf{L}(0, t)$ is the forward tangent linear operator or the linear propagator. So the perturbations that will maximize the difference can be found using the singular value decomposition of $\mathbf{L}(0, t)$.

$$\mathbf{L} = \mathbf{W}\Lambda\mathbf{Y}^* \tag{6}$$

where Λ is a diagonal matrix with the singular values of $\mathbf{L}$ $(\lambda_1, \lambda_1, \ldots)$. $\mathbf{Y}^*$ is the conjugate transpose of $\mathbf{Y}$. The columns of $\mathbf{Y}$ correspond to the initial (or right) singular vectors. The columns of $\mathbf{W}$ are the evolved (or left) singular vectors.

The singular vectors of $\mathbf{L}$ are the same as the eigenvectors of $\mathbf{L}^*\mathbf{L}$. And, specifically, $\mathbf{Y}$ and $\mathbf{W}$ are related in the following manner:

$$\mathbf{L}^*\mathbf{L}\mathbf{y}_i = \lambda_i^2\mathbf{y}_i \tag{7}$$

$$\mathbf{L}\mathbf{L}^*\mathbf{w}_i = \lambda_i^2\mathbf{w}_i \tag{8}$$

To find the perturbations with the maximum amplitude growth, we need to compute them. To this end, we can use any norm $\mathbf{E}$

$$||\chi||_E = \langle \chi, \mathbf{E}\chi \rangle \tag{9}$$

where $\mathbf{E}$ is a matrix operator that defines the inner product.

For a linear operator $\mathbf{L}$, exists its adjoint $\mathbf{L}^*$ such that $\langle \chi, \mathbf{L}y \rangle = \langle \mathbf{L}^*\chi, y \rangle$. Its possible to choose different norms at the initial and the final time

$$||\chi(t_0)||^2_{E_0} = \langle \chi(t_0), E_0\chi(t_0) \rangle \tag{10}$$

$$||\chi(t)||^2_{E_t} = \langle \chi(t), E_t\chi(t) \rangle \tag{11}$$

The objective is to maximize the growth rate, or amplification factor, defined as

$$\begin{aligned}\lambda^2 &= \frac{||\chi(t)||^2_{E_t}}{||\chi(t_0)||^2_{E_0}} = \frac{\langle \chi(t), E_t\chi(t)\rangle}{\langle \chi(t_0), E_0\chi(t_0)\rangle} \\ &= \frac{\langle \mathbf{L}\chi(t_0), E_t\mathbf{L}\chi(t_0)\rangle}{\langle \chi(t_0), E_0\chi(t_0)\rangle} = \frac{\langle \mathbf{L}^* E_t\mathbf{L}\chi(t_0), \chi(t_0)\rangle}{\langle \chi(t_0), E_0\chi(t_0)\rangle}\end{aligned} \tag{12}$$

To maximize λ^2, we solve the following eigenvalue problem:

$$\left(\mathbf{L}^* E_t\mathbf{L}\right) y_i(t_0) = \lambda_i^2 E_0 y_i(t_0) \tag{13}$$

We can rewrite this equation using the variable transformation $y_i(t_0) = E_0^{-\frac{1}{2}}\gamma_i(t_0)$:

$$\left(E_0^{-\frac{1}{2}}\mathbf{L}^* E_t\mathbf{L}E_0^{-\frac{1}{2}}\right)\gamma_i(t_0) = \lambda_i^2\gamma_i(t_0) \tag{14}$$

This equation has the same form as Eq. (7); comparing them we can conclude that the eigenvectors of $E_0^{-\frac{1}{2}}\mathbf{L}^* E_t\mathbf{L}E_0^{-\frac{1}{2}} = \left(E_0^{-\frac{1}{2}}\mathbf{L}^* E_t^{\frac{1}{2}}\right)\left(E_t^{\frac{1}{2}}\mathbf{L}E_0^{-\frac{1}{2}}\right) = \mathbf{L}_s^*\mathbf{L}_s$ are the initial singular vectors of $\mathbf{L}_s$; and they represent the perturbations with a maximum amplification factor in the time interval (t_0, t).

When used in real numerical weather prediction models, the calculation of the singular vector is difficult because the definition of the model $\mathbf{M}$ has to be computed analytically. In operational ensemble prediction systems, this calculation is made using tangent linear and adjoint models and an iterative Lanczos algorithm [6, 12]. A review of the method with applications to El Niño as well as decadal forecasting is presented in [29]. Also, Diaconescu and Laprise [8] review the applications such as forecast error estimation, ensemble forecasting, target adaptive observations, predictability studies and growth arising from instabilities.

1.1.2 Breeding Vector

This method is the most computationally inexpensive [38]. There are two different versions of this method: the simple breeding [35], and the masked breeding [36].

The main idea of the method is that the choice of the initial perturbation has to cover all the space of possible analysis errors. In an operational NWP, the perturbation of the initial state is reduced by the use of observations. Therefore, the most important errors are those associated with the evolution of the model. The breeding method modifies the perturbation using the difference between the perturbed and the unperturbed forecast. Using this technique, all random perturbations develop into the structure of the leading local (time-dependent) Lyapunov vectors (LLVs; see [37]) of the atmosphere after a transient period.

Toth and Kalnay [36] describe the main steps of the breeding method as

1. add a small, arbitrary perturbation to the atmospheric analysis (initial state) at a given day t_0
2. integrate the model from both the perturbed and unperturbed initial conditions for a short period t_1
3. subtract one forecast from the other
4. scale down the difference field so that it has the same norm as the initial perturbation
5. add this difference into the analysis corresponding to the following period t_1

By construction, this method "breeds" the nonlinear perturbations that grow fastest. Therefore, independent perturbations will converge to the same perturbations after enough time steps. This perturbation is related to LLVs. LLVs have been used to characterize the behavior of dynamical systems. The Lyapunov exponents(λ_i) are defined as

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t} \log_2 \left(\frac{p_i(t)}{p_i(0)} \right) \tag{15}$$

where p is a linear perturbation spanning the phase space of the system with orthogonal vectors.

Each Lyapunov exponents can be associated with a perturbation vector. The vector associated to the largest exponent has the property that any random perturbation introduced an infinitely long time earlier develops into it. Lorenz [23] described this property; he noted that initially random perturbations had a strong similarity after 8 days of integration. The breeding method converges to this LLVs after 3 or 4 days of integration.

The masked breeding is the same as the simple breeding described before, but taking into account the geographically dependent uncertainty.

1.2 Multimodel Ensemble Methods

The rationale behind multimodel ensemble methods is that collective information is better than single information, especially the more complex the process. In the concrete case of short- and medium-range weather forecasting Sanders, it was demonstrated that combining different forecast could be beneficial [2, 15, 32]. Combining multiple models, Fritsch et al. [13] suggested that the superiority of the forecast relied on the variations in model physics and numerics between models leading to a substantial role in generating the full spectrum of possible solutions.

However, we should note that model physics and numerics is not enough, another source of uncertainty is the initial state of the atmosphere. This kind of uncertainties is handled by Ensemble Prediction Systems using a technique to perturbate the initial state (such the ones described in Sect. 1.1). So, a good idea could be to combine both

models. Palmer et al. [27] developed a European multimodel ensemble system known as DEMETER.

When developing a multimodel ensemble system, there are several choices to be made. For example, we can consider all the individual forecasts equal, so we just combine them with the same weight. However, more complex methods of optimally combining the single-model output have been described [17, 30, 31]. Another aspect is how the initial state is perturbed; is it better to use the same perturbation in all models? Or we should use the default perturbation technique for each model?

In the concrete case of DEMETER, from each model, except that of the Max-Planck Institute (MPI), uncertainties in the initial state are represented through an ensemble of nine different ocean initial conditions. Three different ocean analyses; a control ocean analysis is forced with momentum, heat, and mass flux data from the ECMWF 40-yr Reanalysis, and two perturbed ocean analyses are created by adding daily wind stress perturbations to the ERA-40 momentum fluxes. The wind stress perturbations are randomly taken from a set of monthly differences between two quasi-independent analyses. Also, to represent the uncertainty in SSTs, four SST perturbations are added and subtracted at the start of the hindcasts. As in the case of the wind perturbations, the SST perturbations are based on differences between two quasi-independent SST analyses. Atmospheric and land surface initial conditions are taken directly from ERA-40.

Palmer [27] concludes that the multimodel ensemble is a viable, pragmatic approach to the problem of representing model uncertainty in seasonal-to-interannual prediction, and leads to a more reliable forecasting system than that based on any one single model.

A study of the superiority of multimodel ensemble systems has been done by Hagedorn et al. in [9, 16].

2 Ensemble Model for Diagnostic Wind Field

Given the importance of introducing the uncertainties in the prediction of the wind field, in this chapter, we describe a simple ensemble method designed for Wind3D, the diagnostic wind model presented in Chap. 4. In the same spirit as Wind3D, the ensemble approach described in this section is a fast procedure designed for the microscale.

Schematically, in any NWP, the main sources of uncertainty comes from observations, model parameters, data assimilation procedures, and boundary conditions.

In the wind model described in Chap. 4, we have detected the parameters with more uncertainty, namely: Gauss moduli parameter (α), roughness length (z_0), and displacement height (d).If we categorize these uncertain parameters in the four categories defined above, α belongs to the model parameters while z_0 and d belong to boundary conditions. An evolutionary algorithm has been presented to characterize these parameters. However, it has been noted that even the “best estimation” has

some uncertainty; in Sect. 4.2 several evolutionary algorithms have been run leading to different parameter estimations.

Another source of uncertainty in Wind3D comes from the observations. Please, remember that these observations can originate from measurement stations or the forecast of a deterministic NWP. In the case of the measurement data, the errors are related to the machine and the daily conditions whereas in the deterministic NWP forecast, we are using the "best forecast" provided by the NWP, but we have already seen that this forecast may be inaccurate. Moreover, due to the differences in horizontal resolution between the local scale diagnostic wind model and the NWP, the height of the grid points between models can be inconsistent. In this case, we do not know if these points are reliable for Wind3D. So, we may ask ourselves "Which are the reliable NWP forecast points?"

Since the method described is an ensemble forecast system, the wind model is used in conjunction with an NWP to have the predictability capability. In this case, to be able to estimate the variables, we need two different sets of data, the set used to run the wind model and the set of observations the results are compared against. Instantly another question arises "How do we generate these sets?".

The ensemble model described here tries to answer the two doubts that have arisen. The model chooses the valid NWP points based on the difference between their height; when the difference between the NWP height and the diagnostic height is lower than a threshold, the point is valid. Once we have chosen the viable points, we construct the two subsets (model observations and validation data) using a random selection. Once the two subsets are created, we estimate the best values for α, ε, z_0, and d using the memetic algorithm discussed in Sect. 4.2. Figure 1 shows the diagram of the method.

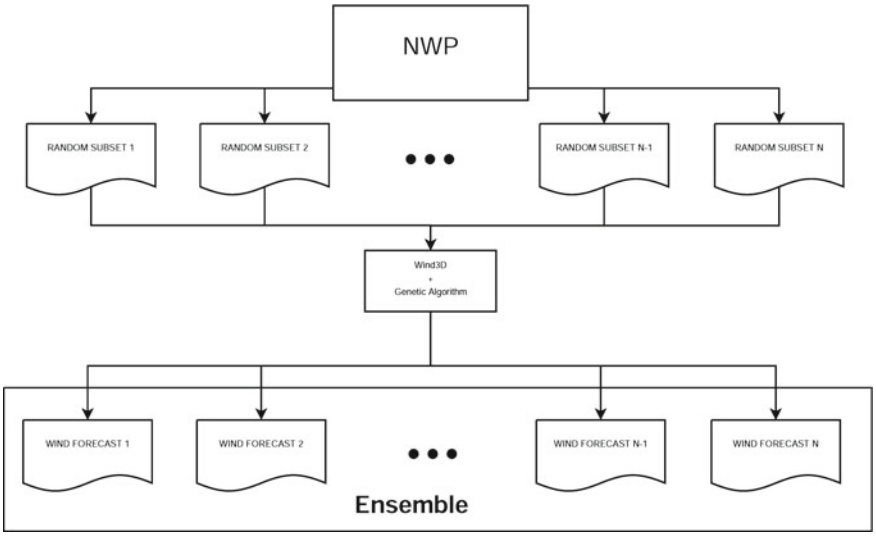


Fig. 1 Diagram of the ensemble system

This method can also be used with various NWP forecast emulating a multimodel ensemble. For example, we can have some ensemble members from ECMWF model, other members from NCEP, and the rest from AROME–HARMONIE.

3 Numerical Experiment

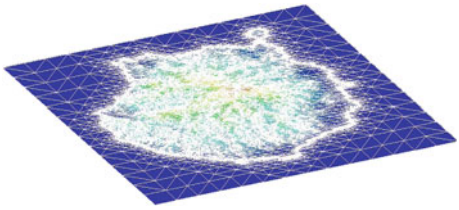
In this section, we present an application of the presented methodology. The application is in Gran Canaria island. The ensemble forecast is generated from AROME–HARMONIE forecast with a horizontal resolution of 2.5 km. The ensemble model is validated against measured data from the AEMET network stations. The day of the simulation is February 20, 2010.

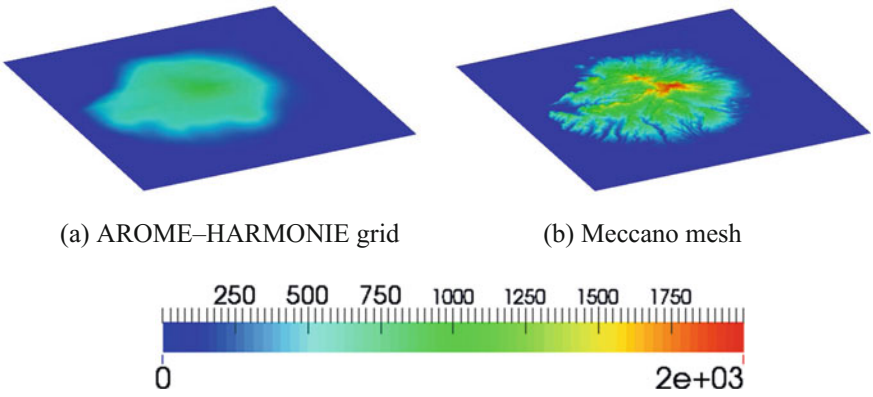
The mesh created for this application is created with the Meccano method (Chap. 3) from a digital terrain model of the Gran Canaria island. The height of the domain is 10.000 m., and the resulting mesh has 251.808 nodes and 1.090.366 tetrahedra (Fig. 2)

Figure 3 shows the terrain height in the Meccano mesh and the AROME–HARMONIE grid. We can observe the differences between the height considered by the Wind3D and AROME–HARMONIE. The maximum height is around 1.000 m in the AROME–HARMONIE discretization and 2.000 m in the Wind3D discretization. This big height difference indicates that, at some points, the AROME–HARMONIE 10 m velocity may not be appropriate. For this reason, instead of using all the 10 m data, we have selected a subset of points attending to a height difference criteria.

Once a set of points has been chosen, we randomly divide them into two different subsets. One subset is used as observations in Wind3D, and the other subset is used by the evolutionary algorithm to compute the fitting function. The fitting function is the Root Mean Square Error (RMSE) between the forecast values by Wind3D and the data in the second subset. In this case, we have selected the points which height difference is less than 50 m. These selected points are shown in Fig. 4 (left). The two randomly generated subsets can be seen in Fig. 4 (right); green points are used as observations for Wind3D, and red points are used to compute the RMSE.

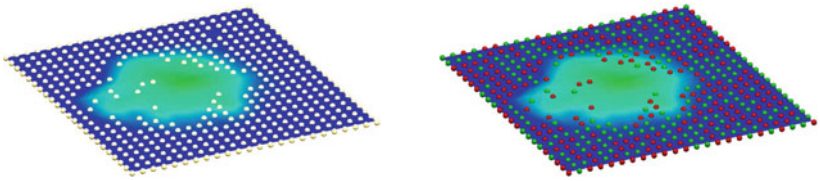
Fig. 2 Terrain discretization





(a) AROME–HARMONIE grid (b) Meccano mesh

Fig. 3 Terrain heights (m.)



(a) Selected points based on height differences

(b) Random selection of subsets: observations (green) and control points (red)

Fig. 4 AROME–HARMONIE points used in simulation

Now we have generated all the members of the ensemble. Then, we estimate the best values of α, ε, z_0 and d, and with these best values, we compute the forecast wind using the Wind3D model.

Finally, to validate the method, we compare the ensemble forecast results with the observed data measured in the AEMET network of automatic stations. Each station provides two data; the average and the maximum wind velocity of the last 10 min. Their UTM coordinates are summarized in Table 1, and their position in a map is shown in Fig. 5.

Figure 6 shows the comparison of measured data and the ensemble box plot forecast. We show the most representative comparisons from four stations. The first thing that we can notice is that, in general, the mean value of the ensemble forecast is reasonably similar to the measured wind velocity. In some cases, the forecasted velocity is close to the maximum (C625O, C639Y), in some others, it is close to the average velocity (C619X), and sometimes it is in between (C635B).

Another observation is that the variation of the mean value of the ensemble forecast is smoother than the measured velocity. In contrast, the measured data exhibits abrupt

Table 1 Location of measurement stations (UTM coordinates)

Station	$X(m)$	$Y(m)$
C619X	429.982	3.108.577
C629Q	429.966	3.073.034
C635B	443.504	3.088.472
C639X	455.377	3.076.514
C639Y	443.283	3.070.534
C625O	436.499	3.081.522

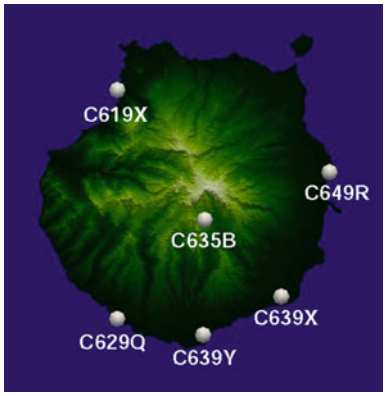
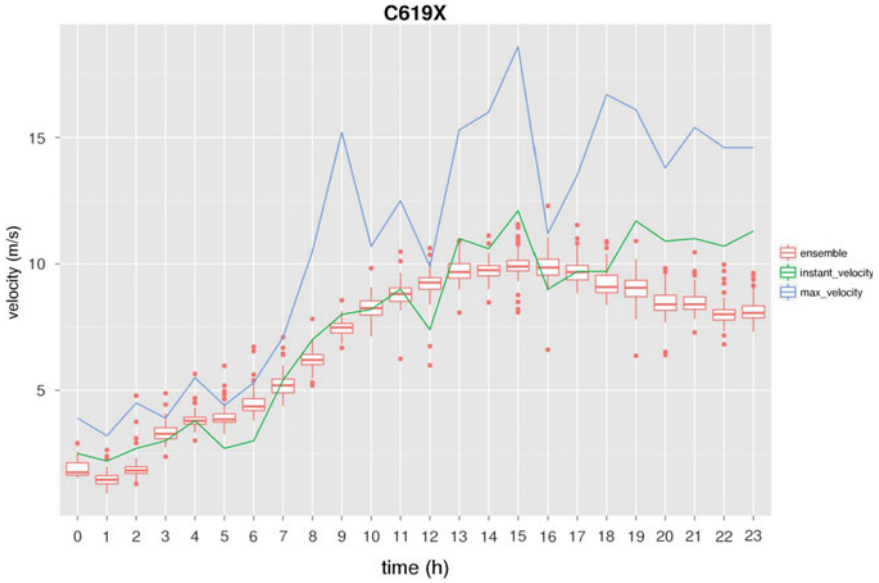


Fig. 5 Location of the AEMET measurement stations

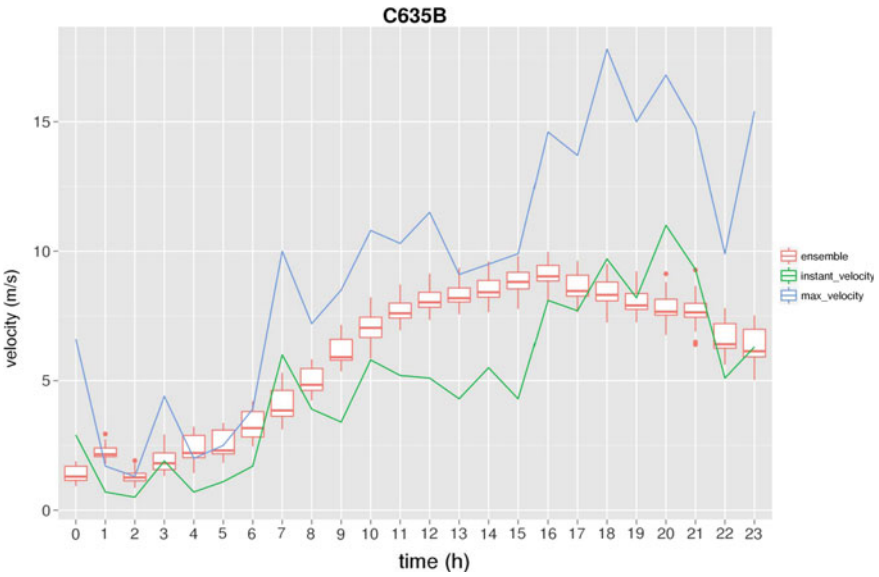
changes among time steps. These abrupt changes are not captured by any member of the ensemble.

A more detailed inspection of the comparatives shows interesting remarks. For example, the ensemble forecast in station C619X has many outliers in all time steps. C639Y also has some of them, but they are close to the mean values. However, C635B and C625O do not have outliers in all the time steps. These outliers sometimes can provide interesting information, for example, in station C619X from 0–7 h they capture the total variation between the average data and the maximum.

C625O station deserves a special mention. Analyzed carefully, we can observe that, between 11 h and midnight, the difference between maximum and average measured data increases. This increase is captured in the ensemble forecast by the higher dispersion of the box plot. This agreement between ranges shows that the resulting ensemble probability can be useful in predicting the uncertainty of the wind velocity.

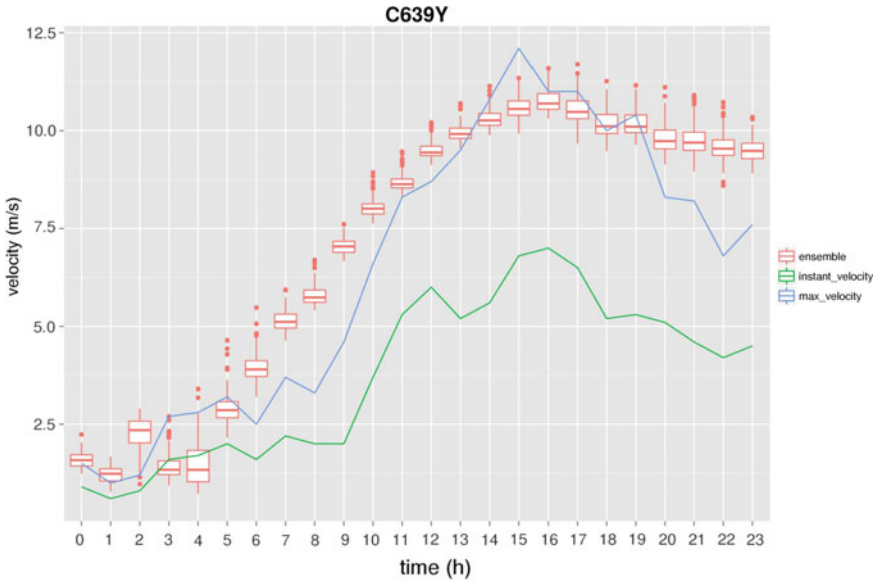


(a) C619X measurement station

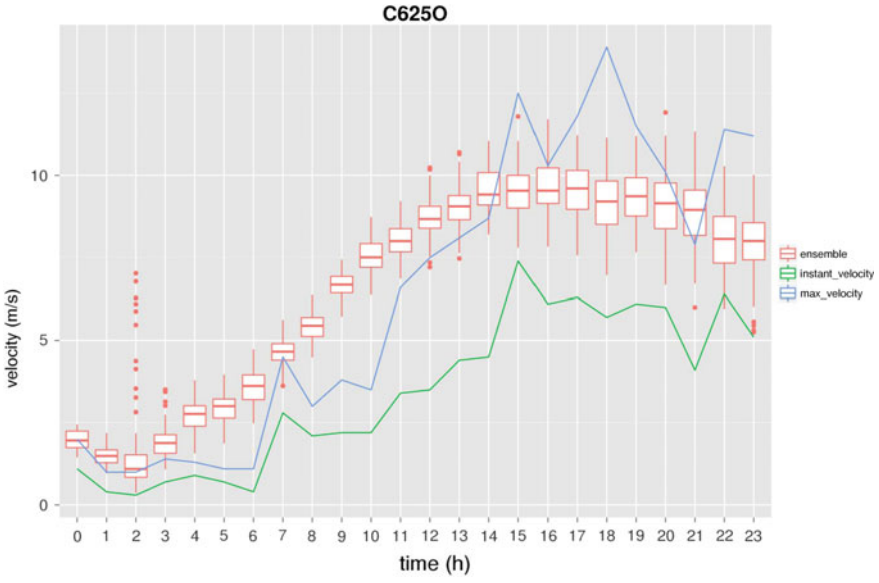


(b) C635B measurement station

Fig. 6 Comparison of the average and maximum measured data and the ensemble box plot forecast



(c) C639Y measurement station



(d) C625O measurement station

Fig. 6 (continued)

4 Conclusions

In this chapter, we have seen the necessity of a probabilistic approach to numerical weather prediction is necessary. It is introduced with a brief review of the progress done in this area: the discovery of the need for a probabilistic approach and the development of these techniques. Then we go into more detail with the description of two of the more used methods to perturbate the initial state; Singular Vector decomposition and Breeding Vectors. To finish the introduction, we describe the basis of a multimodel ensemble method.

Next, we describe an ensemble forecast method specially designed for the microscale. This method is based on the estimation of the uncertain parameters using an evolutionary algorithm. The uncertain parameters are both model parameters, i.e., α and ε, and physical parameters, namely the roughness length (z_0) and the displacement height (d). The evolutionary algorithm minimizes the error of the predicted wind field by a microscale wind model and the forecast of an NWP. The NWP forecast is used for the input data of the model and the control data to compute the fitting function of the evolutionary algorithm. The selection of these two subsets is random and generates the different members of the ensemble system.

Finally, to illustrate the methodology and validate the model, we present a numerical experiment. In this experiment, we use the microscale model Wind3D described in Chap. 4 coupled with the AROME–HARMONIE model described in Chap. 5. The experiment is located in Gran Canaria island during February 20, 2010. The results have shown that, at any predicted time and station, the forecast ensemble probability lies between the average and the maximum velocity, usually closer to the maximum. Also, the range of the forecast increases when the difference between the maximum and average velocity raises, providing a tool to predict variability in the wind field.

References

1. Barkmeijer J, Bouttier F, Gijzen MV (1998) Singular vectors and estimates of the analysis-error covariance metric. Q J Roy Meteorol Soc 124(549):1695–1713. https://doi.org/10.1002/qj.49712454916
2. Bosart LF (1975) Sunya experimental results in forecasting daily temperature and precipitation. Mon Weather Rev 103(11):1013–1020. https://doi.org/10.1175/1520-0493(1975)103<1013:SERIFD>2.0.CO;2
3. Bougeault P, Toth Z, Bishop C, Brown B, Burridge D, Chen DH, Ebert B, Fuentes M, Hamill TM, Mylne K, Nicolau J, Paccagnella T, Park YY, Parsons D, Raoult B, Schuster D, Dias PS, Swinbank R, Takeuchi Y, Tennant W, Wilson L, Worley S (2010) The THORPEX interactive grand global ensemble. Bull Am Meteorol Soc 91(8):1059–1072. https://doi.org/10.1175/2010bams2853.1
4. Buizza R, Houtekamer PL, Pellerin G, Toth Z, Zhu Y, Wei M (2005) A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. Mon Weather Rev 133(5):1076–1097. https://doi.org/10.1175/mwr2905.1

5. Buizza R, Milleer M, Palmer TN (1999) Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. Q J Roy Meteorol Soc 125(560):2887–2908. https://doi.org/10.1002/qj.49712556006
6. Buizza R, Tribbia J, Molteni F, Palmer T (1993) Computation of optimal unstable structures for a numerical weather prediction model. Tellus A 45(5):388–407. https://doi.org/10.1034/j.1600-0870.1993.t01-4-00005.x
7. Candille G (2009) The multiensemble approach: The NAEFS example. Mon Weather Rev 137(5):1655–1665. https://doi.org/10.1175/2008mwr2682.1
8. Diaconescu EP, Laprise R (2012) Singular vectors in atmospheric sciences: A review. Earth Sci Rev 113(3–4):161–175. https://doi.org/10.1016/j.earscirev.2012.05.005
9. Doblas-Reyes FJ, Hagedorn R, Palmer TN (2005) The rationale behind the success of multi-model ensembles in seasonal forecasting—II. calibration and combination. Tellus A 57(3), 234–252. https://doi.org/10.1111/j.1600-0870.2005.00104.x
10. Ehrendorfer M, Tribbia JJ (1997) Optimal prediction of forecast error covariances through singular vectors. J Atmos Sci 54(2):286–313. https://doi.org/10.1175/1520-0469(1997)054<0286:OPOFEC>2.0.CO;2
11. Epstein ES (1969) Stochastic dynamic prediction. Tellus 21(6):739–759. https://doi.org/10.3402/tellusa.v21i6.10143
12. Errico RM, Ehrendorfer M, Raeder K (2001) The spectra of singular values in a regional model. Tellus A 53(3):317–332. https://doi.org/10.1034/j.1600-0870.2001.01199.x
13. Fritsch JM, Hilliker J, Ross J, Vislocky RL (2000) Model consensus. Weather Forecast 15(5):571–582. https://doi.org/10.1175/1520-0434(2000)015<0571:MC>2.0.CO;2
14. Gleeson TA (1966) A causal relation for probabilities in synoptic meteorology. J Appl Meteorol 5(3):365–368. https://doi.org/10.1175/1520-0450(1966)005<0365:ACRFPI>2.0.CO;2
15. Gyakum JR (1986) Experiments in temperature and precipitation forecasting for illinois. Weather Forecast 1(1):77–88. https://doi.org/10.1175/1520-0434(1986)001<0077:EITAPF>2.0.CO;2
16. Hagedorn R, Doblas-Reyes FJ, Palmer TN (2005) The rationale behind the success of multi-model ensembles in seasonal forecasting— i. basic concept. Tellus A 57(3), 219–233 (2005). https://doi.org/10.1111/j.1600-0870.2005.00103.x
17. Krishnamurti TN (1999) Improved weather and seasonal climate forecasts from multi-model superensemble. Science 285(5433):1548–1550. https://doi.org/10.1126/science.285.5433.1548
18. Lacarra JF, Talagrand O (1988) Short-range evolution of small perturbations in a barotropic model. Tellus A Dyn Meteorol Oceanogr 40(2):81–95. https://doi.org/10.3402/tellusa.v40i2.11784
19. Leutbecher M (2005) On ensemble prediction using singular vectors started from forecasts. Mon Weather Rev 133(10):3038–3046. https://doi.org/10.1175/MWR3018.1
20. Lewis JM (2005) Roots of ensemble forecasting. Mon Weather Rev 133(7):1865–1885. https://doi.org/10.1175/mwr2949.1
21. Lorenz EN (1960) Energy and numerical weather prediction. Tellus 12(4):364–373. https://doi.org/10.3402/tellusa.v12i4.9420
22. Lorenz EN (1962) The statistical prediction of solutions of dynamic equations. In: International Symposium on Numerical Weather Prediction in Tokyo. The Meteorological Society of Japan
23. Lorenz EN (1965) A study of the predictability of a 28-variable atmospheric model. Tellus 17(3):321–333. https://doi.org/10.1111/j.2153-3490.1965.tb01424.x
24. Lorenz EN (1995) The essence of chaos. University of Washington Press
25. Magnusson L, Leutbecher M, Klln E (2008) Comparison between singular vectors and breeding vectors as initial perturbations for the ecmwf ensemble prediction system. Monthly Weather Rev **136**(11), 4092–4104 (2008). https://doi.org/10.1175/2008MWR2498.1
26. Murphy JM (1988) The impact of ensemble forecasts on predictability. Q J Royal Meteorol Soc 114(480):463–493. https://doi.org/10.1002/qj.49711448010
27. Palmer TN, Doblas-Reyes FJ, Hagedorn R, Alessandri A, Gualdi S, Andersen U, Feddersen H, Cantelaube P, Terres JM, Davey M, Graham R, Délécluse P, Lazar A, Déqué M, Guérémy

JF, Díez E, Orfila B, Hoshen M, Morse AP, Keenlyside N, Latif M, Maisonnave E, Rogel P, Marletto V, Thomson MC (2004) Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). Bull Am Meteorol Soc 85(6), 853–872. https://doi.org/10.1175/bams-85-6-853

28. Palmer TN, Gelaro R, Barkmeijer J, Buizza R (1998) Singular vectors, metrics, and adaptive observations. J Atmos Sci 55(4):633–653. https://doi.org/10.1175/1520-0469(1998)055<0633:SVMAAO>2.0.CO;2
29. Palmer TN, Zanna L (2013) Singular vectors, predictability and ensemble forecasting for weather and climate. J Phys A Math Theor 46(25):254,018. https://doi.org/10.1088/1751-8113/46/25/254018
30. Pavan V, Doblas-Reyes FJ (2000) Multi-model seasonal hindcasts over the euro-atlantic: skill scores and dynamic features. Clim Dyn 16(8):611–625. https://doi.org/10.1007/s003820000063
31. Rajagopalan B, Lall U, Zebiak SE (2002) Categorical climate forecasts through regularization and optimal combination of multiple gcm ensembles. Mon Weather Rev 130(7):1792–1811. https://doi.org/10.1175/1520-0493(2002)130<1792:CCFTRA>2.0.CO;2
32. Sanders F (1973) Skill in forecasting daily temperature and precipitation: some experimental results. Bull Am Meteorol Soc 54(11):1171–1178. https://doi.org/10.1175/1520-0477(1973)054<1171:SIFDTA>2.0.CO;2
33. Simmons AJ, Hollingsworth A (2002) Some aspects of the improvement in skill of numerical weather prediction. Q J R Meteorol Soc 128(580):647–677. https://doi.org/10.1256/003590002321042135
34. Thompson PD (1985) Prediction of the probable errors of predictions. Monthly Weather Rev 113(2):248–259. https://doi.org/10.1175/1520-0493(1985)113<0248:POTPEO>2.0.CO;2
35. Toth Z, Kalnay E (1993) Ensemble forecasting at nmc: the generation of perturbations. Bull Am Meteorol Soc 74(12):2317–2330. https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>$2.0.CO;2
36. Toth Z, Kalnay E (1997) Ensemble forecasting at ncep and the breeding method. Monthly Weather Rev 125(12):3297–3319. https://doi.org/10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2
37. Trevisan A, Legnani R (1995) Transient error growth and local predictability: a study in the lorenz system. Tellus A Dyn Meteorol Oceanogr 47(1):103–117. https://doi.org/10.3402/tellusa.v47i1.11496
38. Wang X, Bishop CH (2003) A comparison of breeding and ensemble transform kalman filter ensemble forecast schemes. J Atmos Sci 60(9):1140–1158. https://doi.org/10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2

Part III
Solar Radiation Diagnostic and Forecasting

The electrical power generation using solar energy is driven by the fluctuation of solar radiation. This variability is caused by the weather conditions and the seasonal effect, both daily and yearly seasons. The second one is totally predictable but the first one comes from moving clouds and other climatological events. Grid operators need to understand to this unstable behavior to improve this kind of energy integration into power grid and enable an increasing.

The following chapters try to deal with solar radiation using different points of view. Chapter 7 characterizes and models solar radiation temporal and spatial variability. Once solar radiation variability had been discussed the next two Chapters aim to improve solar radiation forecasting. Chapter 8 try to establish a methodology for point forecast solar resource for using different statistical models. The work is focused in hourly forecasting and offers the possibility to improve the forecasting using satellite and NWP data as inputs in the models. On the other hand, Chap. 9 aims to explained different methods and models for probabilistic forecasting. Finally, the last Chapter gives the possibility to develop a Solar Map. The model takes into account solar radiation data, terrain surface conditions and cast shadows using a 2-D adaptive triangles meshes.

Solar Resource Variability

Richard Perez, Philippe Lauret, Marc Perez, Mathieu David, Thomas E. Hoff and Sergey Kivalov

Abstract This chapter aims at characterizing and modeling solar resource variability. It is shown that understanding solar energy variability requires a definition of the temporal and spatial context for which variability is assessed. This research describes a predictable, quantifiable variability-smoothing space–time continuum from a single point to thousands of kilometers and from seconds to days. Implications for solar penetration on the power grid and variability mitigation strategies are also discussed. Models for predicting intra-day or intra-hourly variability as a function of insolation conditions are also depicted.

R. Perez (✉) · S. Kivalov
Atmospheric Sciences Research Center, State University of New York,
251 Fuller Road, Albany, NY 12203, USA
e-mail: rperez@albany.edu

S. Kivalov
e-mail: skivalov@albany.edu

P. Lauret · M. David
Laboratoire de Physique et ingénierie mathématique pour l'énergie,
l'environnement et le bâtiment, Université de La Réunion,
15 avenue René Cassin, 97715 Saint-Denis Cedex, France
e-mail: philippe.lauret@univ-reunion.fr

M. David
e-mail: mathieu.david@univ-reunion.fr

M. Perez · T. E. Hoff
Clean Power Research, 1541 Third Street, Napa, CA 94559, USA
e-mail: MarcP@CleanPower.com

T. E. Hoff
e-mail: tomhoff@cleanpower.com

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*,
Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_7

1 Introduction

The output of solar power plants is driven by weather and by the cycle of days and seasons. This output can vary from zero to full power outside the control of plant operators. Therefore, understanding solar resource's variability is key to optimally integrate it onto electrical grids and enable its high penetration.

Variability has two causes. One is fully predictable. It is the result of the apparent seasonal and daily motion of the sun in the sky and the earth's distance from the sun. The other is less predictable. It is the result of the motion of clouds and weather systems.

In this chapter, we examine how temporal and spatial scales influence variability and how these scales matter to developing effective variability mitigation strategies.

2 Temporal and Spatial Scales

Beginning with an intuitive example for the temporal scale, a single location on a given partly cloudy day will experience a high degree of variability, due to changes in the sun's position and the motion of clouds. However, solar energy cumulated over several days at that same location exhibits less variability. Variability becomes insignificant as the temporal integration increases to 1 year and beyond [12]. Figure 1 illustrates this intuitive example. Similarly in the spatial realm, increasing the solar generation footprint from a single location to a region, and further, to a continent reduces intermittency considerably. Increasing this footprint to the entire planet eliminates it almost entirely (Fig. 2). This spatial integration effect is often referred to as the smoothing effect [38].

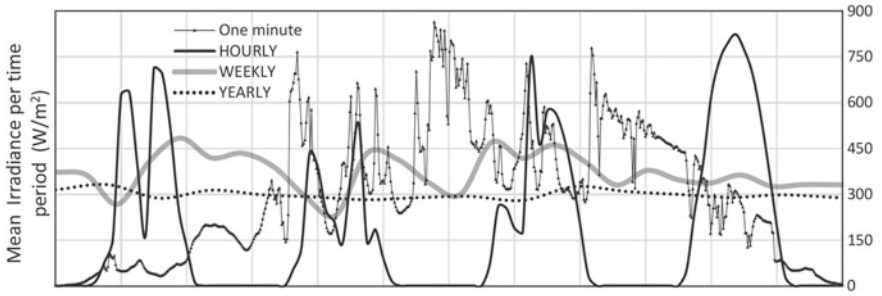


Fig. 1 Comparing the variability of global irradiance time series in a North American location, as a function of integration time. The figure includes 1 day's worth of 1-min data, 4 days worth of hourly data, 26 weeks worth of weekly data, and 16 years worth of yearly integrated data

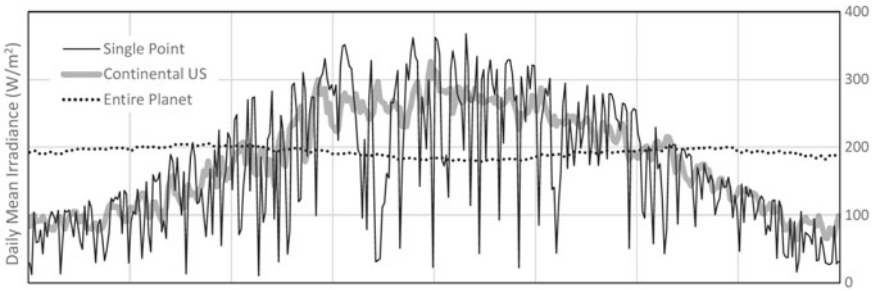


Fig. 2 Comparing the variability of daily global irradiance time series for 1 year as a function of the considered footprint

3 Quantifying Variability

How is variability quantified? An appropriate metric should adapt to a wide range of temporal and spatial scales and embed: (1) the *physical quantity* that varies, (2) the variability *time scale*, and (3) the *time span* over which variability is assessed. In some cases, condition-specific characterizations of variability may be appropriate, e.g., localized variability induced by Cumulus Fields [64].

Physical quantity: For energy producers and grid operators the pertinent quantity is the power output, $\boldsymbol{p}$, of a power plant or of a fleet of power plants at a given point in time. Power output variability reflects the underlying variability of irradiance impinging on the plant(s). Therefore, understanding and quantifying the variability of irradiance amounts to quantifying and understanding the variability of $\boldsymbol{p}$. Irradiance is fully quantified by global horizontal irradiance (*GHI*) and direct normal irradiance (*DNI*). *DNI* variability is relevant for concentrating technologies, while *GHI* variability is representative of flat plate technologies. This chapter focuses on the latter.

Solar geometry-induced variability is fully predictable. Here, we will concentrate on cloud/weather-induced variability that is stochastic in nature, hence not fully predictable. In order to better understand this variability component, it is useful to first remove the solar geometry effects. The *clearness index*, *Kt* (ratio of *GHI* to extraterrestrial irradiance) or the clear sky index, Kt^* (ratio of *GHI* to clear sky *GHI*) both embed the stochastic variability of irradiance, but are largely independent of solar geometry, see Chap. 8. Many authors prefer Kt^* because it more effectively removes solar geometry effects at lower solar elevations (e.g., see [51]) and because it has a more representative zero-to-one range (1 = clear, and 0 = full extinction by thick clouds).

Timescale: The intuitive temporal example presented above suggests that the temporal scale of the selected physical quantity's time series, Δt, is a fundamental factor. Depending on the application, Δt can range from a few seconds to hours and more. The change of Kt^* corresponding to the selected time scale Δt is noted as $\Delta Kt^*_{\Delta t}$. This change is often referred to as the *ramp rate*.

Time span: A proper measure of variability should include ramp events covering a statistically significant *time span*. This time span should be a large multiple of Δt.

Nominal variability metric: *Nominal variability* refers to the variability of the selected dimensionless clear sky index. The maximum or mean $\Delta Kt^*_{\Delta t}$ ramp rate over a given time span has been proposed as such a measure (e.g., see [17]). However, most authors have recently settled on the ramp rate's variance, or its square root, the ramp rate standard deviation, over a given time span as the preferred metric for variability.

$$Nominal\ Variability = \sigma(\Delta Kt^*_{\Delta t}) = \sqrt{Var\left[\Delta Kt^*_{\Delta t}\right]}. \tag{1}$$

Earlier authors, in particular, Skartveit and Olseth [60] had proposed the standard deviation of Kt, $\sigma(Kt)$, rather than $\sigma(\Delta Kt)$ or $\sigma(\Delta Kt^*)$ as a measure of variability. However, $\sigma(Kt)$ is not the most appropriate gauge of variability for a given timescale over a given time span, because it can be driven by one single ramp event—consider, for instance, the case of perfectly clear conditions (i.e., no variability) followed by a one-time change to uniform, heavily overcast conditions (i.e., again, no variability.) In this case, $\sigma(Kt)$ would be the same as if conditions were highly variable and changing from clear to cloudy at every Δt. On the other hand, $\sigma(\Delta Kt^*)$ would capture the difference between the two situations, reflecting low variability in the first case and high variability in the second.

Power output (absolute) variability metric: Equation 1 describes a nominal dimensionless metric. When dealing with power generation, it is necessary to scale up the nominal metric and quantify power variability in absolute terms. This is expressed by Eq. 2:

$$Power\ Variability = \sigma(\Delta \boldsymbol{p}_{\Delta t}) = \sqrt{Var\left[\Delta \boldsymbol{p}_{\Delta t}\right]}. \tag{2}$$

Recall that $\boldsymbol{p}$ can be modeled from Kt^*, via extraction of GHI, extrapolation of plane of array irradiance, and inclusion of PV specs, i.e., without changing the inherent cause of variability. Therefore Eq. 2 does not include additional intrinsic variability information relative to Eq. 1, but only a scaling up from a nominal dimensionless value to a physical value.

4 Variability Mitigation—The Smoothing Effect

When considering a fleet of multiple solar electric installations, the power variability of N plants is given by [18, 48]:

$$Fleet\ Power\ Variability = \sigma\left(\sum_{n=1}^{N} \Delta \boldsymbol{p}^n_{\Delta t}\right) = \sqrt{Var\left[\sum_{n=1}^{N} \Delta \boldsymbol{p}^n_{\Delta t}\right]}. \tag{3}$$

In Eq. 3, $\boldsymbol{p}^{n}_{\Delta t}$ represents the power output time series of the nth plant in the fleet.

Uncorrelated locations: In the limit case where all the plants in the fleet are identical, exhibit the same variability $\sigma(\Delta \boldsymbol{p}_{\Delta t})$, and their power output time series are uncorrelated, Eq. 3 simplifies to

$$Fleet\ Power\ Variability = \sqrt{NVar\left[\Delta \boldsymbol{p}_{\Delta t}\right]} = \sqrt{N}\sigma(\Delta \boldsymbol{p}_{\Delta t}). \tag{4}$$

In this limit case, the *relative variability* defined as the ratio of power variability to installed capacity is given by

$$Fleet\ Relative\ Variability = \frac{\sqrt{N}\sigma(\Delta \boldsymbol{p}_{\Delta t})}{NP_{installed}}. \tag{5}$$

$P_{installed}$ is the installed capacity of each identical plant. Therefore, the relative variability of a fleet of identical power plants with uncorrelated power outputs, but experiencing the same level of individual variability, equals each individual plant's relative variability divided by the square root of the number of plants.

$$Fleet\ Relative\ Variability = \frac{(Single\ Plant\ Relative\ Variability)}{\sqrt{N}}. \tag{6}$$

In nominal terms, Eq. 6 can be applied to quantify the resulting variability of N locations experiencing identically variable, but uncorrelated Kt^{*} time series.

$$\sigma^{N}_{\Delta T} = \frac{\sigma^{1}_{\Delta T}}{\sqrt{N}}, \tag{7}$$

where $\sigma^{N}_{\Delta T}$ is the nominal variability resulting from the ensemble of N uncorrelated locations and $\sigma^{1}_{\Delta T}$ is a single location's nominal variability.

This relative variability reduction as a function of $\sqrt{N}$ underlies the well-known *spatial smoothing effect* noted by many authors, e.g., [35, 38, 66, 68, 69].

General case nonzero correlation: Nearby locations are highly correlated, experiencing the same ramp rates at the same time and varying in sync; in this case a dense fleet of neighboring systems would exhibit nearly the same relative variability as the individual systems. Distant locations' time series are uncorrelated; hence the fleets relative variability is reduced by $\sqrt{N}$. Between these two extremes lie a general case, where fluctuations correlation are partially correlated. Considering a single pair of stations experiencing the same nominal variability $\sigma^{1}_{\Delta T}$, Eq. 7 may be generalized to the case when the two locations are partially correlated, leading to

$$\sigma^{pair}_{\Delta T} = \frac{\sqrt{\rho + 1}}{\sqrt{2}}\sigma^{1}_{\Delta T}, \tag{8}$$

where $\sigma_{\Delta T}^{pair}$ is the nominal variability of the pair and ρ is the correlation between each time series.

How does correlation evolve as a function of location distance and other factors? A considerable amount of work has been devoted to this issue in recent years—e.g., [3–5, 10, 13, 15, 16, 21–23, 26, 29, 31, 33, 36, 40, 50, 53, 54, 58, 59, 62, 67]—leading to the assertion that the correlation of $\Delta Kt^*_{\Delta T}$ time series between two locations depends upon

- The distance, d, between the two locations,
- The considered time scale, ΔT,
- The speed, V, of the variability-inducing clouds/weather systems[1]

The influence of time, speed, and distance had been identified by [17]. They postulated that a dimensionless dispersion factor, D, captures the variability relationship between a single point and a dispersed PV fleet. The dispersion factor is given in Eq. 9 for a homogeneous fleet of systems, where L represents the linear dimension of the fleet in the wind direction.

$$D = \frac{L}{V\,\Delta T}. \tag{9}$$

They identified three possible fleet configurations (see Fig. 3):

1. *A crowded configuration* where the number of systems, N, exceeds the dispersion factor. In this case, the relative variability of the fleet equals the single point's relative variability divided by D.
2. An *optimum configuration* where D equals N and where the fleets variability equals the single point's variability divided by N.
3. A *dispersed configuration* where D is larger than N and where the fleets variability asymptotically tends toward the single point's variability divided by $\sqrt{N}$ as $\frac{D}{N}$ increases.

The dispersion factor model reflects the underlying correlation (or inverse correlation) existing between any two points within the fleet.

The dependence of ρ upon ΔT, d, and V has been empirically inferred from a growing base of experimental observations.

Mills and Wiser [37] analyzed 20 second data from the 32-station ARM network [63]. They observed the exponential decay of station pair correlation as a function of station distance and noted that the rate of decay was a continuous function of the considered time scale. Hoff and Perez [18] used 10-km hourly satellite-derived irradiances over the continental US. They observed a similar asymptotic decay with distance and a predictable dependence of this decay upon ΔT for time intervals of 1, 2, and 3 h. They also noted that the rate of decrease of correlation with distance was different for different US regions and attributed these differences to prevailing

[1]This velocity is a priori defined as the vector in the direction of the two considered locations. However, as will be discussed below, empirical evidence shows that a mean, local—directionless—velocity, can be an adequate input for assessing regional station pair correlations.

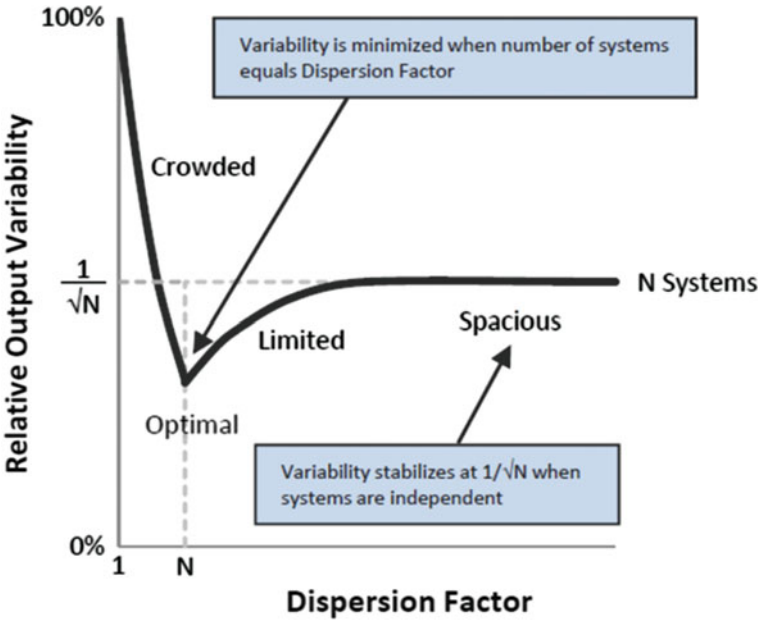


Fig. 3 Relative output variability as a function of the Dispersion Factor for a fleet of N identical PV systems experiencing the same individual variability [17]

regional cloud speeds. Hoff and Norris [15] analyzed high-frequency data (seconds) from a 25-station modular network and confirmed that asymptotic decay with distance was a strong function of ΔT depending on cloud speed that they had estimated independently from satellite-derived cloud motion vectors. They proposed the following relationship between linking distance, time interval, and cloud speed:

$$\rho = \frac{1}{1 + \frac{d}{(\Delta T)(V)}}. \tag{10}$$

Perez et al. [50] analyzed high-resolution high-frequency satellite-derived irradiances (1 km, 1 min) in climatically distinct regions of North America and Hawaii to investigate site-pair correlation decay as a function of distance, timescale, and mean monthly regional cloud speed (see Fig. 4) independently derived from satellite cloud motion vectors. They proposed an alternate formulation for ρ given in Eq. 11:

$$\rho = e^{\frac{d \ln(0.2)}{1.5(\Delta T)(V)}}. \tag{11}$$

Lave and Kleiss [28] and Lave et al. [30] analyzed high-resolution distributed irradiance measurements with a variety of statistical tools such as spectra, coherence spectra, wavelet, correlations, probability density functions, and spatial and temporal

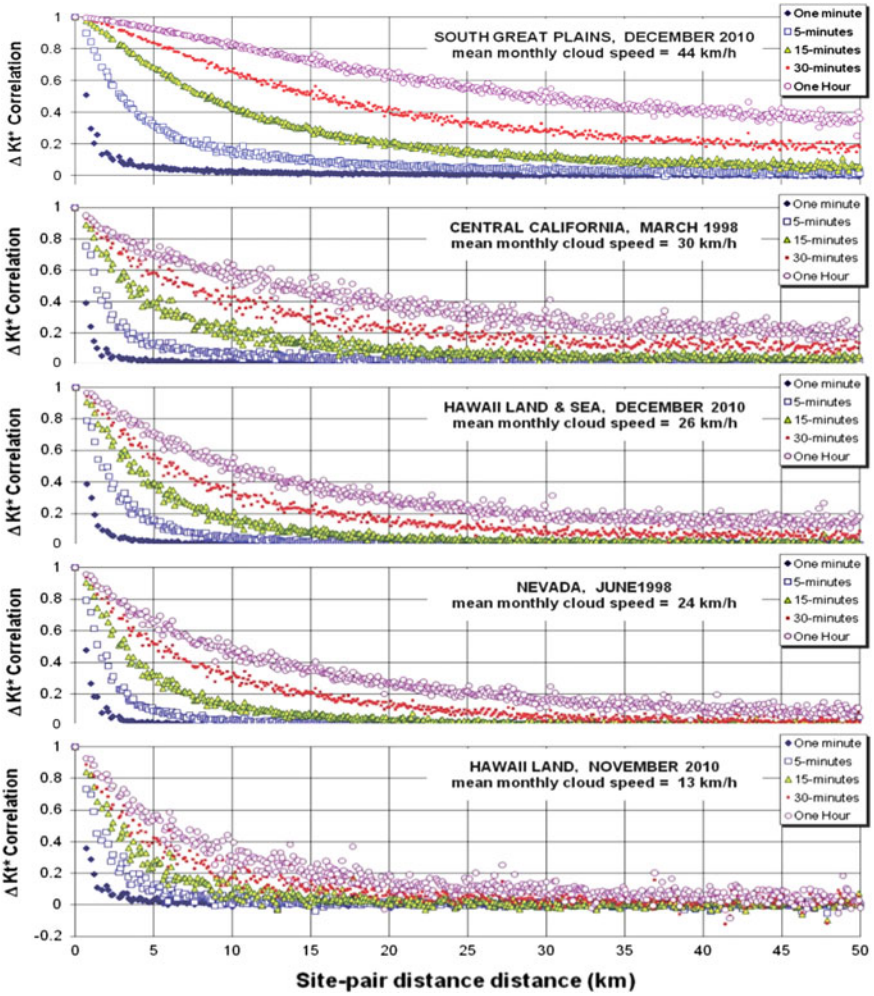


Fig. 4 Site-pair correlation as a function of time period and distance for sample regions in North America and Hawaii. Mean monthly cloud speed was estimated from satellite-derived cloud motion vectors computed for each data point [50]

averaging with the objective of developing a model for simulating the power output of large power plants from single-point measurements. The wavelet variability model (WVM) was then proposed in [31] and it uses wavelet decomposition of the irradiance signal into different timescales (duration of shading from clouds or clouds systems), that were proven to be associated with different amounts of variability reduction. The associated preliminary spatiotemporal correlation function dictates the amount

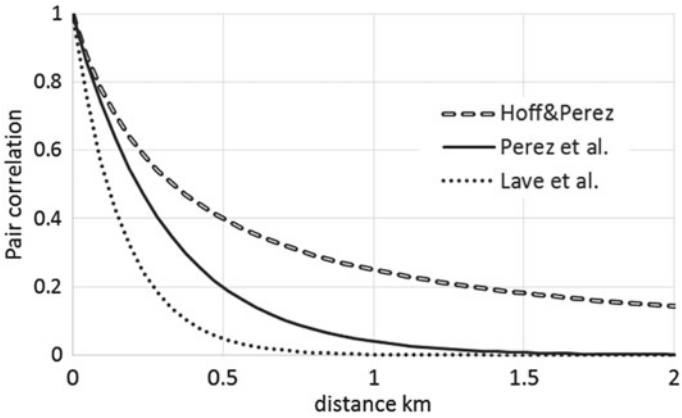


Fig. 5 Comparing correlation decay with distance as formulated in Eqs. 10, 11, and 12 for 1-min data and a cloud speed of 20 km/h

of variability reduction and uses a parameter A that scales the correlation function and that had to be determined from a sensor network collecting high- frequency irradiance data.

$$\rho = exp\left(-\frac{d}{A\bar{t}}\right). \tag{12}$$

Through a virtual cloud model, Lave et al. [29] inferred that the parameter A in Eq. 12 could be approximated to $\frac{1}{2}V$ for station pairs in the direction of the cloud speed.

Figure 5 illustrates and contrasts the formulations in Eqs. 10, 11, and 12 for an example with a timescale of 1 min and a cloud speed of 20 km/h. Note that the difference between Eqs. 11 and 12 may be traceable to the fact that V represents a monthly prevailing cloud speed in the first case and a time-coincident cloud speed in the second, further noting that Eq. 10 was derived empirically without consideration whether pairs were located along or across wind direction.

Formulations such as Eq. 10 or 11 that define cloud speed in the direction of a station pair, do not explain the variability and correlation reduction with distance that is nevertheless observed when speed is zero, i.e., in crosswind directions, e.g., see [14, 32]. As an attempt to describe correlation anisotropy with respect to cloud speed, Arias-Castro et al. [2] applied a kinematic-stochastic model based upon given cloud cover fraction λ_n, cloud size r, stream-wise and cross-stream distance, cloud speed, and time difference. Through dimensional analysis, the correlation functions were expressed through just 4 independent variables: cloud cover fraction, the along-wind and crosswind distance normalized by cloud diameter, and the distance of cloud motion within the ramp interval Δt relative to the cloud diameter (see [2] for details regarding the correlation functions).

Further, David et al. [8] noted that the three key factors governing the correlation decay—timescale, cloud speed, and distance—are not fully independent variables. David et al. [8] analyzed station pair correlations from an irradiance measurement

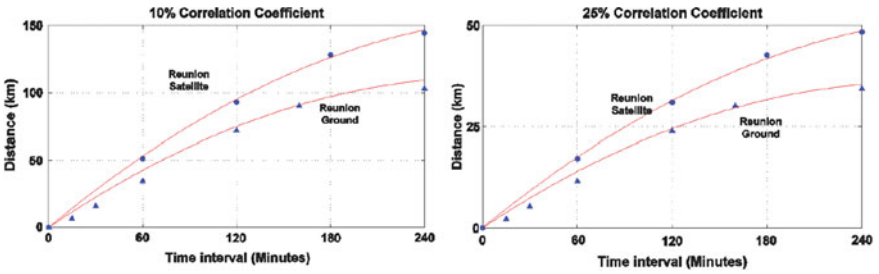


Fig. 6 Mean distance to reach a 10% (left) and a 25% (right) correlation threshold as a function of the time interval of observations in La Reunion (ground and satellite) [8]

network in the Island of La Reunion. Whereas exponential formulation in Eqs. 10 and 11 would imply that, for a given cloud speed, a linear relationship should exist between the distance at a given correlation level and timescale, with the slope depending on cloud speed [50], they observed that the time scale versus distance slope tended to diminish as a distance increased (Fig. 6). This dependence can be explained by the fact that cloud speed evolves as a function of the considered spatial and temporal scales. Indeed for the smallest scales, the variability drivers are cloud substructures. As the spatial scale increases, the drivers become cloud fields, and further, entire weather systems. Because the speed of these drivers is known to decrease with scale, the observed relationships are nonlinear. The cloud speed dependence upon time scale becomes fully apparent when considering very large spatiotemporal scales. Perez and Fthenakis [44, 45] analyzed millions of possible pair correlations from the NASA SSE data set [61] for the entire planet. Figure 7 compares the observed exponential correlation decay for ΔT of 1 and 7 days respectively. It is remarkable that these results are fully consistent with the much lower spatiotemporal scales shown in Fig. 7. This observation could be an expression of the underlying self-similar (fractal) nature of clouds and cloud systems at all scales [34]. At these large scales, as for smaller scales, decorrelation distances are a function of prevailing cloud/weather system speed. It is helpful to compare East–West pairs and North–South pairs (Fig. 8): decorrelation distances are considerably shorter for the latter. A likely explanation is the fact that weather systems tend to move in East–West directions and have limited, slower, North–South motion. Finally, as noted in Fig. 6 for small scales, it is also apparent that the cloud system velocity underlying variability decreases with timescale. For instance, cloud system speeds inferred from Fig. 8 using Eq. 10 would indicate that the speed of the East–West weather system for daily ΔT is of the order of 20 km/h. For ΔT of 7 days the prevailing of systems of weather systems is of the order of 8 km/h.

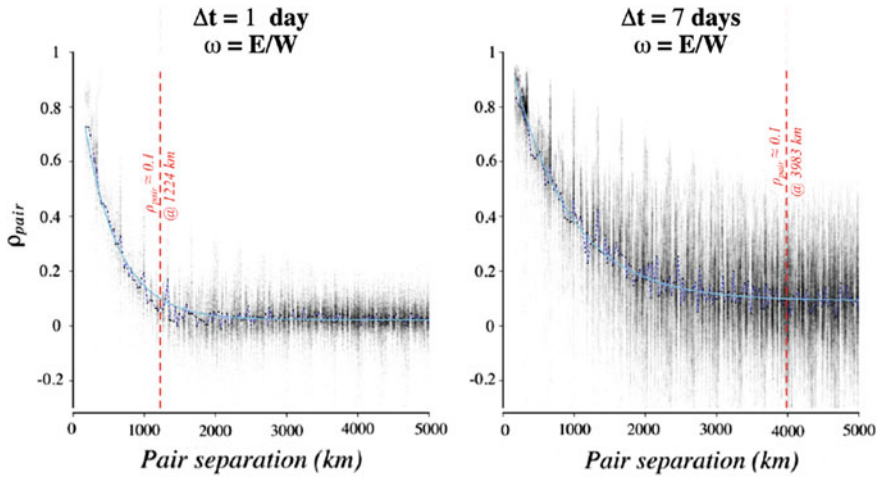


Fig. 7 Site-pair correlation as a function distance for daily and weekly time periods. Station pairs are selected to have a predominantly East–West orientation [45]

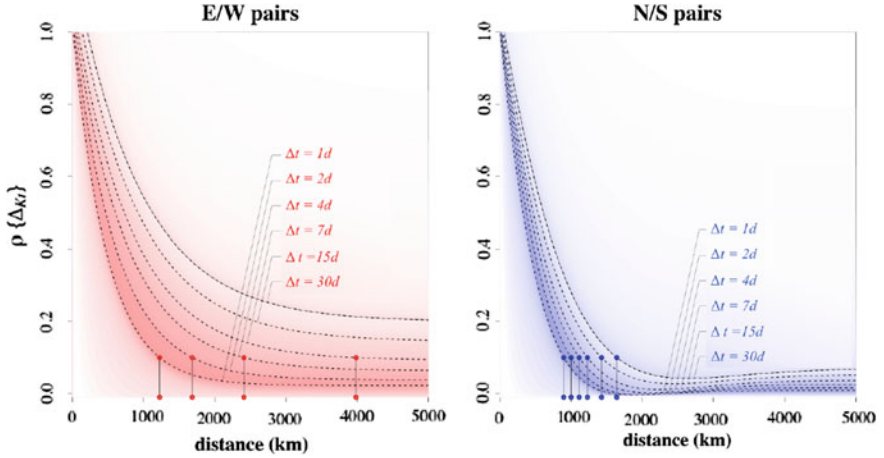


Fig. 8 Impact of prevailing cloud speed on correlation decay for time periods ranging from 1 to 30 days contrasting East–West pairs (prevailing direction of weather systems) and North–South pairs [45]

5 Variability as a Function of Insolation Conditions

Predicting intra-day or intra-hourly variability: Another elementary characteristic of variability is its dependence upon insolation conditions. Intuitively, we know that clear sky conditions will exhibit very little nominal variability, while partly cloudy conditions are inherently variable. Indeed, empirical evidence shows that the prevail-

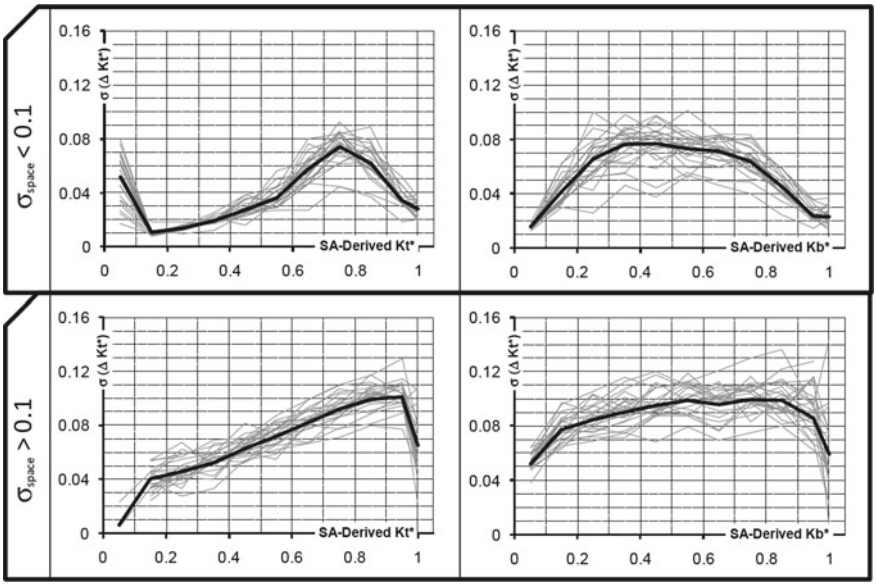


Fig. 9 1-minute $\sigma\left|\Delta Kt^*\right|$ as a function of hourly Kt^*, Kb^* and σ_{space}. Each gray line represents an individual site. The bold black line represents the mean trend derived for all sites [53]

ing insolation conditions quantified by the clear sky index over a given time interval are good indicators of the variability within this time interval. In Figs. 9 and 10, we show two examples relating respectively (a) intra-hour variability ($\Delta t = 1$ minute) and the clear sky index for that hour [53], and (b) intra-day variability ($\Delta t = 1$ hour) and the clear sky index for that day [27]. These observations have led to the development of simple empirical models to infer short-term variability when knowing overall conditions over a given time span. Skartveit and Olseth [60] pioneered this approach in Task 9 of the IEA, by proposing a model to produce 1 min data from hourly data. More recently, Perez et al. [53] proposed a similar parameterization that also takes into account the spatial hourly variability accessible from gridded satellite data. Lauret et al. [27] introduced an approach to generate hourly time series from a day's clear sky index.

These simple models are useful because high-frequency (e.g., 1-minute) measured data are not as commonly available than lower frequency (e.g., hourly) data. The use of daily inputs to produce hourly variability may also be useful in specific applications such as estimating variability for day-ahead forecasts. Indeed, Numerical Weather prediction (NWP) models tend to underestimate short-term variability to deliver best overall performance. Hour-to-hour variability may be inferred from these forecasts by applying a day-to-hour decomposition model such as described by [27] and based on Fig. 10 's observations. It is also noteworthy to mention that the models developed in the 1980s to produce synthetic time series data [1, 46], for e.g., the European

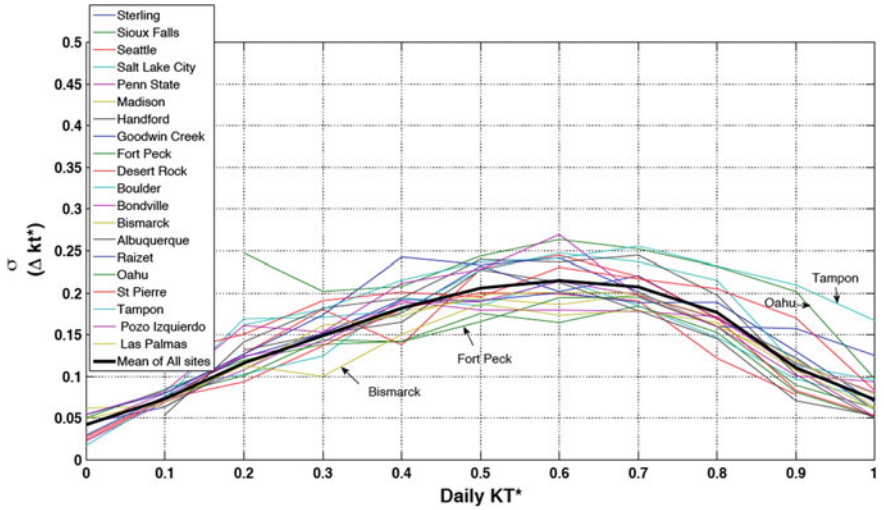


Fig. 10 Hourly $\sigma(\Delta Kt^*)$ as a function of daily KT^*. Each thin line represents an individual site. The bold black line represents the mean trend derived for all sites [27]

solar radiation database [9] indirectly exploit the relationship that inherently exists between hour-to-hour, or day-to-day variability and long-term insolation conditions respectively daily, and monthly.

6 Applied Tools and Models for Power Grid Management

Understanding the fundamentals of solar resource variability's spatiotemporal characteristics have led to the development of methodologies for addressing operational solar power generation issues. One issue, in particular, is the prediction of the variability and ramp rates of spatially extended PV fleets from a limited number of input data points.

Predicting variability of an extended source from a single or multiple measurement points: We present examples of direct application of the spatial/temporal variability correlations presented above. An approach depicted in Fig. 11 and developed by [31] applies the wavelet analysis to decompose the irradiance signal into different timescales to simulate a power plant's output given (1) a spatiotemporal correlation function (e.g., from Eq. 12, (2)) measurements from a single irradiance sensor, (3) the power plant footprint and PV density, (4) a time and location-dependent scaling parameter (parameter A in Eq. 12). The WVM uses these inputs to estimate the variability ratio over the area of the plant. The simulated power plant may have any density of PV coverage: it may be distributed generation with low PV density

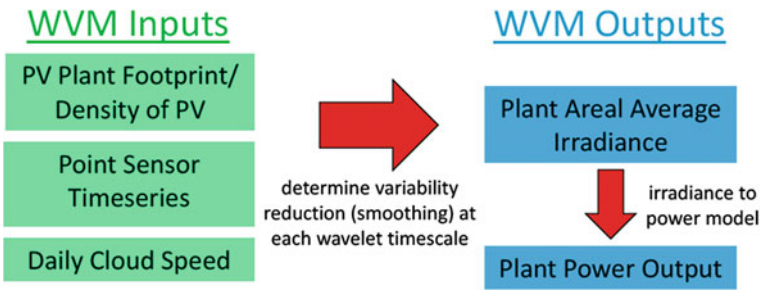


Fig. 11 Wavelet variability model (WVM) for modeling reduction in PV power output variability through geographic smoothing [31]

(i.e., a neighborhood with rooftop PV), centrally located PV as in a utility-scale power plant with high PV density, or any combination thereof.

Another operational approach proposed by [18] is based upon expressing Eq. 3 as the sum of the covariance of all possible plant pair combinations in a PV fleet.

$$\sigma_{\Delta T}^{fleet} = \sqrt{Var\left[\sum_{n=1}^{N} \Delta \boldsymbol{p}_{\Delta t}^{n}\right]} = \sqrt{Var\left[\sum_{i=1}^{N}\sum_{j=1}^{N} COV\left(\Delta \boldsymbol{p}_{\Delta t}^{i}, \Delta \boldsymbol{p}_{\Delta t}^{j}\right)\right]} \tag{13}$$

The covariance between any two plants equals the standard deviations of each of the location's times the correlation coefficient between the two locations (i.e. $COV\left(\Delta \boldsymbol{p}_{\Delta t}^{i}, \Delta \boldsymbol{p}_{\Delta t}^{j}\right) = \sigma_{\Delta T}^{i}\sigma_{\Delta T}^{j}\rho_{\Delta T}^{i,j}$. Therefore, the standard deviation of the changes in fleet output can be defined entirely by the standard deviation of the change in plant output at each location and the correlation between the locations (obtained e.g., from Eq. 10). This method can be applied by deriving nominal variability from a small sample of instrumented power plant and assuming that sampled variability is representative of nearby locations. Kato et al. [24] proposed a comparable approach to determine fluctuation of high-penetration photovoltaic power generation systems dispersed over a large area. His approach is known as the representative blocks method. PV is distributed over a number of subgroups each consisting of N blocks with a given installed capacity and a given variability. The size of each block is set such that block-to-block correlation is negligible and a form of Eq. 5, accounting for different system sizes, may be applied to aggregate blocks and determined the variability of the ensemble.

Inferring spatial variability from high- resolution gridded data sources: Recent years have seen the development of instruments and models capable of producing accurate irradiance data on extended spatial grids. Sky imagers can yield high spatial and temporal resolution data on footprints approaching 100 km^2 (e.g., [39]). Satellite models (e.g., [55]) can produce gridded data for entire continents. The spatial and temporal resolution of satellites is lower than sky imagers' but now approach a few 100 m and a few minutes for the most recent geostationary satellites. With this type

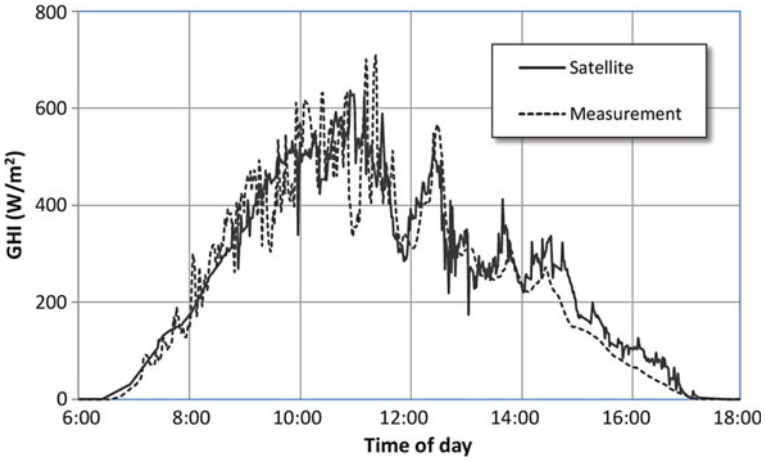


Fig. 12 Comparing 1-minute satellite-derived and measured global irradiance

of input, and within their domain of applications, the problem of applying models to infer spatially extended variability does not pose itself since the extended variability information is inherent the gridded data. In addition, the methodologies developed to understand variability have contributed to the enhancement of these instruments and models. The sky imager functionality has been greatly enhanced [65], to produce high- frequency (seconds) locally gridded irradiance data (5 km radius) and cloud speed can be measured as input to variability models [6, 7, 11]. Satellite-to-irradiance models (see Fig. 12) were also enhanced to produce high-resolution (1 km) data with a timescale approaching 1-minute [50]. In addition to producing the experimental data that led to a better understanding of variability, these new capabilities, in particular, the satellite capability, have also led to a direct, massive approach of PV fleet simulation to directly evaluate and manage variability issues for timescales in excess of a few minutes and spatial scales of a few kilometers, by directly simulating any dispersed PV fleets from satellite-derived irradiance time series [57]. In addition, combining the high-resolution satellite data capability with an understanding and parameterization of underlying variability and localized spatiotemporal correlations could lead to improved short-term solar forecasts under high variability (partly cloudy) conditions [72].

Other approaches, e.g., spatiotemporal kriging: It is sometimes useful to describe time site-specific situations (decorrelation trends, directional effects traceable to localized conditions). Kriging offers the means of doing so from a limited number of measured data points [25]. This method is useful for spatial and temporal scale below what can be achieved by remote sensing today, where a deterministic use of high-resolution gridded data would be a straightforward way to proceed, also noting both spatial and temporal resolution of remote sensing will be improving in the future. A detailed description of spatiotemporal kriging applied to solar resource variability determination is available in [47].

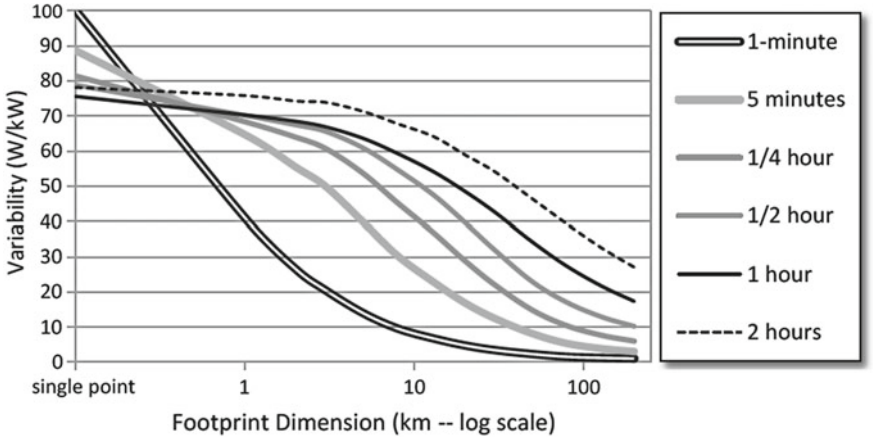


Fig. 13 Nominal variability of a 1 kW power plant as a function of its footprint [49]

Solar resource variability and power grid management: The spatiotemporal characteristics of solar resource variability have direct relevance to PV power generation and its impact on the power grid. These characteristics should be front-and-center in the discussion shaping and optimizing future grid operation and regulations in an inevitable high PV penetration paradigm. This is in full agreement with the conclusions of the International Energy Agency's PVPS Task 14 conclusions [56].

Observations and models describe a space–time continuum underlying the smoothing effect of solar resource variability: shortest term variability matters for the smallest spatial scales while the minimum relevant timescale gradually increases with the size of the considered footprint. This is illustrated in Fig. 13 where the absolute power variability of a nominal 1kW PV power plant (from Eq. 2) is plotted as a function of the resource's footprint from a single point up to 200 $\times$ 200 km [49].

The solar generation footprint and timescale should, therefore, be the primary concerns of grid operators as they pose different load management challenges and imply different solutions: for single distribution systems and large centralized plants, 1-minute fluctuations are relevant as they may create voltage control issues. For grid balancing areas including both fleets of large and small distributed systems, variability effects below 30 min should be of no concern, while hourly and above timescales remain relevant.

Likewise variability mitigation solutions should reflect the solar resource time–space context.

- Up to a few tens of meters small and medium PV installations—ramp rates of the order of seconds are relevant in particular over-irradiance issues, where ramps can exceed power ratings by up to 50% [41, 71] and can create voltage control issues at interconnection points. These are generally passively mitigated by the installations' hardware that curtails excess spikes. For very large systems, buffering via capacitors may be warranted to minimize curtailment losses.

- From hundreds of meters to a few km—distribution feeders and large PV plants—minute ramps are relevant, with impact on distribution system voltage, or transmission system voltage for large centralized power plants. In the latter case, active output buffering via capacitor or battery storage could be considered some utilities impose maximum allowable 1-minute ramp rate requirements [42]. However, these should only be warranted for very large plants or very dense PV fleets where solar production is of the order of the base demand energy flow on the local power grid. For most distributed systems, as long as penetration remains reasonable, experience shows that the ramping noise induced by PV systems on distribution grids is less than the background demand-side ramp noise that utilities have been accustomed to handle for a long time [19, 20]. For very large dispersed penetration—exceeding local demand—grid management would be similar to a centralized plant case and would require buffering.
- From 5 to 20 Km substations, cities—1 min variability vanishes while 10 min and longer ramps remain. Depending on penetration, local regulation via storage may be needed.
- For 50 km large cities and dense transmission networks—15–30 min fluctuations and above are still a concern. Solutions include contingency stand-by generation, storage, or load management in order to react to ramps and ensure a balance between supply and demand—note that these solutions need not be collocated with PV installations.
- For 100's km regional transmission organization's balancing areas fluctuations of less than 1 h should not be of concern. Variability mitigation at these scales can be effectively handled by an optimized basket of active generation, storage, load management, PV output curtailment, and increased interconnection bandwidth [43].

For all temporal and spatial scales where active variability mitigation would be needed, it has been shown that solar forecasting could substantially reduce mitigation measures and operational cost [49]. For small centralized scales, minutes-ahead forecasts could be obtained from sky imaging sensors [70], while for a few km and more, satellite-derived (1–2 h ahead) and numerical weather prediction forecasts (5+ h ahead) would be warranted [52].

7 Conclusions

This chapter presented observations and results assembled from a large body of work from recent years. The underlying narrative of this body of work is the remarkable continuity and self-similarity of the relationships linking spatial footprint and temporal variability across a very large range of spatial and temporal scales, from seconds to days and from meters to thousands of kilometers. The long-observed smoothing effect implying that variability is mitigated over space can be appropriately quanti-

fied from the knowledge of (1) the temporal scale, (2) the spatial scale, and (3) the motion of clouds or cloud systems.

This implies that a proper understanding of how solar energy's resource variability impacts energy systems require a definition of both temporal and spatial contexts for which variability is assessed. The shortest relevant timescale for which variability should be assessed is a direct function of the considered solar generation footprint, and, to a lesser extent, the speed of the clouds/weather systems inducing variability. When developing strategies to mitigate the impact of solar variability on electric power grids, defining this context is critical: indeed, whether 1-minute ramps would be relevant for large centralized plants injecting solar kWh on the grid, focusing on these short-term ramps would be irrelevant for fleets of small and medium power plants distributed over a utility service area, where mitigation occurs naturally.

The understanding of underlying variability structures has led to development of methodologies and models capable of extrapolating the resulting variability on arbitrarily defined spatial footprints (e.g., an ensemble of power plants) from a small sample of point measurements. These models fill a gap in high temporal and spatial scales, improving the temporal/spatial resolution of gridded solar data resources where all relevant variability information would be included in the data themselves without the need for models. The current limit of gridded data sources is of the order of a few minutes on the temporal scale and a few km on the spatial scale. New satellite-derived data are pushing this operational boundary towards finer spatial and temporal resolutions.

Acknowledgement This chapter includes material originally developed for articles written in collaboration with Jan Kleissl of University of San Diego, California.

References

1. Aguiar R, Collares-Pereira M, Conde J (1988) Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices. Sol Energy 40(3):269–279. https://doi.org/10.1016/0038-092X(88)90049-7
2. Arias-Castro E, Kleissl J, Lave, M (2014) A Poisson model for anisotropic solar ramp rate correlations. Sol Energy 101:192–202. https://doi.org/10.1016/j.solener.2013.12.028. http://linkinghub.elsevier.com/retrieve/pii/S0038092X13005549
3. Badosa J, Haeffelin M, Chepfer H (2013) Scales of spatial and temporal variation of solar irradiance on Reunion tropical island. Sol Energy 88:42–56. https://doi.org/10.1016/j.solener.2012.11.007. http://linkinghub.elsevier.com/retrieve/pii/S0038092X12003982
4. Becker S, Frew BA, Andresen GB, Zeyer T, Schramm S, Greiner M, Jacobson MZ (2014) Features of a fully renewable US electricity system: optimized mixes of wind and solar PV and transmission grid extensions. Energy 72:443–458. https://doi.org/10.1016/j.energy.2014.05.067. http://linkinghub.elsevier.com/retrieve/pii/S0360544214006343
5. Bing J, Krishnani P, Bartholomy O, Hoff T, Perez R (2012) Solar monitoring, forecasting, and variability assessment at SMUD, Denver, CO
6. Bosch J, Kleissl J (2013) Cloud motion vectors from a network of ground sensors in a solar power plant. Sol Energy 95:13–20. https://doi.org/10.1016/j.solener.2013.05.027. http://linkinghub.elsevier.com/retrieve/pii/S0038092X13002193

7. Bosch J, Zheng Y, Kleissl J (2013) Deriving cloud velocity from an array of solar radiation measurements. Sol Energy 87:196–203. https://doi.org/10.1016/j.solener.2012.10.020. http://linkinghub.elsevier.com/retrieve/pii/S0038092X12003854
8. David M, Andriamasomanana FHR, Liandrat O (2014) Spatial and temporal variability of PV output in an insular grid: case of Reunion Island. Energy Procedia 57:1275–1282. https://doi.org/10.1016/j.egypro.2014.10.117. http://linkinghub.elsevier.com/retrieve/pii/S1876610214014842
9. ESRA (1999) ESRA, European solar radiation atlas. https://hal.archives-ouvertes.fr/hal-00363667/document
10. Frank J, Freedman J, Brower M, Schnitzer M (2011) Development of high frequency solar data, Raleigh, NC
11. Fung V, Bosch JL, Roberts SW, Kleissl J (2014) Cloud shadow speed sensor. Atmos Meas Tech 7(6):1693–1700. https://doi.org/10.5194/amt-7-1693-2014. http://www.atmos-meas-tech.net/7/1693/2014/
12. Gueymard CA, Wilcox SM (2011) Assessment of spatial and temporal variability in the US solar resource from radiometric measurements and predictions from models using ground-based or satellite data. Sol Energy 85(5):1068–1084. https://doi.org/10.1016/j.solener.2011.02.030. http://linkinghub.elsevier.com/retrieve/pii/S0038092X11000855
13. Halsz G, Malachi Y (2014) Solar energy from Negev desert, Israel: assessment of power fluctuations for future PV fleet. Energy Sustain Dev 21:20–29. https://doi.org/10.1016/j.esd.2014.04.005. http://linkinghub.elsevier.com/retrieve/pii/S0973082614000404
14. Hinkelmann L (2013) Differences between along-wind and cross-wind solar variability. Proc Sol Energy:192–203
15. Hoff T, Norris B (2010) Mobile high-density: irradiance sensor network: cordelia junction results
16. Hoff T, Perez R (2012) Predicting short-term variability of high-penetration PV, Denver, CO
17. Hoff TE, Perez R (2010) Quantifying PV power output variability. Sol Energy 84(10):1782–1793. https://doi.org/10.1016/j.solener.2010.07.003. http://linkinghub.elsevier.com/retrieve/pii/S0038092X10002380
18. Hoff TE, Perez R (2012) Modeling PV fleet output variability. Sol Energy 86(8):2177–2189. https://doi.org/10.1016/j.solener.2011.11.005. http://linkinghub.elsevier.com/retrieve/pii/S0038092X11004154
19. Holger R, Schroedter-Homscheidt M, Beyer H, Meier F, Heilscher G (2014) Wolkenindikatoren fr die Sammelschienenspannung von Niederspannungsverteilnetztransformatoren, 29. In: Symposium Photovoltaische Solarenergie, Kloster Banz, Bad Staffelstein
20. Holger R, Schroedter-Homscheidt M, Meier F, Heilscher G (2014) Application of meteorological data for state estimation of an electrical low voltage grid with a high amount of photovoltaic systems. In: 14th EMS annual meeting and 10th European conference on applied climatology, Prague
21. Huang J, Troccoli A, Coppin P (2014) An analytical comparison of four approaches to modelling the daily variability of solar irradiance using meteorological records. Renew Energy 72:195–202. https://doi.org/10.1016/j.renene.2014.07.015. http://linkinghub.elsevier.com/retrieve/pii/S0960148114004017
22. Jamaly M, Bosch JL, Kleissl J (2013) Aggregate ramp rates of distributed photovoltaic systems in San Diego County. IEEE Trans Sustain Energy 4(2):519–526. https://doi.org/10.1109/TSTE.2012.2201966. http://ieeexplore.ieee.org/document/6232480/
23. Kankiewicz A, Sengupta M, Li J (2011) Cloud meteorology and utility scale variability, Raleigh, NC
24. Kato T, Inoue T, Suzuki Y (2011) Estimation of total power output fluctuation of high penetration photovoltaic power generation system, Detroit, Michigan, USA
25. Krige D (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand. J Chem Metall Min Soc South Africa, 119–139
26. Kuszamaul S, Ellis A, Stein J, Johnson L (2010) Lanai high-density irradiance sensor network for characterizing solar resource variability of MW-scale PV system, Honolulu, HI

27. Lauret P, Perez R, Mazorra Aguiar L, Tapachs E, Diagne HM, David M (2016) Characterization of the intraday variability regime of solar irradiation of climatically distinct locations. Sol Energy 125:99–110. https://doi.org/10.1016/j.solener.2015.11.032. http://linkinghub.elsevier.com/retrieve/pii/S0038092X15006490
28. Lave M, Kleissl J (2010) Solar variability of four sites across the state of Colorado. Renew Energy 35(12):2867–2873. https://doi.org/10.1016/j.renene.2010.05.013. http://linkinghub.elsevier.com/retrieve/pii/S0960148110002314
29. Lave M, Kleissl J (2013) Cloud speed impact on solar variability scaling application to the wavelet variability model. Sol Energy 91:11–21. https://doi.org/10.1016/j.solener.2013.01.023. http://linkinghub.elsevier.com/retrieve/pii/S0038092X13000406
30. Lave M, Kleissl J, Arias-Castro E (2012) High-frequency irradiance fluctuations and geographic smoothing. Sol Energy 86(8):2190–2199. https://doi.org/10.1016/j.solener.2011.06.031. http://linkinghub.elsevier.com/retrieve/pii/S0038092X11002611
31. Lave M, Kleissl J, Stein JS (WVM) A wavelet-based variability modelfor solar PV power plants. IEEE Trans Sustain Energy 4(2):501–509 (2013). https://doi.org/10.1109/TSTE.2012.2205716. http://ieeexplore.ieee.org/document/6269913/
32. Lonij VP, Brooks AE, Cronin AD, Leuthold M, Koch K (2013) Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors. Sol Energy 97:58–66. https://doi.org/10.1016/j.solener.2013.08.002. http://linkinghub.elsevier.com/retrieve/pii/S0038092X13003125
33. Lorenz E, Scheidsteger T, Hurka J, Heinemann D, Kurz C (2011) Regional PV power prediction for improved grid integration. Prog Photovolt Res Appl 19(7):757–771. https://doi.org/10.1002/pip.1033. http://doi.wiley.com/10.1002/pip.1033
34. Lovejoy S (1982) Area-perimeter relation for rain and cloud areas. Science, 185–187. New Series
35. Marcos J, Marroyo L, Lorenzo E, Garca M (2012) Smoothing of PV power fluctuations by geographical dispersion: smoothing of PV power fluctuations. Prog Photovolt: Res Appl 20(2):226–237. https://doi.org/10.1002/pip.1127. http://doi.wiley.com/10.1002/pip.1127
36. Mazumdar BM, Saquib M, Das, AK (2014) An empirical model for ramp analysis of utility-scale solar PV power. Sol Energy 107:44–49. https://doi.org/10.1016/j.solener.2014.05.027. http://linkinghub.elsevier.com/retrieve/pii/S0038092X14002564
37. Mills A, Wiser R (2010) Implications of wide-area geographic diversity for short-term variability of solar power. Technical report, Lawrence Berkeley National Laboratory Technical Report LBNL-3884E
38. Murata A, Yamaguchi H, Otani K (2009) A method of estimating the output fluctuation of many photovoltaic power generation systems dispersed in a wide area. Electr Eng Japan 166(4):9–19
39. Nguyen A, Velay M, Schoene J, Zheglov V, Kurtz B, Murray K, Torre B, Kleissl J (2016) High PV penetration impacts on five local distribution networks using high resolution solar resource assessment with sky imager and quasi-steady state distribution system simulations. Sol Energy 132:221–235. https://doi.org/10.1016/j.solener.2016.03.019. http://linkinghub.elsevier.com/retrieve/pii/S0038092X16002061
40. Norris B, Hoff T (2011) Determining storage reserves for regulating solar variability
41. Ole-Morten M, Yordanov G, Ranaweera I (2014) Short-term intermittency of solar irradiance in Southern Norway, pp 2635–2638
42. PEA: PREPA puerto rico electric power authority (2013). www.prepa.com
43. Perez M (2015) Geographic dispersion and curtailment of VLS-PV electricity. Chapter 4 future technical options for the entire energy system, 2015. Technical report
44. Perez M, Fthenakis V (2012) Quantifying long time scale solar resource variability, Denver, CO
45. Perez MJ, Fthenakis VM (2015) On the spatial decorrelation of stochastic solar resource variability at long timescales. Sol Energy 117:46–58. https://doi.org/10.1016/j.solener.2015.04.020. http://linkinghub.elsevier.com/retrieve/pii/S0038092X1500208X
46. Perez R, Aguiar R, Collares-Pereira M, Dumortier D, Estrada-Cajigal V, Gueymard C, Ineichen P, Littlefair P, Lund H, Michalsky J, Olseth J, Renn D, Rymes M, Skartveit A, Vignola F, Zelenka

A (2001) Solar resource assessment a review. Solar energy-the state of the art. James & James, London, pp 497–575
47. Perez R, David M, Hoff TE, Jamaly M, Kivalov S, Kleissl J, Lauret P, Perez M (2016) Spatial and temporal variability of solar energy. Found Trends Renew Energy 1(1):1–44. https://doi.org/10.1561/2700000006. http://www.nowpublishers.com/article/Details/REN-006
48. Perez R, Hoff T (2013) Solar resource variability. In: Kleissl J (ed) Solar resource assessment and forecasting. Elsevier
49. Perez R, Hoff T, Dise J, Chalmers D, Kivalov S (2013) Mitigating short-term PV output variability, Paris, France
50. Perez R, Hoff T, Kivalov S (2011) Spatial & temporal characteristics of solar radiation variability, Kassel, Germany
51. Perez R, Ineichen P, Seals R, Zelenka A (1990) Making full use of the clearness index for parameterizing hourly insolation conditions. Sol Energy 45(2):111–114. https://doi.org/10.1016/0038-092X(90)90036-C. http://linkinghub.elsevier.com/retrieve/pii/0038092X9090036C
52. Perez R, Kankiewicz A, Schlemmer J, Hemker K, Kivalov S (2014) A new operational solar resource forecast service for PV fleet simulation
53. Perez R, Kivalov S, Schlemmer J, Hemker K, Hoff T (2011) Parameterization of site-specific short-term irradiance variability. Sol Energy 85(7):1343–1353. https://doi.org/10.1016/j.solener.2011.03.016. http://linkinghub.elsevier.com/retrieve/pii/S0038092X11000995
54. Perez R, Kivalov S, Schlemmer J, Hemker K, Hoff TE (2012) Short-term irradiance variability: preliminary estimation of station pair correlation as a function of distance. Sol Energy 86(8):2170–2176. https://doi.org/10.1016/j.solener.2012.02.027. http://linkinghub.elsevier.com/retrieve/pii/S0038092X12000928
55. Perez R, Schlemmer J, Hemker K, Kivalov S, Kankiewicz A, Gueymard C (2015) Satellite-to-irradiance modeling a new version of the SUNY model
56. Remund J, Calhau C, Marcel D, Perret L (2015) Spatio-temporal variability of PV production. In: Spatio-temporal variability of PV production. Proceedings of the European photovoltaic solar energy conference and exhibition (EUPVSEC), Hamburg, Germany
57. Research CP, Behind-the-meter intelligence for distributed PV grid integration (2012). http://www.cleanpower.com/wp-content/uploads/SA-FleetView-Whitepaper-v060812.pdf
58. Rowlands IH, Kemery BP, Beausoleil-Morrison I (2014) Managing solar-PV variability with geographical dispersion: An Ontarioa (Canada) case-study. Renew Energy 68:171–180. https://doi.org/10.1016/j.renene.2014.01.034. http://linkinghub.elsevier.com/retrieve/pii/S0960148114000639
59. Sengupta M (2011) Measurement and modeling of solar and PV output variability, Raleigh, NC
60. Skartveit A, Olseth J (1992) The probability density and autocorrelation of short-term global and beam irradiance. Sol Energy 49(6):477–487. https://doi.org/10.1016/0038-092X(92)90155-4. http://linkinghub.elsevier.com/retrieve/pii/0038092X92901554
61. SSE, N: Surface meteorology and solar energy (2012). http://eosweb.larc.nasa.gov/sse/
62. Stein J, Ellis A, Hansen C, Chadliev V (2011) Simulation of 1-minute power output from utility-scale photovoltaic generation systems, Raleigh, NC
63. Stokes G, Schwartz S (1994) The atmospheric radiation measurement (ARM) program: programmatic background and design of the cloud and radiation test bed. Bull Am Meteorol Soc 1201–1221
64. Tomson T (2010) Fast dynamic processes of solar radiation. Sol Energy 84(2):318–323. https://doi.org/10.1016/j.solener.2009.11.013. http://linkinghub.elsevier.com/retrieve/pii/S0038092X09002795
65. Urquhart B, Kurtz B, Dahlin E, Ghonima M, Shields JE, Kleissl J (2014) Development of a sky imaging system for short-term solar power forecasting. Atmos Meas Tech Discuss 7(5):4859–4907. https://doi.org/10.5194/amtd-7-4859-2014. http://www.atmos-meas-tech-discuss.net/7/4859/2014/
66. Vignola F (2001) Variability of solar radiation over short time intervals, Washington, D.C

67. Vindel J, Polo J (2014) Intermittency and variability of daily solar irradiation. Atmos Res 143:313–327
68. Wiemken E, Beyer H, Heydenreich W, Kiefer K (2001) Power characteristics of PV ensembles: experiences from the combined power production of 100 grid connected PV systems distributed over the area of Germany. Sol Energy 70(6):513–518. https://doi.org/10.1016/S0038-092X(00)00146-8. http://linkinghub.elsevier.com/retrieve/pii/S0038092X00001468
69. Woyte A, Belmans R, Nijs J (2007) Fluctuations in instantaneous clearness index: analysis and statistics. Sol Energy 81(2):195–206. https://doi.org/10.1016/j.solener.2006.03.001. http://linkinghub.elsevier.com/retrieve/pii/S0038092X0600079X
70. Yang D, Dong Z, Reindl T, Jirutitijaroen P, Walsh WM (2014) Solar irradiance forecasting using spatio-temporal empirical kriging and vector autoregressive models with parameter shrinkage. Sol Energy 103:550–562. https://doi.org/10.1016/j.solener.2014.01.024. http://linkinghub.elsevier.com/retrieve/pii/S0038092X14000425
71. Yordanov G, Mitdgard O, Norum L (2013) Overirradiance (cloud enhancement) events at high latitudes. IEEE J Photovolt 78:203–218
72. Zagouras A, Pedro HT, Coimbra CF (2015) On the role of lagged exogenous variables and spatiotemporal correlations in improving the accuracy of solar forecasting methods. Renew Energy 78, 203–218. https://doi.org/10.1016/j.renene.2014.12.071. http://linkinghub.elsevier.com/retrieve/pii/S0960148115000051

Solar Radiation Forecasting with Statistical Models

Luis Mazorra-Aguiar and Felipe Díaz

Abstract Renewable energy electrical generation has experienced significant growth in the recent years. Renewable energies generate electrical energy using different natural resources, such as solar radiation and wind fields. These resources present an unstable behavior because they depend on different meteorological conditions. In order to maintain the balance between input and output electrical energy into the power system, grid operators need to control and predict these fluctuating events. Indeed, forecasting methods are completely necessary to increase the proportion of renewable energies into the system (Heinemann et al. in Forecasting of solar radiation: solar energy resource management for electricity generation from local level to global scale. Nova Science Publishers, New York, 2006 [17], Wittmann et al. in IEEE J Sel Top Appl Earth Obs Remote Sens 1:18–27, 2008 [46]). Reducing the uncertainty of natural resources, operators could reduce maintenance costs, improve the interventions in the intra-day market and optimize management decisions with nonrenewable energies supply. Many forecasting methods are used to obtain solar radiation forecasting for different time horizons. In this chapter, we will focus on several solar radiation forecasting statistical methods for intra-day time horizons using ground and exogenous data as inputs.

1 Introduction

Solar radiation forecasting could be used for different purposes with a wide range of methods. Depending on these purposes, forecasting models are based on different input parameters and used for several time horizons [22, 42].

L. Mazorra-Aguiar (✉) · F. Díaz
University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain
e-mail: luis.mazorra@ulpgc.es

F. Díaz
e-mail: felipe.diaz@ulpgc.es

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_8

- For time horizons less than hour models based on ground-based sky images obtain very good results. These models offer high precision information about cloud cover variability using sky images with 180 cameras [10, 45].
- Satellite image models are considered a very useful tool to improve solar radiation for time horizons up to several hours ahead. Geostationary meteorological satellites obtain images from atmosphere and satellite models estimate solar radiation using these images. In recent years, these models obtain accurate results with temporary resolution less than an hour and spatial resolution around 1–5 km. A review of several satellite models is shown in Sect. 1.2.
- Statistical models obtain accurate results for time horizons up to hours ahead. These models are not good enough to estimate the cloud motion but the high correlation between ground solar radiation data series made them very good tools for solar forecasting over 1 hour. The bibliography offers different statistical models for solar radiation purposes, as autoregressive models (AR) and autoregressive moving average (ARMA) [4, 5], autoregressive-integrated moving average (ARIMA), or several machine learning techniques such as neural networks, support vector machines, or Gaussian process [6, 23, 26].
- For time horizons over 1 day ahead up to 15 days, numerical weather predictions (NWP) models estimate atmosphere conditions and give different meteorological variables as solar radiation. These models are based on physical models using differential equations and solved with numerical methods, see Sect. 1.3.

NWP models accuracy vary depending on the temporal resolution and the geographical area. Different works are presented in bibliography showing almost no deviation for clear sky days [17] and errors around 30–40% for different stations between Europe, U.S.A., and Canada [33–35]. NWP data have also been used in recent years for post-processing forecasting results with hourly ground measurements from 6 h ahead onwards [12]. On the other hand, satellite images could also provide information about cloud variability using cloud motion vectors and improve hourly forecasting [16, 33].

This chapter is focused on solar radiation forecasting for global horizontal irradiance up to 6 h ahead. The statistical models provide good forecasting results for short-time horizons with different temporal granularities (from 5 min to hourly data). Statistical models find a relation between input data and the desired forecast solar radiation data. Many references estimate this relation using past ground solar radiation data for the same time series as inputs. However, in recent years several works have pointed out the improvement obtained combining ground measurement data with exogenous data as inputs [11, 31, 48]. This chapter is intended to provide a procedure to use statistical models for solar radiation forecasting using ground measurements and exogenous data, such as NWP and satellite data. An automatic methodology is proposed for the selection of satellite pixels using Pearson's correlation values.

2 Ground Solar Radiation Data

As statistical models explained in this chapter are based on ground solar radiation measurements, it is important to establish a good quality series of data. Indeed, before applying forecasting models, a solar radiation data assessment and quality check procedure must be used, see Sect. 1.1.2.4.

GHI data series is not considered stationary because they are affected by several variabilities. One variability is completely predictable and it is caused by the annual and daily solar cycle. On the other hand, motion of clouds and atmospheric parameters such as aerosols or water vapor caused a nonpredictable variability. All statistical models suggested in this chapter work with stationary time series of data, so autocorrelation should be constant over the time [9].

To work with statistical models, separating solar geometry dependence from the nondeterministic influences generated by atmospheric phenomena is considered appropriate [13]. So, two different new variables have been introduced to get transformed solar radiation temporal series in stationary series, clearness index k, and clear sky index K_t^*.

Clearness index is calculated dividing the global solar horizontal radiation GHI from measurement data by exoatmospheric horizontal radiation GHI_0 in the same point, see Eq. 1. This index removes deterministic variability caused by solar cycle because exoatmospheric radiation is based on solar angles.

$$K_t^* = \frac{GHI}{GHI_0} \tag{1}$$

GHI_0 is calculated for every day of the year over an horizontal surface with a simple expression using slight variations of distance between the Sun and Earth.

$$GHI_0 = I_0 \varepsilon_0 \cos(\theta_{zs}) \tag{2}$$

$$\begin{aligned} \varepsilon_0 = 1.00011 + 0.034221 \cos\tau + 0.001280 \sin\tau + \\ +0.000719 \cos 2\tau + 0.000077 \sin 2\tau \end{aligned} \tag{3}$$

$$\tau = \frac{2\pi(n-1)}{365} \tag{4}$$

where I_0 represents solar energy received from sun in a specific surface outside of the atmosphere per unit of time. The solar constant is considered normally as $I_0 = 1367\ W/m2$. While ε_0 is the variation of the distance between the Sun and Earth over the year calculated with Eq. 3, and $cos\theta_{zs}$ is zenith angle. Finally, zenith angle equation is substituted in main equation, Eq. 5.

$$GHI_0 = I_0 \varepsilon_0 (\sin\delta \sin\Phi + \cos\delta \cos\Phi \cos\omega) \tag{5}$$

The second variable introduced allows us to remove seasonal and atmospheric variability from solar radiation data series. This index is called **Clear sky index** K_t^* and is widely used in the bibliography for facilitating the learning process of statistical methods. As the clear sky index includes a clear sky model, Eq. 6, some atmospheric conditions are included in this calculation and we obtain a stationary data series.

$$K_t^* = \frac{GHI}{GHI_c} \tag{6}$$

Clear sky models estimate solar radiation GHI_c in a surface taken into account a day without any clouds. Most of the clear sky models are based on different climatic variables that represent the conditions of the atmosphere in clear times, such as aerosol optical depths (AODs), water vapor, ozone, Linke turbidity factor, or pressure. AODs represent solar radiation attenuation for different wavelengths from the scattering and absorption of sunlight within an atmospheric column. AODs and water vapor could be obtained from AERONET measurement stations net [19], while ozone could be retrieved from World Ozone Monitoring Mapping provided by the Canadian Government [8]. MACC project also provides AODs, water vapor, and ozone data for the whole world from 2004, available in [43].

These kinds of models have been tested all over the world and good results were obtained compared with ground measurement for clear sky times [39, 47]. One of the most common clear sky models in solar energy community is Bird and Hulstrm model [2]. This model is easy to implement and use water vapor column in *cm*, two aerosol optical depths, for 380 nm and 500 nm respectively, and total ozone column for the point we are estimating clear sky radiation. Based on these data, Bird model estimates different variables, such as Rayleigh dispersion, absorption of ozone, oxygen, carbon dioxide, and water vapor or absorption and dispersion of aerosols.

Another example widely used in the solar energy field is a method based on the REST2 model [15]. First version of REST, developed by Gueymard, only estimated beam component of solar radiation for clear sky. Later, Gueymard developed REST2 as a dual-band model based on the CPCR2 model. REST2 includes spectral distribution of extraterrestrial radiation, solar constant, water vapor, Angstrom turbidity coefficient, and reduced NO_2 and ozone column as inputs. In [43], is also available data series for global horizontal irradiance (GHI), direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI) calculated with McClear clear sky model [27]. McClear data are available from 2004 to current day $d - 2$ with minute, hourly, daily, or monthly time step for whole world and with a spatial resolution of 1.125. This model is based on look-up tables and radiative transfer model libRadtran using atmospheric composition variables provided by the MACC projects over whole world, such as AOD at 550 and 1240 nm, water vapor, and ozone column.

It is important to evaluate the accuracy of the clear sky model comparing with the measurement data for clear sky days. To evaluate these methods, it is necessary first to find out cloud-free solar radiation times (clear sky). Several methods could be found in the solar energy field to separate clear sky conditions from cloudy sky. Ineichen method detects clear sky hours by establishing a relation between global, beam, and diffuse, studying the stability of clearness index and the broadband aerosol optical depth [21]. In a similar way, Lefevre et al. [27] employ clearness index, corrected clearness index, direct normal clearness index, and diffuse fraction to detect clear sky instants. The method is described for 1 min. data and an adaptation for hourly data is used for Eissa to validate HelioClim-3 database in Egypt [14]. On the other hand, to detect individual times or period of times with clear or cloudy sky conditions only using GHI, Reno, and Hansen, [38] uses a moving window of period of times with 1 min. data series. This methodology detects clear and cloudy sky if data series meets certain conditions based on maximum value of GHI, mean value of GHI, and three different parameters to study the variability of each period. If the period studied in this window meets all the conditions, this period is considered clear sky weather. The limits of each condition should be established experimentally with ground measurement data in each location. Another methodology is also proposed to separate clear and cloudy sky conditions for GHI in periods of time [36, 37]. The model compares hourly data from ground measurement stations and clear sky model to detect whole clear sky days. For each day, the correlation coefficients matrix between ground data and clear sky data estimated by the model is calculated, Eq. 7. The determinant of this matrix should be lower than a threshold established experimentally once the data have been observed.

$$C = \begin{bmatrix} \rho_{GHI,GHI} & \rho_{GHI,GHI_c} \\ \rho_{GHI_c,GHI} & \rho_{GHI_c,GHI_c} \end{bmatrix} = \begin{bmatrix} 1 & \rho_{GHI,GHI_c} \\ \rho_{GHI_c,GHI} & 1 \end{bmatrix} \tag{7}$$

$$\rho_{GHI_c,GHI} = \frac{Cov(GHI_c, GHI)}{\sigma_{GHI_c}\sigma_{GHI}} \tag{8}$$

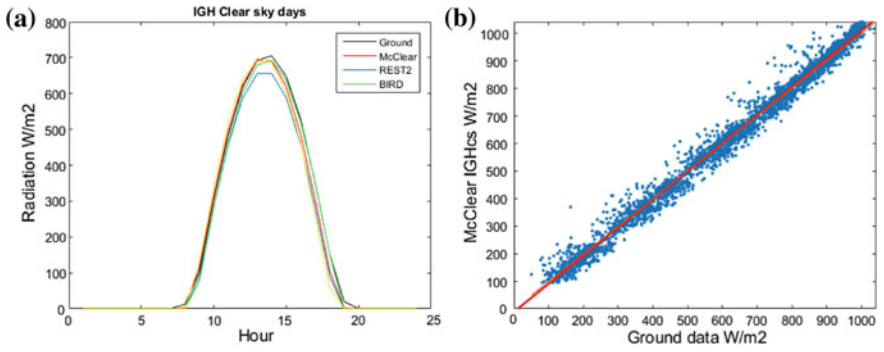


Fig. 1 Hourly GHI estimated with **a** different clear sky models and compared with ground measurement and **b** estimated with McClear model compared with ground measurement in a location in Canary Island, Spain

In Fig. 1a several clear sky models were tested for a single station in Canary Islands, Spain. All clear sky models reproduce very good results in terms of % rRMSE, variating from 4% with McClear model to 8% Bird model. While in Fig. 1b, GHI estimated with McClear compared with ground measurement data obtained in the same location is shown. McClear reproduces accurate results comparing hourly data.

3 Numerical Weather Prediction Model Data

NWP provided several atmospheric variables forecasting up to 15 days ahead. All these models are operated by 15 different meteorological agencies around the world. For global purposes, we can find the Global Forecast System (GFS) used by the US National Oceanic and Atmospheric Administration (NOAA) and Integrated Forecast System (IFS) operated by European Centre for Medium-Range Weather Forecast (EDMWF). On the other hand, some mesoscale models are available only for some zones around the world but offer better spatial resolution. In this case, we can find MM5 developed by Pennsylvania State University and National Centre for Atmospheric Research (NCAR) or WRF model. Accuracy of these different models change depending on the temporal scale and geographic area, as explained in Sect. 1.

Recently, several works have been published associating NWP models predicted data with a post-processing method to improve hourly ground solar radiation forecasting for time horizons hours ahead. Some other references establish a forecasting improving using NWP models data as inputs in different statistical methods.

In this chapter, it is explained the methodology for working on the second manner. NWP models data predicted for the next day are used as inputs in statistical models to improve solar forecasting. In this case, the methodology is described using the European Centre for Medium-Range Weather Forecast (ECMWF). ECMWF-provided data comes within 3 h intervals, so an interpolation of the value into hourly data was necessary. ECMWF provides information about several meteorological variables for different altitudes, however in this case, we only explained a methodology for using the following variables described by latitude, longitude, and time:

- Total Cloud Cover (TCC), with values between 0 and 1 using a cloud index.
- Surface Solar Radiation Downwards (SSRD), for accumulative values of J/m2 within two instants.

4 Satellite Solar Radiation Data

As proposed first with NWP data, satellite-derived data will be used to improve solar radiation forecasting accuracy with statistical models. The most important characteristic of satellite data is their great spatial resolution and possibility of introducing

many information to see the evolution of the surroundings of the desired location. Indeed, the most important decision is the optimal selection of satellite pixels with the best information for the forecasting performance. The analysis of satellite data in a region surrounding the location where the solar radiation forecasting takes place is an important issue to establish an optimal selection.

4.1 Satellite Data Analysis

Satellite data offer solar radiation data with different spatial and temporary resolutions depending on the geographical area. These models provide GHI, DNI, and some great information about clouds and atmosphere conditions. Ineichen [20] provides an assessment study of several satellite-derived data for BSRN stations with hourly errors around 17% for global and 34% for direct normal irradiance. Anyway, depending on the location and climatic conditions, the uncertainties and deviation from ground measurement change significantly. Eissa [14] reports errors between 17 and 30% for different stations in Egypt, obtaining worst results for northern stations closer to the sea. Moreover, Mazorra [31] show errors with an average 12.2% rRMSE at C0-Pozo Izquierdo and 27.8% rRMSE at C1-Las Palmas, two stations in Gran Canaria island. The first station belongs to the southern area of the island with more occurrence of clear sky days, while the second station is situated in northern station with more cloudy days. Both works use satellite-derived hourly data from Helioclim3. On the other hand, Antonanzas [1] report around 4% rRMSE for a set of stations in Spain with yearly GHI data obtained from CMSAF database. In Gran Canaria island, for hourly data using CMSAF database with GHI an error was obtained from 15% in the south and 33% in the north, Fig. 2.

Calculate the error between satellite-derived data, both GHI or DNI, and ground measurement can show us the quality of the estimation. The more accuracy provided

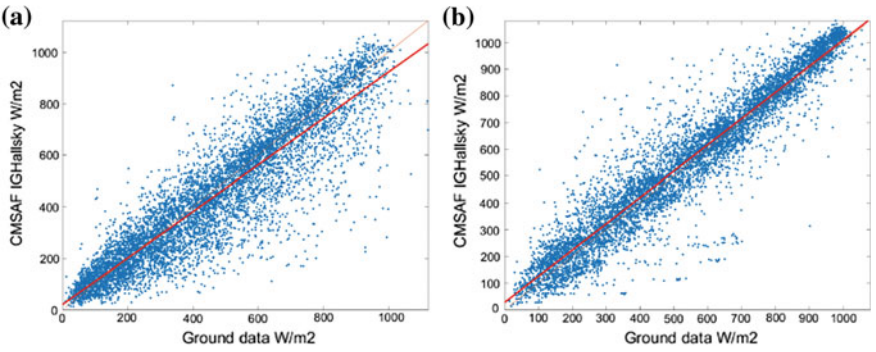


Fig. 2 CMSAF SIS hourly data comparison with ground data for northern station (**a**) and southern station (**b**) in Gran Canaria, Spain

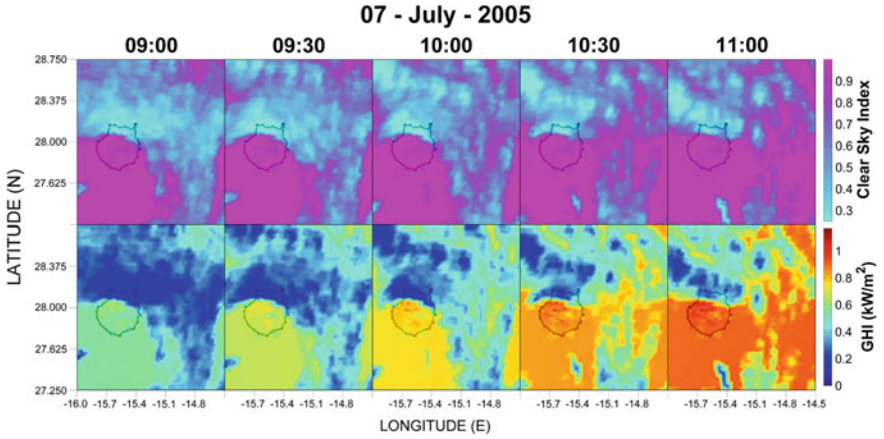
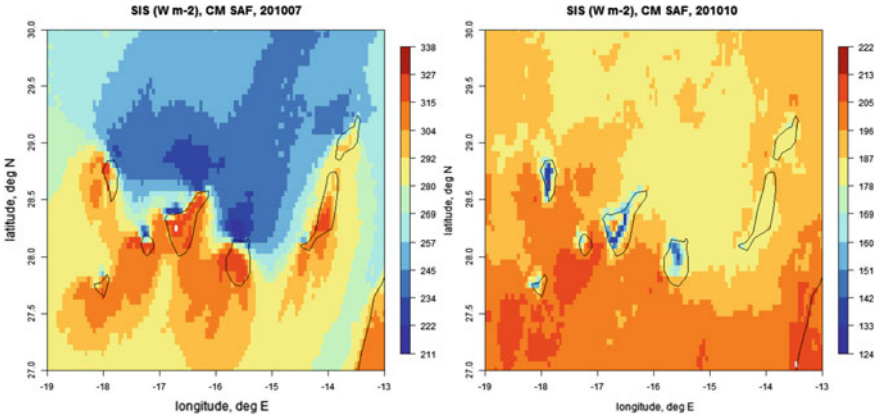


Fig. 3 Intra-day evolution 07-07-2005 of satellite-derived data each 30 min for Gran Canaria Island [31]

by the satellite derived, the more improvement in solar radiation forecasting will be obtained, instead of using only ground data. Moreover, an observation of satellite data all over the years could give an overview of the quality of the dataset. We can observe if satellite data represent specific climatic conditions in the location we are studying. Figure 3 shows that satellite data could estimate a particular behavior of known meteorological pattern in Canary Islands during summer. The northern part of the island presents an accumulation of clouds brought by the predominant trade winds. In this case, satellite-derived data used was obtained from the Helioclim-3 database version 5 (HC3v5). All this information has been processed by the Heliosat-2 method using images from the Meteosat geostationary [2, 3]. The selected area contains the entire island of Gran Canaria as well as a significant portion of sea at the north–east, motivated by the knowledge and influence of the trade winds in the Canary Islands. This area is defined, in decimal degrees, by the coordinates' latitude [+28.7500 to +27.2500], and longitude [−16.0000 to −14.5000], resulting in a grid of 61×55 pixels of information, where each pixel possesses a spatial resolution of 3×3 km^2.

In the same way, satellite-derived data obtained from the Satellite Application Facility on Climate Monitoring (CM SAF) showed the same good results for representing the cloud cover during summer. CMSAF information has been processed using images taken from the Meteosat Second-Generation (MSG) geostationary satellite network with SEVIRI sensor on board and NOAA polar satellites with AVHRR sensor [40]. These data are converted into global solar radiation and direct normal irradiance using Heliosat method and the Magic approach, validated with BSRN ground stations and provided in SARAH-2 database [41, 44]. The selected area contains the entire Canary Islands as well as a significant portion of sea. This area is defined, in decimal degrees, by the coordinates latitude [+27.0000 to +30.0000], and longitude [−19.0000 to −13.0000], where each pixel possesses a spatial resolution of 5×5 km^2 (Fig. 4).



(a) july (b) october

Fig. 4 CMSAF SIS monthly means of gridded satellite data for Canary Islands, between July and October 2010

4.2 Spatiotemporal Correlation Analysis

Statistical models usually forecast GHI data using only ground clear sky data as inputs. As explained, the aim is to improve statistical models hourly forecasting using different satellite information as well as ground data. Satellite-gridded data includes a huge amount of pixels, so statistical model computation would be very difficult using the whole radiation data. In order to introduce in these models the most representative information, one of the most important decisions is to select the optimal pixels from the total set. The variable used to establish the best satellite pixel is the Pearson correlation between ground data of each station and satellite data of the selected area [11, 31, 48]. Pearson correlation provides information about the weather relationship and establish a useful tool to enhance a prediction.

The higher Pearson correlation factor between a satellite pixel and the soil ground data at the studied location is, the more information about the surroundings provide this pixel. Indeed, pixels chosen to improve further prediction are those with the higher correlation factor. As proposed in [31, 48], clear sky index is the variable used for studying the correlation factor. To evaluate correlations between both the parameters, satellite and ground data sets, in different temporal moments a time lag is established. This time lag provides information about the best closest reactions in the area and gives us an important overview between ground data at the present moment and solar radiation from the surroundings in the past and possible incoming events, Eq. 9. It is suggested a selection of four time lags, for hourly data the time lags go from the same temporal moment to a 3 h earlier maximum. From 3 to 6 h, intercorrelation between satellite and ground dataset obtain values below 0.5, so the relation is not considered relevant in the studied cases.

$$C_{K_t^*}(i,j)_h = corr(K^*_{t,ground}(t), K^*_{t,satellite}(t-h)) \; para \; h = 0, 1, 2 \; \& \; 3 \qquad (9)$$

In Figs. 5 and 6, it is shown an example for two stations in Gran Canaria Island (Spain). Figure 5 represents the correlation between each pixel in the whole grid for each time lag with a station in the south. Each image corresponds to a correlation calculation using the whole time-lagged satellite grid and the ground measurements at the present time. In the same way, Fig. 6 shows the correlation with a northern station. In both the cases the results provided by the calculation resemble with the expected behavior. The higher correlated pixels belong to the part of the island surrounding each station and the correlation decreases while time lag increases. All these observations lead to dividing islands into two different zones and allows a better comprehension of islands' microclimate. North part of the island is heavily influenced by clouds created by trade winds and the complicated orography. On

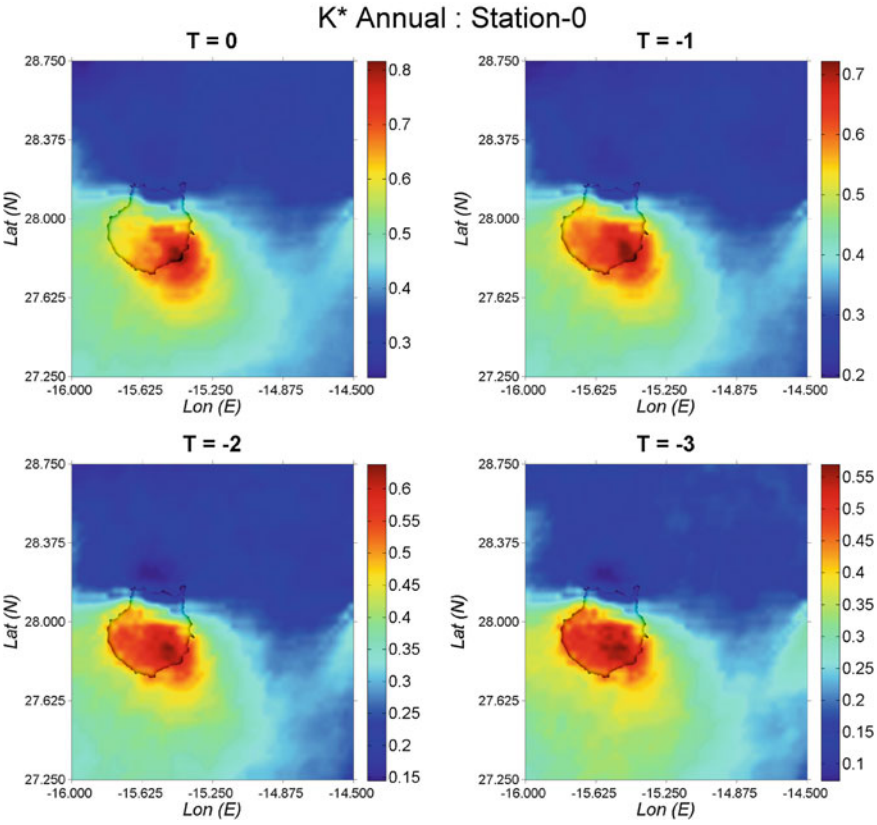


Fig. 5 Intercorrelation annual map for clear sky index between ground measurement and each satellite pixel around the measurements at station C0 Pozo Izquierdo for time lag h = 0, 1, 2, & 3 h [31]

the other hand, south sector remains protected from these clouds by the mountains. The empirical information about the island suggests that northern part possesses more cloudy days than south area and it is confirmed by the results obtained with Pearson's correlation. Indeed, correlated values estimated with satellite pixels and ground stations give us coherent information for selecting the best ones to improve solar forecasting.

Annual correlation is calculated using the whole year of each data set, both for satellite grid and ground measurement station. The results provide climatic conditions information from satellite grid radiation data in this area but it represents the statistical average on a whole year. Moreover, Zagouras et al. [48] propose a correlation using a temporal frame in order to evidence specific climatic conditions along the year. In this case, a correlation between satellite and ground data using different data sets for each meteorological quarters is shown, so the weather patterns where more coherent and accurate.

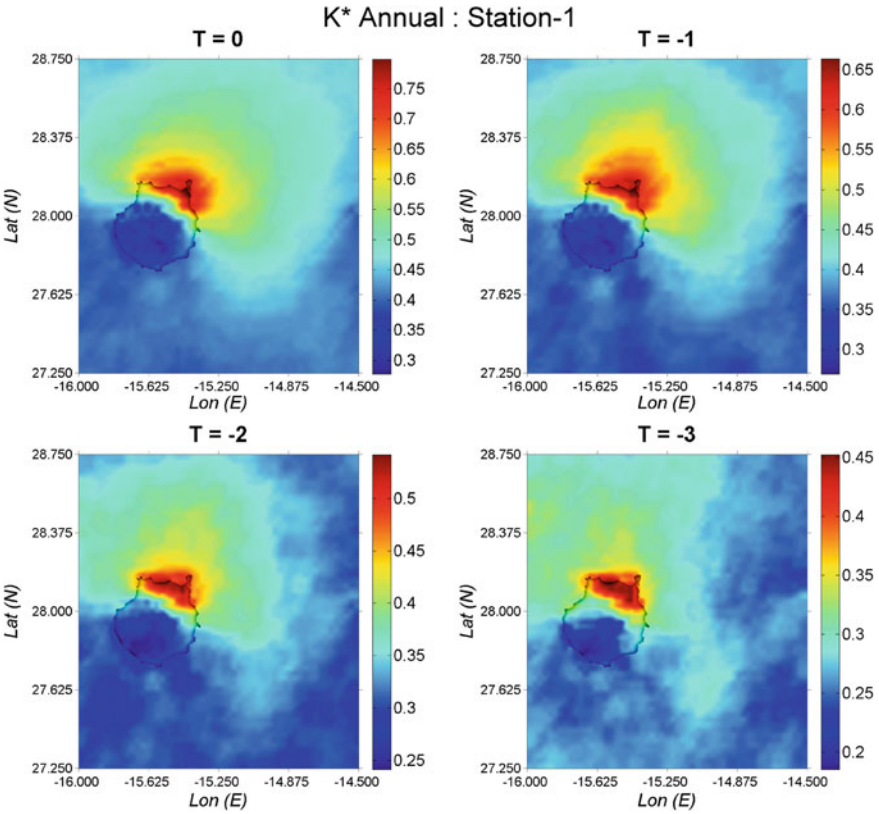


Fig. 6 Intercorrelation annual map for clear sky index between ground measurement and each satellite pixel around the measurements at station C1 Las Palmas for time lag h = 0, 1, 2, & 3 h [31]

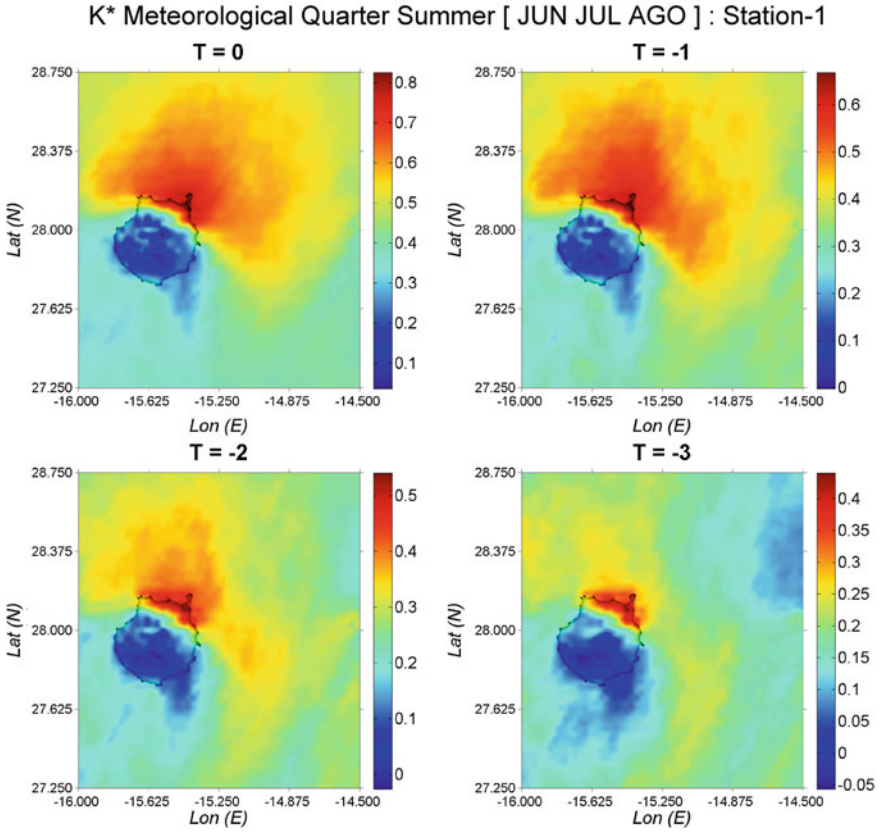


Fig. 7 Intercorrelation summer map for clear sky index between ground measurement and each satellite pixel around the measurements at station C1 Las Palmas for time lag h = 0, 1, 2, & 3 h [31]

For the example of northern station in Gran Canaria Island, Fig. 7 represents the correlation for summer data and Fig. 8 the correlation for autumn data. Both quarterly time-lagged correlation images offer climatic information consistency with empirical observations in all the seasons. In summer, due to the strong effect of trade winds, which creates a big area in the north of the island with a similar behavior. It is also important to remark the shelter provided but the orography, giving other climatic conditions to south of the Island. This shelter generates also a trail with similar climatic conditions on the sea, where we can also find clouds. In autumn and rest of the seasons, the Island is still divided into two regions but the higher correlations value correspond to pixels more concentrated around the ground station. In this season, the presence of winds is not so strong and there is not also influence in the surroundings.

The results obtained give us a similar behavior to the empirical climatic conditions in this region. Indeed, Pearson's correlation factor with time-lagged grid data give

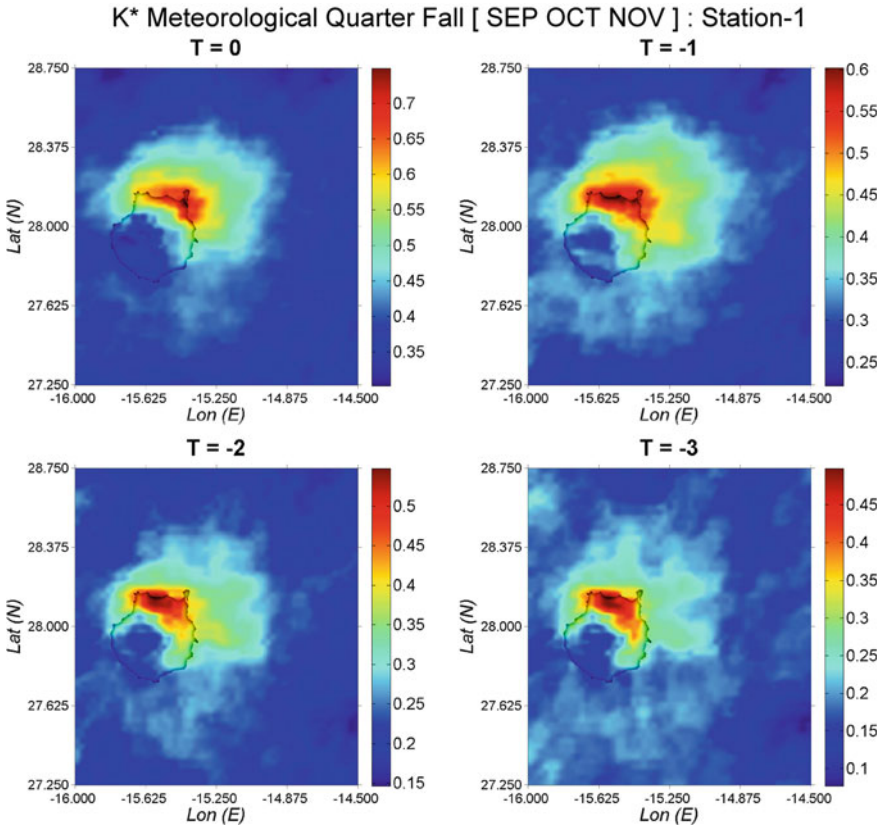


Fig. 8 Intercorrelation fal map for clear sky index between ground measurement and each satellite pixel around the measurements at station C1 Las Palmas for time lag h = 0,1, 2, & 3 h

us an important information for different time frames for selecting the most related pixels with the solar radiation we want to predict in the ground station. This technique will be used later to estimate the pixels to introduce in statistical models as inputs.

In the explained case, the time-lagged correlation was estimated using clear sky index data, both from satellite grid and ground data. In an attempt to get more additional data to enhance the prediction, Dambreville et al. [11] proposed a calculation using step by step clear sky index difference, Eqs. 10 and 11. This information should offer the direction of significant incoming patterns in the weather of the island, therefore, annual and quarterly analysis where made as in the previous case. These works were based on 15-min solar radiation data sets, instead of hourly data sets used by Mazorra et al. [31] and Zagouras et al. [48].

$$\Delta K_t^* = K_t^*(t+1) - K_t^*(t) \tag{10}$$

$$C_{\Delta K_t^*}(i,j)_h = corr(\Delta K^*_{t,ground}(t),\, \Delta K^*_{t,satellite}(t-h))\ para\ h = 0, 1, 2\ \&\ 3 \quad (11)$$

The results report a stretched area around the ground measurement station, indeed giving a more precise information about the most important pixels to improve the forecasting. However, these correlation values are not high, so the difference between selected and not selected pixels is not so relevant. The time-lagged images show a different behavior between east and west area (with higher correlated pixels in the west while lag increases), confirmed by the fact that wind mainly blows from west. Indeed, by selecting the most correlated pixels in each time lag we are giving information to the statistical models about incoming clouds.

5 Forecasting Statistical Models

The main purpose of this chapter is to explain a methodology to improve solar radiation forecasting for several hours ahead. As it was explained in Sect. 2, statistical models work with stationary data series. Hence, in this case the variable used in solar radiation forecasting models is the clear sky index. However, for calculating and discussing results and errors, the variable used is the global solar radiation GHI, estimated with Eq. 12.

$$\hat{GHI} = (\hat{K_t^*}) \cdot GHI_c \quad (12)$$

The general function used to connect input and outputs is Eq. 13.

$$\begin{aligned}\widehat{K_t^*}(t+h) = F[&K^*_{t,g}(t), \ldots, K^*_{t,g}(t-i), K^*_{t,e1}(t), \ldots \\ &K^*_{t,e1}(t-j), \ldots, K^*_{t,en}(t), \ldots, K^*_{t,en}(t-j)]\end{aligned} \quad (13)$$

where $\widehat{K_t^*}(t+h)$ is the clear sky index calculated for time horizon h, $K^*_{t,g}(t-i)$ is clear sky index from ground data set at the location for i past values and $K^*_{t,en}(t-j)$ corresponds to the n exogenous data with a j time lag. The number of exogenous data could vary depending on the selection of satellite pixels and NWP variables. One the most important decisions for the modeler is to choose the number of ground past data, number of satellite pixels, and NWP variables. Irrelevant inputs may unnecessarily increase model complexity and as a consequence may hamper the model performance. The general function F depends on the statistical model used and it is established during the training process. It is important to split measurement data set in training and testing sets. First one is used to establish the optimal function to relate input and output values, while testing set let us to calculate the accuracy of the model when new data are presented and controlled the overfitting.

5.1 Simple Forecasting Models

It is normal to compare any statistical model or methodology developed to improve solar radiation forecasting with a simple model. These simple models, naïve models, establish a reference forecasting limit that the new model should improve.

Two simple models for GHI hourly forecasting and different time horizons are suggested. Naïve models presented in this chapter only work with ground past values data, using clear sky index series. The first is the simple persistence (Pers) model [28], Eq. 14.

$$\widehat{K^*}(t+h) = K^*(t) \tag{14}$$

Persistence model is based on the assumption that atmospheric conditions remain invariant in two consecutive instants, indeed, that clear sky data for $t + h$ only depend on clear sky for the previous data. An easy improvement of this model is the smart persistence (smart pers). It consists of the forecast of the clear sky index for time horizon h using only the mean of h previous clear sky index with the ground data [18], Eq. 15.

$$\widehat{K^*}(t+h) = mean[K^*(t), ..., K^*(t-h)] \tag{15}$$

5.2 Linear Models

Statistical linear models have been widely developed for temporary series estimation. In this case, two different linear models are used for solar radiation forecasting using past ground data as inputs. The procedure explained in this chapter is based on linear models regression as described by Boland [4, 5] for solar radiation estimation in Australia using hourly and daily data.

- Autoregressive models (AR), a regression linear model based only on ground clear sky past data to forecast solar radiation for time horizon h, Eq. 16.

$$\widehat{K_t^*}(t+h) = \sum_{i=0}^{p-1} \left[\Phi_{i+1} K_t^*(t-i)\right] + \varepsilon_{t+h} \tag{16}$$

- Autoregressive moving average (ARMA), based on two linear models, an autoregressive model (AR) and a moving averages model (MA). This model estimates solar radiation forecast using a linear combination of different numbers of past data and error, Eq. 17.

$$\widehat{K_t^*}(t+h) = \sum_{i=0}^{p-1} \left[\Phi_{i+1} K_t^*(t-i)\right] + \varepsilon_{t+h} + \sum_{j=0}^{q-1} \left[\Theta_{j+1} \varepsilon_{t-j}\right] \tag{17}$$

In both the equations, $\widehat{K_t^*}(t+h)$ represents the solar radiation forecasting for time horizon h in terms of clear sky index, ε is a white noise and $K_t^*(t-i)$ are the ground clear sky past data from the measurement station used as inputs in AR model. For the AR model, Φ_{i+1} for $i = 1, 2, ..., p$ displays the autoregresive parameters and established the relation between clear sky past ground data and output data. While, for the MA model Θ_{j+1} for $j = 1, 2, ..., q$ shows moving average parameters that accompany the errors in MA regression. Both kind of values, Φ and Θ, are obtained during the training process. The methodology used to obtain both parameters is least square regression resulting from comparison of the set of past data used as input and future data that you want to predict.

The order p for the AR model shows the number of past data used to predict. One of the most important decisions during the training process is the model complexity. In this case, the optimal order p is obtained by calculating the partial autocorrelation function (PACF) and the Bayesian information criterion (BIC). Indeed, the model is defined by AR(p) depending on the number of past inputs used to obtain the best forecast data. On the other hand, the optimal order q for the MA model is obtained by calculating the autocorrelation function (ACF) and define the number of errors used during the prediction. In case of ARMA model, the optimal p and q orders should be established during the training process and the model is defined as ARMA(p,q).

These models are widely used for solar radiation forecasting because of the flexibility for working with temporary series depending on model orders. In many cases, very good results are described using AR and ARMA models with low-order parameters [7], which means not a long number of past clear sky values as inputs.

5.3 Artificial Neural Networks

Machine learning techniques have been described as very useful models for solar radiation forecasting. The method of machine learning explained in this chapter is artificial neural networks (ANN) [3]. However, many other techniques, such as Gaussian process or support vector machines, have been described in many papers with very good results and the methodology would be similar. ANNs is a statistical model that establishes a relation between a group of inputs and outputs during a training process. The model is based in a group of units, called neurons, that generate an output and received inputs from a group of input data or from other units. The units are connected between them by an associated weight. Each unit, neuron, receives the sum of different variables affected by these connection weights and produces an output. The output is obtained using a nonlinear activation function of transfer function to limit its amplitude and the input sums. The activation function used in this case is the hyperbolic tangent function, Eq. 18.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{18}$$

The neural network used in this case is the multilayer perceptron (MLP), as described in many engineering and forecasting applications. MLP consists of a group of input data, which makes up the input layer, connected by weights with at least one layer of neurons, called hidden layer, finally connected with the output layer neurons. The input layer is not neurons because it only contains the meteorological input data. On the other hand, neurons in hidden layer present a nonlinear transfer function (hyperbolic function), while the final output neuron uses a linear activation function. The output layer consists of a single neuron with the solar radiation (in terms of clear sky data) data for the time horizon we want to forecast, $\hat{K_t^*}(t+h)$. Each variable in the input layer is connected with each neuron in hidden layer by a first group of weights. All hidden layer outputs are also connected with the single output by a second group of weights, Eq. 19.

$$\hat{K_t^*}(t+h) = \sum_{j=1}^{H} \omega_{sj}^2 f_j [\sum_{i=0}^{T-1} (\omega_{ji}^1 K_t^*(t-i) + \omega_0^1] + \omega_0^2 \tag{19}$$

where $\hat{K_t^*}(t+h)$ is the forecast solar radiation for time horizon h, ω_{sj}^2 is the group of weights that connect the output of hidden layer neurons with general output, ω_{ji}^1 is the group of weights that connect each input i with each hidden unit j, $K_t^*(t-i)$ represents the inputs variables for the ANN and ω_0^1 and ω_0^2 are the biases for hidden and output layers. Input variables $K_t^*(t-i)$ could be only ground measurement past data or also other meteorological, satellite, or NWP data.

5.3.1 Backpropagation Training Process

Both groups of weights, ω^1 and ω^2, associated to each connection between input, hidden, and output layer are modified during the training process. The optimal group of weights is obtained by minimizing a cost function. The mean square error between the target forecast data $K_t^*(t+h)$ and the estimated data obtained with ANN is one the most common methods $\hat{K_t^*}(t+h)$, Eq. 20.

$$E(\omega) = \frac{1}{2} \sum_{i=1}^{N} [\hat{K_i^*}(t+h) - K_i^*(t+h)]^2 \tag{20}$$

The backpropagation algorithm is the optimizing method used to minimize the cost function. In this algorithm, first the ANN weight vectors are randomly initialized. During the training process, the weights ω_k are changed with each iteration by calculating a new group ω_{k+1} by minimizing $E(\omega)$ with a gradient descent process, Eq. 21. Where η is the learning parameter. Scaled conjugate gradient gives us an optimal solution to estimate gradient direction and learning parameter in each iteration. In this way, we get the optimal solution faster.

$$\omega_{k+1} = \omega_k - \eta \frac{\partial E}{\partial \omega_k} \tag{21}$$

5.3.2 Regularization Techniques

The network architecture is one of the most important issues to obtain the optimal accuracy of ANNs to approach continuous functions. If ANNs obtain very good results with the training data set and approximate the noise of the function, poor accuracy will be obtained when new data are presented. This problem is called overfitting. So, ANNs structure will determine the possibility of the function to be useful with a general data set. It is widely described the use of regularization techniques to avoid overfitting problem [3, 25]. This complexity control has been treated with different regularization techniques, as pruning methods [24], regularization coefficients, or Bayesian regularization framework [30]. Classical regularization techniques need to estimate the regularization coefficients using a cross-validation method. Control model complexity reduces the computational load and find inputs without any influence to improve forecasting, because their associated weights are pruned.

In this case, the number of hidden units and inputs are decided by using Bayesian regularization framework. This method controls the complexity of the model. Bayesian framework considers a probability density function over the weight space. Indeed, the optimal group of ANNs weight values agree to the maximum probability density function. In practice, Bayesian framework [29, 30] introduces two hyperparameters, α and β, to the cost function in order to control the model complexity, Eq. 22. Term E_ω in the cost function induces a decay in unnecessary weights, so at the end of training process it is possible to prune weights under a certain value. Bayesian framework permits to estimate hyperparameters at the same time that we are training our network.

$$S(\omega) = \frac{\beta}{2} E_D + \frac{\alpha}{2} E_\omega \tag{22}$$

$$E_\omega(\omega) = \frac{1}{2} \sum_{j=1}^{m} (\omega_j^2) \tag{23}$$

where m is the number of parameters of the whole ANN structure. Bayesian framework permits to estimate hyperparameters at the same time that we are training our network. So, not only overfitting is controlled but also it is studied the complexity of the model to reduce hidden and input units. As described in [25], Bayesian framework approach uses an iterative procedure to estimate hyperparameter's optimal values, α and β, and optimal group of weights ω_{MP}. This iterative procedure takes place only in the training dataset.

1. Hyperparameters α y β are initialized using small values and vector of weights is randomly set using a Gaussian distribution. In this iteration number k, estimated

weights ω^k and defined hyperparameters α^k and β^k give us first an ANN output and calculate the error function $S^k(\omega)$, Eq. 22.

2. The optimal vector of weights ω_{MP}^{k+1} is obtained in this step using an optimization algorithm, as scaled conjugate gradient. The number of iterations in this step depends on the convergence criterion decided for backpropagation process. With this optimal weight, we estimate cost function for iteration $k+1$, E_ω^{k+1} y E_D^{k+1}.
3. In this step hyperparameters, α^{k+1} and β^{k+1}, are recalculated using the following steps:

   a. $\gamma^{k+1} = \sum_{p=1}^{m} \left(\frac{\lambda_p}{\lambda_p + \alpha^k} \right)$, where λ_p are eigenvalues of error Hessian matrix without regularization term, $H = \beta^k \nabla\nabla E_D$.
   b. $\alpha^{k+1} = \frac{\gamma^{k+1}}{2E_\omega^{k+1}}$.
   c. $\beta^{k+1} = \frac{N\gamma^{k+1}}{2E_D^{k+1}}$.

4. Repeat step 2 using new parameters ω^{k+1}, α^{k+1}, and β^{k+1}, calculated in the previous step until reaching the convergence criterion.

These steps are repeated until the regularized error is equal to half of the number of data points. The theory states that $S(w) = N/2$ when $\alpha = \alpha_{MP}$ and $\beta = \beta_{MP}$. It is also possible to study the hyperparameters α and β and parameter γ in each iteration k and decide the convergence when they are almost constant.

Bayesian framework gives also the possibility to study the number of inputs and hidden units, model complexity. To study the number of inputs, we study the weights associated to each input to decide the influence in the final result. We can divide weights into different sets, one group for weights associated to each input, one for second layer of weights (connect hidden units with output), and one for each layer biases. Each group is controlled for an independent hyperparameter α_g. This technique is called automatic relevance determination (ARD). As different hyperparameters are assigned to each group of weights, during the training process it is possible to determine the most relevant inputs. Weights associated to a large α_g are supposed to be small. In this case, input related to this weight and hyperparameter is not relevant for network results and can be eliminated.

To control the number of hidden units, Bayesian framework estimates the probability for each model, called evidence of the model. Different ANNs are trained using several numbers of hidden units and the network with the highest evidence provides us the best one [24, 30, 32]. To calculate evidence of each model the final expression is Eq. 24 that calculate the log of evidence. Where N is the number of inputs, m is the total number of parameters, γ is the number of well-determined parameters (weights not close to zero), and $|A|$ is the determinant of the Hessian matrix of the total (regularized) error function $S(w)$.

$$logP(M_i|D) = -\alpha_{MP}E_{\omega}^{MP} - \beta_{MP}E_D^{MP} - \frac{1}{2}log|A| + \frac{m}{2}log\alpha_{MP} + \\ + \frac{N}{2}log\beta_{MP} + \frac{1}{2}log\left(\frac{2}{\gamma}\right) + \frac{1}{2}log\left(\frac{2}{N-\gamma}\right) \quad (24)$$

6 Numerical Statistical Models Implementation

Once the theoretical approach of statistical models has been explained, in this section it is described the implementation using different data sets (ground measurement, satellite-derived, and NWPs data). In case of linear models and ANNs, one of the most important decisions is the model complexity. Following sections explain how to work with both statistical models in order to choose the optimal model complexity and number of inputs. As it is necessary to split data sets into training and testing set, both groups should represent the same climatic conditions and seasonal events to work with similar relations between input and output data (for example, one whole year for each set).

6.1 Linear Models Complexity and Results

The model complexity is one of the most important issues to take into account by the modeler. The complexity of a linear model consists of the number of inputs for the AR model and the number of error terms for the MA model. This complexity is settled by estimating the order q and p of the model. If the model uses a great number of unnecessary parameters, the general accuracy could be worse. As explained in Sect. 5, to study the model complexity we use the sample of partial autocorrelation function (PACF), the sample of the autocorrelation function (ACF) and the Bayesian information criterion (BIC).

Partial autocorrelation function (PACF) sets the correlation between two instants of time series with a ρ delay. The sample PACF for the different time lags gave us the number of past values relevant for the forecasting. The maximum order p of the model is established within a range of 95% of this sample, Fig. 9.

Following the same criterion, the maximum q order is selected using the sample autocorrelation function (SACF). Once the maximum orders have been decided, for AR models we calculated several simulations using order p from 1 to maximum for all time horizons in order to select the optimal number of parameters. In case of ARMA model, we calculate different situations using all possible combinations with order p and q from 1 to maximum.

Finally, to decide the best option between all simulations the Bayesian information criterion (BIC) and the error of the model %rRMSE (with testing data set) give us the optimal solution. For each time horizon AR model, optimal solution is obtained with different p orders.

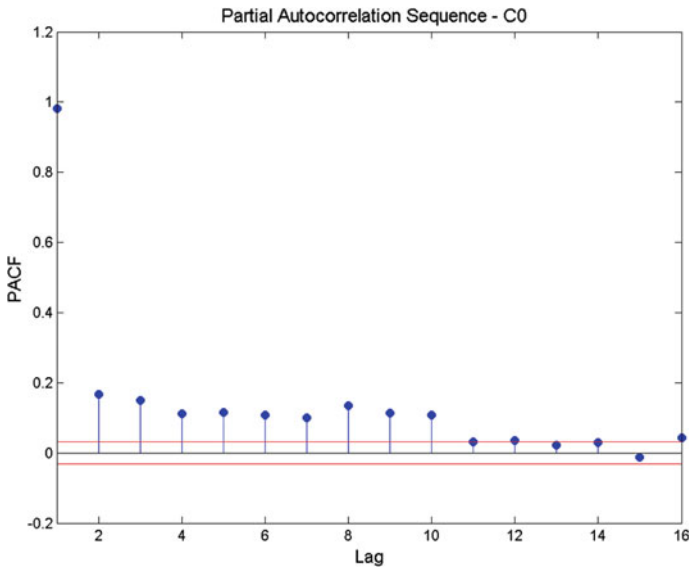


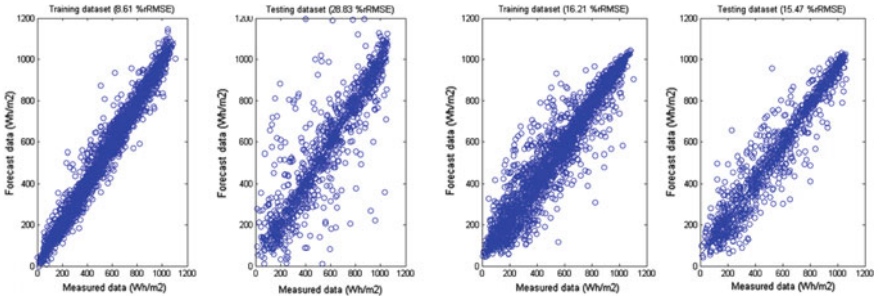
Fig. 9 Sample partial autocorrelation factor (SPACF) using solar radiation clear sky index. In this case, the maximum order was selected in $p = 12$

In the same way, ARMA model optimal solution was established calculating BIC and %rRMSE for all scenarios. In most cases, optimal model shown by BIC gives us different results of p and q orders. However, in many cases when compared with the optimal solution obtained with BIC to a simple ARMA model using $p = 2$ and $q = 1$, there is not a substantial improvement in terms of error %rRMSE. The optimal solution could need different orders for each time horizons and a huge number of input data (i.e., p order around 11), while ARMA(2,1) is a very simple model using only past input data to obtain a solar radiation forecasting.

6.2 ANNs Optimal Selection Using Ground Data

ANNs complexity in one of the most important issues to obtain the optimal forecasting accuracy. As explained in Sect. 5.3.2 we focus in selecting the number of inputs and hidden units. Bayesian framework gives us the possibility of selecting the number of inputs with ARD technique and the number of hidden units calculating the log of evidence.

Moreover, Bayesian framework controls the overfitting of the model [24]. Figure 10 shows the final result obtained with training and testing datasets forecasting using classical NN (a) and bayesian NN (b). In the first case, it is possible to observe a major dispersion in testing set because the model has overfitted the



(a) Classical ANN (b) Bayesian NN

Fig. 10 Measured data versus forecasted data for the training and testing (right) datasets using classical ANN (**a**) and Bayesian NN (**b**)

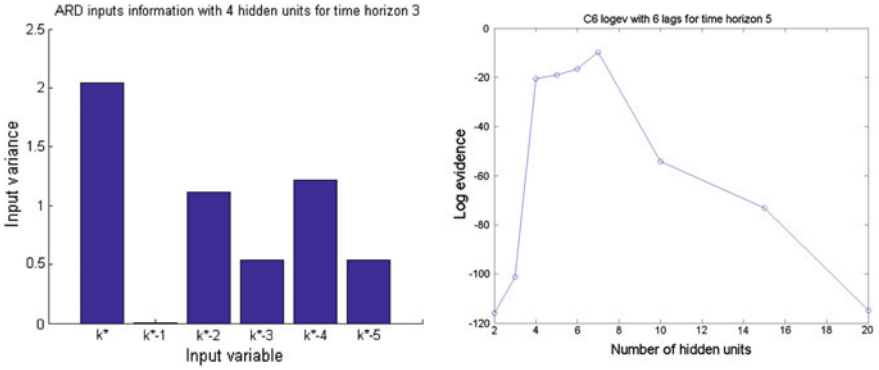
training set (overfitting problem). While with Bayesian NN the dispersion in training and testing sets remain almost similar, overfitting problem is not present.

In case of using only ground data, the number of inputs of each model corresponds to the number of past ground measurement using to forecast the solar radiation for time horizon h. ARD assigns a different hyperparameter α_g to each group of weights associated with one input. At the end of the training session, the weights with a large α_g are close to zero. In this case, the corresponding input is considered not relevant for the network and can be eliminated. In practice, each hyperparameter is represented in a figure with his variance. Inputs with a low bar comparde with other hyperparameters associated to the rest of inputs is considered irrelevant and could be eliminated. Figure 11a shows the result obtained with six past clear sky index inputs. Sometimes, pruned inputs are considered irrelevant with ARD technique, as the second input in Fig. 11, do not reproduce more accurate results and it is better to use all inputs. It is advisable to check the general error of the model when we prune these inputs.

The number of hidden units is settled once it is decided the number of inputs that give us the optimal results. Bayesian framework calculates the log of evidence between several ANNs with different number of hidden units, Eq. 24, to establish the optimal one. The ANNs with the higher Log of evidence is considered the best one. As in Fig. 11b, most of results show low number of units. As explained with ARD technique, log of evidence give us information about the best number of units but it is recommended to calculate the error of several models around the best one to establish the optimal number.

6.3 Exogenous Data Optimal Selection

The aim of using exogenous data is to improve ANN's hourly forecasting obtained only with ground data. Exogenous data used in this case are NWP's data and a grid of



(a) ARD (b) Log of Evidence

Fig. 11 ARD information for six ground data (**a**) and log of evidence for different hidden units (**b**)

satellite-derived data. These exogenous data will be added to the number of ground measurement inputs obtained in Sect. 6.2.

NWPs data suggested are total cloud cover (TCC) and surface solar radiation downwards (SSRD) obtained for the location of study. Both data are the one day ahead prediction for the hour we want to predict, $k_t^*(t+h)$, estimated by a NWP model.

Moreover, it is proposed to use also satellite-derived data in order to include information of the surroundings to the ANN. Satellite-gridded data includes a huge amount of pixels, so ANN computation would be very difficult using the whole radiation data. In order to introduce the most representative information obtained from satellite data, one of the most important decisions is to select optimal pixels from the total set. The variable used to establish the best satellite pixel is the Pearson correlation between satellite-gridded data and ground data, [11, 31, 48].

This Pearson correlation is calculated for each station between ground data at the present time and satellite pixel with time lags. During the training proces, we used time lags from $t = 0$ to a maximum of 3 h obtaining four time-lagged images. This correlation quantifies a relation between ground data and satellite for different time lags. After 3 h, the correlation between present ground data and past satellite data is not representative. Consequently, Pearson correlation gives us information about meteorological event incoming from the surroundings included in satellite images.

In that way, we can select pixels from the surroundings that represent the highest relation with ground station data for different time lags. ANN improves solar forecasting depending on the satellite information we use as inputs. The selection of optimal group of pixels is one the most important issues in this field. Dambreville et al. [11] proposes to use Pearson correlation of clear sky index variation between satellite and ground data with a time lag from 15 to 60 min. each 15 min. It is suggested to use a fixed number of pixels from each time-lagged image to improve solar forecasting with a linear statistical model. While Zagouras et al. [48] choose the 100

most correlated pixels from all time-lagged images. They work with hourly data and choose the best pixel using clear sky index Person correlation between satellite and ground data. Number of optimal pixels in each time- lagged image is settled with the genetic Algorithm.

Mazorra et al. [31] considered a maximum number of 30 satellite-derived radiation data. During the training process, six different tests at each station based on selecting different pixels were made. For the first test, the number of pixels over 0.5 correlation values were retrieved for each image and later the distribution of pixels between the four time-lagged images was computed. The considered 30 satellite pixels are the highest correlated for each image according to this percentage distribution (e.g., in Test-1 56% of pixels at time lag t = 0; 30% at $t = -1$; 8% at $t = -2$ and 6% at $t = -3$ should be selected, which means to retrieve 17, 9, 2, and 2 pixels respectively). As most of the optimal pixels were taken from first two time-lagged images, it was considered five tests using different percentage distributions of pixels. Test-2 and Test-3 estimated a new distribution taking into account more pixels from the other two images. While Test-4 only added satellite pixels only from time lags, $t = 0$ and $t = -1$. Finally, Test-5 formulates the same procedure as Test-1 but calculates a different distribution for every quarterly group of images for each station. Test-6 composes a new distribution using the best previous percentage distribution but selecting pixels from quarterly images.

The huge amount of satellite-derived data makes the computation difficult, so a median filter for each 3×3 satellite pixels is applied. Consequently, a superpixel was created computing GHI median value of every 3×3 group of pixels, Fig. 12.

To improve the previous work that use a different distribution of pixels for each test, it is suggested an automatic methodology. The estimation of the optimal number of pixel is based on the same Pearson correlation calculation. Instead of selecting a fix limit correlation value (0.5) to generate the distribution, it was considered different tests changing this limit. The percentile of the whole Pearson correlation distribution for all time-lagged images was established as the limit. In that way, it is possible to change the distribution of pixel from the four images depending on the percentile considered. During the training process, different percentiles from 0.1 to 0.9 were

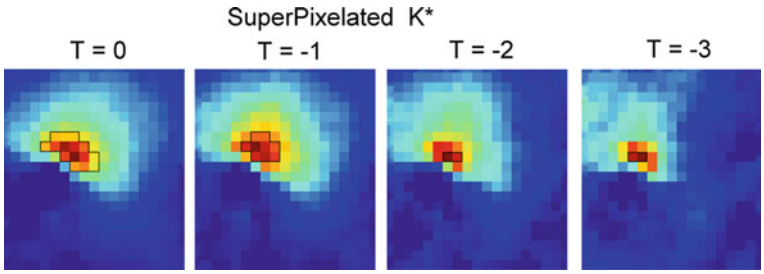


Fig. 12 Superpixel (3×3) selection at station in Gran Canaria (Spain) for time-lagged correlation images, t = 0, 1, 2, & 3 h. Black area shows selected superpixels

suggested. For the first one, more pixels from first images (time lag 0 and 1 h) are selected, while for the second one more pixels from latest images were extracted.

Once selected a different test, ANNs are trained using ground past data and satellite data for each test. Comparing forecasted hourly GHI with measured data for the testing dataset using the relative root mean square error, the best ANN's architecture and the optimal satellite information is selected for each case. In each location, this procedure should be repeated for the different time horizons' solar radiation forecasting.

6.4 Solar Radiation Forecasting Results

Statistical forecasting models explained in this chapter should be validated with measurement data. It is possible to find several error metrics suggested in specialized bibliography to establish the accuracy of each model. The error of the models are calculated with testing data set, because it is necessary to evaluate the capacity of each model to generalize the results with unknown data. All metrics are expressed in terms of GHI (W/m^2) even if clear sky model was the variable used during the training season. The most common error metrics are root mean square error (RMSE), mean absolute error (MAE), or mean bias error (MBE) and their relative metrics calculated dividing by the generally measured mean for the testing data set. It is also widespread the use of SKILL metric. This metric calculates the difference of each model with a simple model used as a reference. In this case, it is explained SKILL error metric compared with persistence model. Indeed, this value gives us how the described model improves a simple persistence model. This chapter shows some examples obtained using two measurement stations in Gran Canaria (Spain), calculating the accuracy of each model using %rRMSE, Eq. 25, and SKILL, Eq. 26. The models in terms of these error metrics are the following:

- Persistence Model—*Pers*
- Smart Persistence Model—*Smart Pers*
- Autoregressive Moving average with orders (2,1)—*ARMA(2,1)*
- Artificial Neural Networks with ground data—*NN*
- Artificial Neural Networks with ground data, satellite data and NWP data—*NN+ECMWF+SAT* Model

$$RMSE_{modelo} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{GHI}_{g,i} - GHI_{measure,i})^2} \tag{25}$$

$$SKILL(\%) = \left(1 - \frac{RMSE_{modelo}}{RMSE_{persistence}}\right) x100 \tag{26}$$

Table 1 RMSE for time horizons $h = 1...6$ in two stations in Gran Canaria (Spain)

Stations	Models	1 h	2 h	3 h	4 h	5 h	6 h
C0	Persistence	92.47	128.04	149.88	168.08	176.17	177.26
	Smart persistence	92.47	124.64	140.10	144.44	141.50	138.69
	ARMA(2,1)	85.78	109.30	119.77	126.64	129.32	130.38
	NN	88.20	113.39	125.13	127.12	131.68	130.03
	NN+ECMWF+SAT media IGH = 543.10 Wm^{-2}	84.00	106.17	110.51	114.93	118.89	120.43
C1	Persistence	118.95	167.03	195.15	213.39	224.71	228.18
	Smart persistence	118.95	169.11	190.69	195.34	190.21	182.18
	ARMA(2,1)	111.44	145.14	159.90	167.17	170.65	171.49
	NN	110.63	143.90	157.06	162.11	162.09	162.88
	NN+ECMWF+SAT media IGH = 433.79 Wm^{-2}	104.75	134.37	142.82	145.41	147.31	147.88

Table 1 shows the results in terms of RMSE in (W/m^2) and the ground measurement mean for the testing dataset. While, Fig. 13 describes the results in terms of %rRMSE. Both of them give the results for time horizons between 1 and 6 h ahead and for two ground measurement stations. First station (C0) is located to the south of the island and presents better results because the weather is more stable along the year with more presence of clear sky day. On the other hand, C1 station is on the north of the island a presents more cloudy and unstable days during the year, so error metrics are worse. As it is obvious, all the models obtain better results for shorter time horizons and get worse results while increase time horizon. In case of persistence simple models this growth is much more pronounced, while ARMA and NN with or without exogenous data control the error for large time horizons. Even if smart persistence presents an improvement compared to persistence, in larger time horizons it presents still some problems. ARMA(2,1) and NN only use ground measurement past data as inputs get similar results in terms of RMSE for both the stations and time horizons. Both the models improve significantly as simple models. Moreover, the inclusion of exogenous data in NN as inputs improve also the model and obtain the best accuracy for both stations and time horizons.

In Fig. 14 it is possible to see the SKILL(%) parameter for both the stations and all time horizons. In this figure, it is shown the different combinations of NN inputs in order to discuss the importance of everyone: NN only with ground data, NN with ground and satellite, NN with ground and NWP data, and NN with ground and all exogenous data. The SKILL forecast increases with time horizon, which means that the more far ahead in time, the better results we get with ANN+ECMWF+SAT method compared with persistence model. For both the stations the best model is the neural networks with ground, satellite and NWP data as inputs. Moreover, it is also observable that satellite data (NN+SAT) give better results for the first three time horizons, from 1 to 3 h, while NWP data (NN+ECMWF) is the best model from time horizon 4 to 6 h. It could also be interesting the results separating testing data sets in

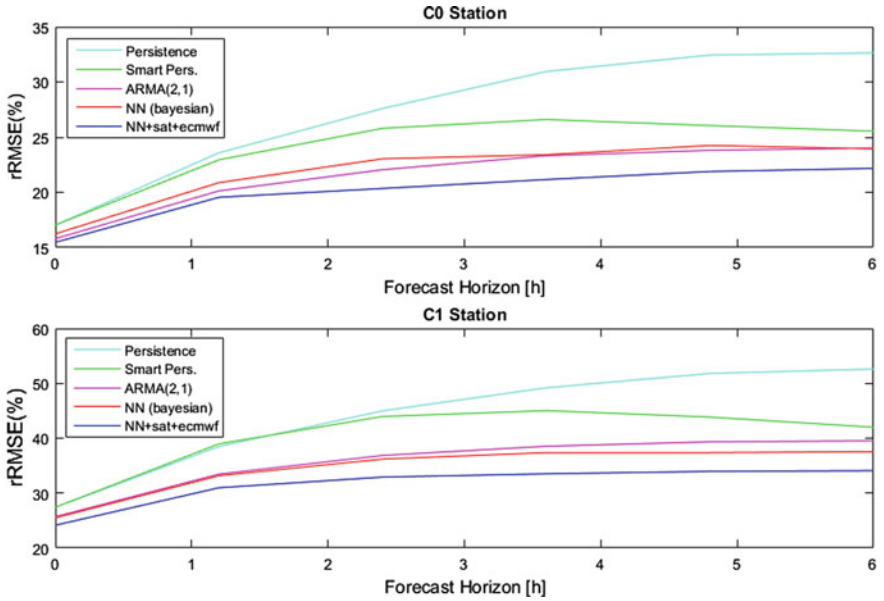
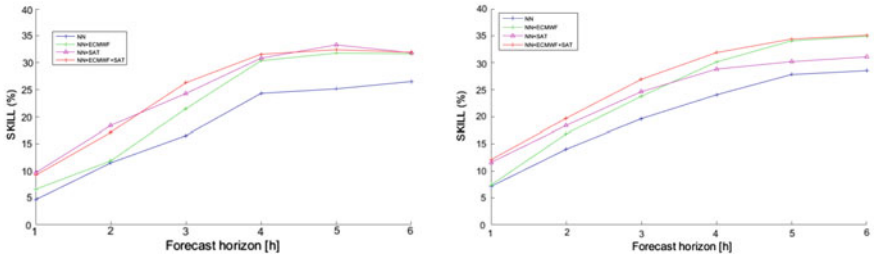


Fig. 13 %rRMSE results using testing data set for two different stations in Gran Canaria (Spain) with several forecasting models. C0 Station (up) & C1 Station (down)



(a) C0 Station (b) C1 Station

Fig. 14 SKILL(%) results for two different stations in Gran Canaria (Spain) using exogenous data with ANNs

the different seasons of the year or type of days (i.e., cloudy or sunny days). In this way, it is possible to establish a different model depending on the weather conditions or the time of the year.

7 Conclusions

The main conclusion of this chapter is that ANN and ARMA model present very good results in solar radiation hourly forecasting compared with persistence simple models. On the other hand, when exogenous data, as satellite data and NWP data, are introduced to the ANN as inputs we obtain an important improvement. In this case, we used solar radiation from several pixels around the measurement station with time lagged from $t = 0$ h to $t = -3$ h compared to present time. NWP data used to improve solar radiation forecasting are 24 h ahead the prediction of total cloud cover and surface solar radiation forecasting for the time step we want to forecast. One of the most important decision in order to obtain more accurate results is to find the optimal satellite pixels. A huge number of pixels without relevant information for solar radiation forecasting causes a high computation cost with ANNs and worse estimation errors. Pearson's correlation between ground and satellite data give us critical information to select optimal satellite pixels. The architecture of neural networks influences the final result of the estimation. The Bayesian methods explained in this section are considered an adequate tool to estimate the number of inputs and hidden neurons. With this method, it is possible to avoid overfitting problem and obtain accurate prediction results with both training and testing data.

References

1. Antonanzas-Torres F, Cañizares F, Perpiñán O (2013) Comparative assessment of global irradiation from a satellite estimate model (cm saf) and on-ground measurements (siar): a spanish case study. Renew Sustain Energy Rev 21:248–261
2. Bird RE, Hulstrom RL (1981) Simplified clear sky model for direct and diffuse insolation on horizontal surfaces. Solar Energy Research Inst., Golden, CO (USA), Technical report
3. Bishop CM (1995) Neural networks for pattern recognition. Oxford university press
4. Boland J (1995) Time-series analysis of climatic variables. Sol Energy 55(5):377–388
5. Boland J (2008) Time series modelling of solar radiation. Springer
6. Bosch J, Lopez G, Batlles F (2008) Daily solar irradiation estimation over a mountainous area using artificial neural networks. Renew Energy 33(7):1622–1628
7. Box G, Jenkins G (1998) Time series analysis, forecasting and control. Wiley
8. Canada's E (2015) World ozone monitoring mapping. http://es-ee.tor.ec.gc.ca/e/ozone/ozoneworld.htm/
9. Chatfield C (2013) The analysis of time series: an introduction. CRC press
10. Chow CW, Urquhart B, Lave M, Dominguez A, Kleissl J, Shields J, Washom B (2011) Intra-hour forecasting with a total sky imager at the uc san diego solar energy testbed. Sol Energy 85(11):2881–2893
11. Dambreville R, Blanc P, Chanussot J, Boldo D (2014) Very short term forecasting of the global horizontal irradiance using a spatio-temporal autoregressive model. Renew Energy 72:291–300
12. Diagne M, David M, Boland J, Schmutz N, Lauret P (2014) Post-processing of solar irradiance forecasts from wrf model at reunion island. Sol Energy 105:99–108
13. Diagne M, David M, Lauret P, Boland J, Schmutz N (2013) Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renew Sustain Energy Rev 27:65–76

14. Eissa Y, Korany M, Aoun Y, Boraiy M, Abdel Wahab MM, Alfaro SC, Blanc P, El-Metwally M, Ghedira H, Hungershoefer K et al (2015) Validation of the surface downwelling solar irradiance estimates of the helioclim-3 database in egypt. Remote Sens 7(7):9269–9291
15. Gueymard CA (2008) Rest2: high-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation-validation with a benchmark dataset. Sol Energy 82(3):272–285
16. Hammer A, Heinemann D, Lorenz E, Lückehe B (1999) Short-term forecasting of solar radiation: a statistical approach using satellite data. Sol Energy 67(1):139–150
17. Heinemann D, Lorenz E, Girodo M (2006) Forecasting of solar radiation: solar energy resource management for electricity generation from local level to global scale. Nova Science Publishers, New York
18. Hoff TE, Perez R (2012) Modeling pv fleet output variability. Sol Energy 86(8):2177–2189
19. Holben BN, Eck T, Slutsker I, Tanre D, Buis J, Setzer A, Vermote E, Reagan J, Kaufman Y, Nakajima T et al (1998) Aeroneta federated instrument network and data archive for aerosol characterization. Remote Sens Environ 66(1):1–16
20. Ineichen P (2014) Long term satellite global, beam and diffuse irradiance validation. Energy Procedia 48:1586–1596
21. Ineichen P (2016) Validation of models that estimate the clear sky global and beam solar irradiance. Sol Energy 132:332–344
22. Kostylev V, Pavlovski A et al (2011) Solar power forecasting performance–towards industry standards. In: 1st International workshop on the integration of solar power into power systems Aarhus, Denmark
23. Lauret P, David M, Fock E, Bastide A, Riviere C (2006) Bayesian and sensitivity analysis approaches to modeling the direct solar irradiance. J Sol Energy Eng 128(3):394–405
24. Lauret P, Fock E, Mara TA (2006) A node pruning algorithm based on a fourier amplitude sensitivity test method. IEEE Trans Neural Netw 17(2):273–293
25. Lauret P, Fock E, Randrianarivony RN, Manicom-Ramsamy JF (2008) Bayesian neural network approach to short time load forecasting. Energy Convers Manag 49(5):1156–1166
26. Lauret P, Voyant C, Soubdhan T, David M, Poggi P (2015) A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol Energy 112:446–457
27. Lefevre M, Oumbe A, Blanc P, Espinar B, Gschwind B, Qu Z, Wald L, Schroedter-Homscheidt M, Hoyer-Klick C, Arola A et al (2013) Mcclear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. Atmos Meas Tech 6(9):2403–2418
28. Lorenz E, Heinemann D (2012) Prediction of solar irradiance and photovoltaic power—Comprehensive Renewable Energy. Elsevier, Oxford, pp 239–292
29. MacKay DJ (1992) A practical bayesian framework for backpropagation networks. Neural Comput 4(3):448–472
30. MacKay DJ (2003) Information theory, inference and learning algorithms. Cambridge university press
31. Mazorra Aguiar L, Pereira B, David M, Daz F, Lauret P (2015) Use of satellite data to improve solar radiation forecasting with bayesian artificial neural networks. Solar Energy
32. Penny WD, Roberts SJ (1999) Bayesian neural networks for classification: how useful is the evidence framework? Neural Netw 12(6):877–892
33. Perez R, Kivalov S, Schlemmer J, Hemker K, Renné D, Hoff TE (2010) Validation of short and medium term operational solar radiation forecasts in the US. Sol Energy 84(12):2161–2172
34. Perez R, Lorenz E, Pelland S, Beauharnois M, Van Knowe G, Hemker K, Heinemann D, Remund J, Müller SC, Traunmüller W et al (2013) Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. Sol Energy 94:305–326
35. Perez R, Moore K, Wilcox S, Renné D, Zelenka A (2007) Forecasting solar radiation-preliminary evaluation of an approach based upon the national forecast database. Sol Energy 81(6):809–812
36. Petruccelli JD, Nandram B, Chen M (1999) Applied statistics for engineers and scientists. Prentice Hall New Jersey

37. Polo J, Zarzalejo LF, Salvador P, Ramírez L (2009) Angstrom turbidity and ozone column estimations from spectral solar irradiance in a semi-desertic environment in Spain. Sol Energy 83(2):257–263
38. Reno MJ, Hansen CW (2016) Identification of periods of clear sky irradiance in time series of ghi measurements. Renew Energy 90:520–531
39. Reno MJ, Hansen CW, Stein JS (2012) Global horizontal irradiance clear sky models: implementation and analysis. SANDIA report SAND2012-2389
40. Schulz J, Albert P, Behr HD, Caprion D, Deneke H, Dewitte S, Durr B, Fuchs P, Gratzki A, Hechler P et al (2009) Operational climate monitoring from space: the eumetsat satellite application facility on climate monitoring (cm-saf). Atmos Chem Phys 9(5):1687–1709
41. Scientist CS (2016) Annual product quality assessment report 2015. EUMETSAT Satellite Application Facility on Climate Monitoring
42. Sengupta M, Habte A, Kurtz S, Dobos A, Wilbert S, Lorenz E, Stoffel T, Renné D, Gueymard C, Myers D et al (2015) Best practices handbook for the collection and use of solar resource data for solar energy applications. NREL
43. Transvalor MP (2014) Soda solar radiation data http://www.soda-pro.com/. Accessed 2017
44. Trentmannn J, Huld T (2016) Meteosat-east solar surface irradiance data records. EUMETSAT Satellite Application Facility on Climate Monitoring
45. Urquhart B, Ghonima M, Nguyen D, Kurtz B, Chow C, Kleissl J (2013) Sky imaging systems for short-term forecasting. J Elsevier, Waltham, Massachusetts, Kleissl
46. Wittmann M, Breitkreuz H, Schroedter-Homscheidt M, Eck M (2008) Case studies on the use of solar irradiance forecast for optimized operation strategies of solar thermal power plants. IEEE J Sel Top Appl Earth Obs Remote Sens 1(1):18–27
47. Younes S, Muneer T (2007) Clear-sky classification procedures and models using a world-wide data-base. Appl Energy 84(6):623–645
48. Zagouras A, Pedro HT, Coimbra CF (2015) On the role of lagged exogenous variables and spatio-temporal correlations in improving the accuracy of solar forecasting methods. Renew Energy 78:203–218

Solar Radiation Probabilistic Forecasting

Mathieu David and Philippe Lauret

Abstract In contrast to deterministic forecasts, probabilistic forecasts give additional information about the inherent uncertainty embodied in weather predictions. In the realm of solar forecasting, prediction intervals are especially important to assess risks in grid operations and to optimize the energy storages needed to ensure the supply–demand balance. Even if the development of probabilistic solar forecasts is relatively recent, the main available methods come from other fields of meteorology, particularly from the wind domain. This chapter reviews some of the methods used to generate probabilistic solar forecasts. A special emphasis is put on short term (from several hours to several days) and very short term (from several minutes to several hours) forecasts. As the verification of the quality of the probabilistic forecasts is of major interest, graphical tools like reliability diagram and rank histogram are depicted. These diagnostic tools are relevant for assessing the good calibration of the probabilistic forecasts. In addition, a quantitative score, the CRPS, is also proposed. The CRPS takes into account the different sources of uncertainties and as a proper score, the CRPS is useful to rank competing forecasting methods.

1 Nature of a Probabilistic Forecast

As shown by Lorenz [42], the atmosphere is a chaotic system with flow-dependent error growth. Indeed, an initial state of the atmosphere can lead to extremely different weathers in the future. This characteristic of the atmosphere is well-known and is commonly referred as the butterfly effect. As a consequence, no deterministic forecast is perfect. We can even say that a forecast is uncertain by nature. Thus, prediction intervals are necessary to assess the uncertainty associated to the point forecasts.

M. David (✉) · P. Lauret
Laboratoire PIMENT (Physique et Ingénierie Mathématique pour l'Énergie, l'environnemeNt et le bâtimenT), Université de La Réunion, 15 avenue René Cassin, 97715 Saint-Denis Cedex, France
e-mail: mathieu.david@univ-reunion.fr

P. Lauret
e-mail: philippe.lauret@univ-reunion.fr

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_9

Regarding the high penetration of solar renewables in energy grids, the knowledge of the uncertainty of the solar forecasts is of high importance. Indeed, the prediction intervals of the power output of the solar systems allow assessing the risk related to the scheduling of energy sources and to the planning of units commitment. Thus, they should be key feature for the optimization of the operations of such grids.

1.1 Sources of Uncertainties in Solar Forecasts

The main source of uncertainties is the chaotic behavior of the atmosphere. From analog initial conditions, a small event could trigger important weather changes. The main effects could be predicted, but important details might lie forever beyond calculation [56]. As many of other weather variables, the cyclic and seasonal evolution of the solar irradiance can be accurately reproduced [31]. The clear sky models were developed to predict this periodic and deterministic component. Therefore, most of the forecasting models are built to predict the clear sky index (see Chap. 8), i.e., the stochastic component of the solar irradiance variability.

Another source of uncertainty originates from the forecasting models themselves. Their uncertainty partly lies in their approximate reproduction of the atmospheric phenomena, in the inaccuracy of the initial conditions or in the assumption of analogy between past and future events.

As data from environmental monitoring are used to build the forecasting models, a last source of error is the measurement noise. This source of uncertainty is constant in time and in space. It depends only on the technology of the sensors. In the field of solar energy, the most common sensors are photopiles (pyranometers) and calibrated PV cells. It must be noted, however, that the accuracy of the current state of the art forecasting models is largely worse than the accuracy of these sensors. Consequently, here we will not address the issue related to the measurement noise.

Regardless of the time horizon and the used model, forecast accuracy is strongly dependent on the temporal variability of the sky conditions. Forecasting errors increase with the temporal variability. This behavior has been highlighted for short-term forecasts done by a NWP [37] and very short-term forecasts done by learning machine techniques [38].

1.2 Quantifying the Uncertainties of a Forecast

Probabilistic forecasts correspond to the estimation of the statistical distribution of the forecast error. Without any assumption on the distribution of the forecast error, a probabilistic forecast is provided as a set of quantiles spanning the unit interval. This type of probabilistic forecast is usually called predictive distributions [47]. For example, during the Global Energy Forecasting Competition 2014 (GEFCom 2014) [30], the solar forecasts were expressed in the form of 99 quantiles with various nominal proportions between zero and one.

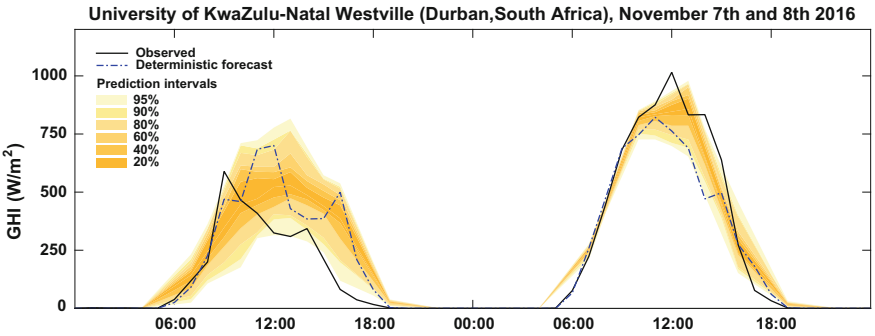


Fig. 1 Example of probabilistic solar forecasts: 2 days of measured GHI at the University of KuwaZulu-Natal Westville (South Africa) [9] and associated forecasts with prediction intervals provided by ECMWF-EPS

Let us consider F the cumulative distribution function of a random variable X, such that $F(x) = Pr(X \leq x)$. The quantile q_τ, at probability level $\tau \in [0, 1]$ is defined as follows:

$$q_\tau = F^{-1}(\tau) = inf\{y : F(y) >= \tau\}, \tag{1}$$

A quantile q_τ indicates that there is τ probability that the observation falls below the quantile q_τ. As mentioned above, a set of quantiles represents the predictive distribution of the variable to predict. Thus, prediction intervals (or interval forecasts) with different nominal coverage rates can be inferred from this set of quantiles. Prediction intervals give a range of possible values within which the true value of the variable to forecast is expected to lie with a certain probability, that is, its nominal coverage rate [47]. To completely characterize a coverage rate, it is also necessary to define the way it should be centered on the probability density function [47]. The most common way is to center the prediction interval on the median (Fig. 1). Consequently, there is the same probability of risk below and above the median. We will implicitly use this convention in this chapter. Thus, the $(1 - \alpha)100\%$ central prediction interval is generated by taking the $\alpha/2$ quantile as the lower bound and the $1 - \alpha/2$ quantile as the upper bound. More precisely, a prediction interval with $(1 - \alpha)100\%$ nominal coverage rate is estimated by:

$$\widehat{PI}_{(1-\alpha)100\%} = \left[\hat{q}_{\tau=\alpha/2}, \hat{q}_{\tau=1-\alpha/2}\right] \tag{2}$$

2 Ensemble Prediction System (EPS)

In the realm of meteorology, an ensemble forecast is a set of forecasts that depicts the range of future weather possibilities. Each forecast of this ensemble is called a member. Thus, in the field of solar radiation, an ensemble forecast is a set of

predictions of the solar irradiance at the same horizon. The cumulative distribution function of the forecasts can be inferred from the ensemble by ranking the members in ascending order. For example, with an odd number of members, the median ($q_{\tau=0.5}$) is the middle-ranked member.

2.1 EPS and Numerical Weather Predictions (NWP)

With a NWP model, an ensemble forecast corresponds to a perturbed set of forecasts computed by slightly changing the initial conditions of the control run and of the modeling of unresolved phenomena [39]. An EPS allows representing the uncertainties of the prediction scheme. For example, the European Centre of Medium-range Weather Forecasts (ECMWF) provides an ensemble forecast from the Integrated Forecasting System (IFS) model. It consists in 1 control run and 50 "perturbed" members. Today, the main meteorological facilities (NCEP in the US, Met Office in the UK, Météo France, China Meteorological Association, etc.) make their own ensemble weather forecasts at a global scale (i.e., worldwide).

As the generation of ensemble forecasts by a NWP is extremely computationally intensive, their resolution is coarser than the deterministic forecasts. Table 1 provides a comparison between the resolutions of the deterministic forecasts and the ensemble forecasts of the operational system of the ECMWF that can be considered as the finest global NWP of the moment in terms of resolutions. As expected, the spatial resolution of the ensemble forecasts is two times coarser than the point forecasts.

With the THORPEX Interactive Grand Global Ensemble (TIGGE), a new step has been reached. Indeed, the TIGGE dataset consists of ensemble forecasts from 10 global NWP centers that are freely available for scientific research since October 2006 [7]. Thorey et al. [54] built probabilistic solar forecasts from a part of this huge ensemble forecast (i.e., 158 members) and compared them with GHI maps derived from satellite images.

When considering complex terrain, the coarse resolution of the global NWP models are not able to reliably take into account the small-scale phenomena. Regional or mesoscale NWPs were developed to produce high-resolution (until 1 km^2) weather forecasts for a limited area. The boundary conditions of the considered area are set

Table 1 Overview of the operational forecasts of the Global Horizontal Irradiance (GHI) provided by the ECMWF (IFS Cycle 43)

		HRES (Deterministic)	EPS (Ensemble)
Spatial resolution		0.1° × 0.1°	0.2° × 0.2°
Time step	T0 to T0 + 144 h	hourly	3 h
	T + 150 h to T0 + 240 h	6 h	6 h
	T + 246 h to T0 + 360 h	–	6 h
Number of analysis per day		2 (00–12 h)	2 (00–12 h)

by a global NWP and the regional model performs a downscaling. Some national meteorological services maintain mesoscale models that cover their countries (e.g., AROME for France, NAM for the United States, COSMO for several European countries, etc.). However, being freely available, the Weather Forecasting and Research (WRF) model [50] which is developed by NCAR is the most utilized mesoscale model by the solar forecasting community.

If mesoscale models are extensively used to provide deterministic solar forecasts, their use remains today almost unexplored for ensemble predictions in the field of solar energy. Nevertheless, several approaches, originated mainly from other research fields, are available but they do not totally respect the strict definition of an ensemble forecast given at the beginning of this subsection. First, it is possible to downscale the perturbed members of an EPS provided by a global NWP. MeteoSwiss uses this approach to provide probabilistic forecasts for the Alpine region of Europe. The mesoscale model COSMO is used to downscale the members of the ECMWF-EPS. Figure 2 provides a sketch of their approach. A second way to proceed is to gather the deterministic forecasts provided by different mesoscale models for the same area. Meteoblue uses this method in combination with global NWP-EPS for ten regions around the world [1]. Unfortunately, for these two first approaches, no comparison with measured solar data has been published. A last approach, the Lagged Average Forecast (LAF) [29], allows generating ensemble forecasts from the deterministic runs of a NWP. Each ensemble member of the LAF is computed with different initial

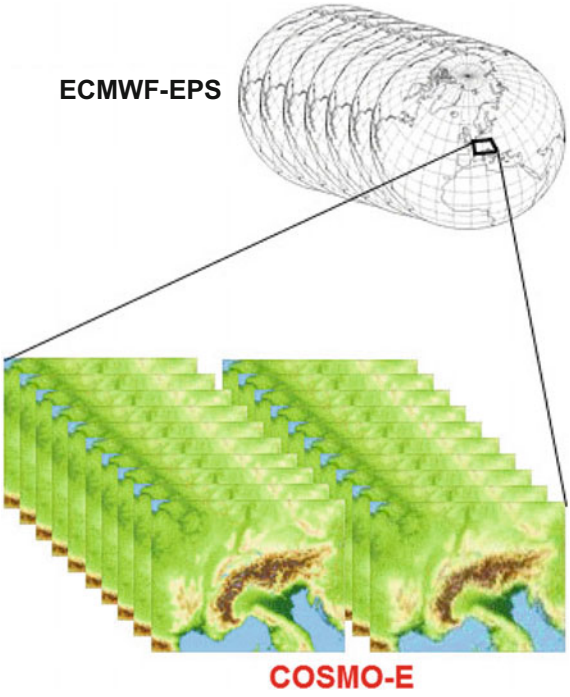


Fig. 2 The numerical weather forecasting models COSMO-E from MeteoSwiss, which are embedded within the ECMWF models. *Source* http://www.cosmo-model.org/

conditions corresponding to different initial time. This method has been set up with WRF to produce probabilistic solar forecasts in Japan [40].

Global and regional NWPs are designed to forecast a large variety of weather variables and they are not focused on the accurate generation of the different component of the solar radiation. As a consequence, probabilistic forecasts of the solar irradiance done the NWPs suffer from a lack of accuracy. Thus, raw ensemble forecasts are systematically refined by post-processing techniques.

2.2 Probabilistic Forecasts Derived from Post-processing of NWP

Two approaches are proposed in the literature to generate probabilistic forecasts from the refinement of the outputs of the NWPs. On one hand, the forecasted weather parameters of the deterministic run are used as explanatory variables to create a conditional distribution. On another hand, a transformation is applied to ensemble forecasts in order to improve their calibration.

2.2.1 From Deterministic to Probabilistic Forecasts

For the first approach, we will only give here a brief summary of the main methods already proposed for solar forecasting. As these methods contribute to a more generic framework in the field of probabilistic forecasting, details will be given in the following Sects. 3 and 4. Indeed, most of the models used to post-process the deterministic forecasts of a NWP are also used alone or in combination to provide probabilistic forecasts for various lead times (e.g. very short-term or long-term).

Lorenz et al. [41] proposed a parametric model assuming a normal distribution with a zero mean of forecast errors. In their method, only the standard deviation has to be forecasted to assess the whole distribution. They estimate the standard deviation in dependence on the cosine of solar zenith angle and on the forecasted clear sky index (from the NWP) by a polynomial function of fourth order. More sophisticated methods based on a non-parametric approach are most common and illustrated by the models used during the Global Energy Forecasting Competition 2014 (GEFCom 2014) [30]. The aim of the competition was to compute 99 quantiles of the power output of the PV fields on a rolling basis for 24 h ahead. 12 weather variables (including GHI) from the ECMWF-HRES were made available to the participants and were taken as explanatory variables inside their models. The main methods used were the quantile regression (see Sect. 4.2) [33], the gradient boosting techniques (see Sect. 4.4), and the quantile regression forest (see Sect. 4.3) [45]. An original method was also proposed by Alessandrini et al. [2] to post-process the forecasts of a mesoscale NWP (RAMS). They applied an analog ensemble approach

to a set of predicted weather variables (GHI, cloud cover, air temperature, etc.) to derive the quantiles of the forecast errors.

2.2.2 Calibration of Ensemble Forecasts

The second approach, which is less common than the first one, is based on the direct use of ensemble forecasts. The aim is to apply a statistical calibration to the PDF drawn by the initial ensemble forecasts in order to optimize a specific metric (see Sect. 5) used to assess the quality of probabilistic forecasts. Indeed, as well as having a coarse resolution, the ensemble forecasts from the NWPs are known to be underdispersive (see Fig. 8), i.e., they exhibit a lack of spread [39]. For this purpose, Sperati et al. [51] proposed two different methods already used in the realm of wind forecasting. Even if these methods cannot be considered as parametric, they are based on the characteristics of a normal distribution. Indeed, such a distribution is appealing because it can be assessed with only two parameters: the mean and the standard deviation that are, respectively, related to the average bias and the spread of the ensemble. Thus, in order to reproduce a Gaussian distribution, a logit function is first applied to each member of the ensemble. The following equation details the logit function:

$$e^* = log(\frac{k}{1-k}) \text{ with } k = \frac{e - e_{min}}{e_{max} - e_{min}} \tag{3}$$

e is a member of the ensemble forecast and e^* is the corresponding member after the logit transform. e_{min} and e_{max} are respectively the members with the minimum and the maximum values. The inverse logit function is then applied to the new ensemble produced by one of the calibration methods described below.

The first method of calibration, initially designed by Buizza et al. [12], set up a calibration of the spread of the ensemble. The aim is to fit the standard deviation of the new ensemble similar to the standard deviation calculated from a training set of past observed data.

To do so, a new ensemble $\widetilde{E}^* = (\widetilde{e}_1^*, ..., \widetilde{e}_n^*)$ is created from the ensemble $E^* = (e_1^*, ..., e_n^*)$ by modifying each member with the Variant Deficit (VD) procedure described by the Eqs. 4 and 5.

$$\widetilde{e}_i^* = \overline{e}^* + vd(e_i^* - \overline{e}^*) \tag{4}$$

The initial ensemble and the calibrated ensemble have the same size n. e_i^* and $\widetilde{e}_i^*$ are the members, respectively, of the initial and of the calibrated ensemble. Finally, $\overline{e}^*$ is the ensemble mean.

$$vd = \frac{\sqrt{\frac{1}{N}\sum_{j=1}^{N}(\overline{e}_j^* - o_j)^2}}{\frac{1}{N}\sum_{j=1}^{N}\sigma_j^*} \tag{5}$$

vd is the Variance Deficit coefficient that quantifies the lack of spread of the initial ensemble. o_j, σ_i^* and N represents, respectively, the observations, the standard deviation of the members of the ensemble and the size of the training data set.

The second method of calibration used by Sperati et al. [51] was first proposed by Gneiting et al. [22] and it is called Ensemble Model Output Statistic (EMOS). The aim of this method is to create a new Gaussian distribution from the logit transform of the ensemble forecast $E^* = (e_1^*, ..., e_n^*)$. The mean and the standard deviation of this calibrated distribution are obtained using two linear functions. This new PDF is drawn from the following normal distribution:

$$N\left(a + b_1 e_1^* + ... + b_n e_n^*,\ c + d\sigma^{*2}\right), \tag{6}$$

where σ^* is the standard deviation of E^*. A minimization of the Continuous Rank Probability Score (CRPS) allows estimating the coefficients a, b_1, ..., b_n, c and d. The CRPS can be considered as a probabilistic extension of the mean absolute error (MAE) usually used for assessing the accuracy of deterministic solar forecasts. Section 5.3 gives a detailed presentation of the CRPS. Finally, n equally spaced quantiles with probabilities $\frac{i}{1+n}$ (with $i = 1, ..., n$) are derived from the Cumulative Distribution Function (CDF) to build the calibrated ensemble.

We briefly present here, a last method of calibration based on the rank histogram and initially proposed by Hamill and Colucci [26] for forecasts of precipitation. Zamo et al. [58] applied this method to the Météo France's EPS (PEARP) to generate probabilistic solar forecasts. The rank histogram is a verification tool that assesses the consistency of ensemble forecasts. Section 5.2 gives a detailed presentation of this graphical tool. The aim of this last method is to build a calibrated CDF from the rank histogram derived from past forecasts and observations. The transformation of the initial ensemble forecasts lies on two principles:

- Between two ranks, the new CDF is built from a linear interpolation.
- As the rank histogram does not give any information about the tails of the distribution (below the first rank and above the last rank), specific functions have to be fixed to generate the tails of the new CDF.

The tails of an error distribution hold important quantiles (e.g., 0.05, 0.1, 0.9, and 0.95). Thus, the functions that will generate the tails must be in agreement with the physics of the considered parameters. For the case studied by Zamo et al. [58], linear functions have been used. Appendix of the article of Hamill and Colucci [26] gives details regarding the implementation of this procedure.

3 Parametric Models

Parametric models assume that the uncertainty associated with solar irradiance forecasts follows a known law of distribution (e.g., a normal law). As a consequence, this particular class of models needs only few parameters to predict the whole

distribution. They are attractive for operational purpose because they can run very fast with a low computational effort. This approach was used by the first operational solar forecasting system developed in Germany by Lorenz et al. [41].

3.1 Error Distribution of Point Forecasting Methods

When dealing with parametric models, the aim is to forecast a Probability Density Function (PDF) that follows a parametric law. Thus, the first question to ask is—Does the uncertainty of a forecasting system follow a known probabilistic law? Figs. 3 and 4 show the PDFs of the errors of two types of deterministic solar forecasts with different lead times: short-term from a NWP (3) and very short-term from a statistical model (4). These different cases clearly show that the uncertainty of these forecasting systems does not agree with the hypothesis of a normal distribution. Figure 3 presents the distribution of the errors of a NWP for three distinct sky conditions (clear, cloudy, and overcast). We can observe that the shape of the three resulting PDFs depends on the sky conditions. Thus, a unique parametric law will not be able to reproduce the uncertainty of the considered forecasting method.

3.2 Example of Forecasts with a Parametric Model: GARCH

Probabilistic forecasts of financial products are frequently generated using the Generalized Autoregressive Conditional Heteroscedasticity model (GARCH). First proposed by Bollerslev [6], the GARCH model is extensively documented in the field of econometrics. This time-series model is commonly associated with an autoregres-

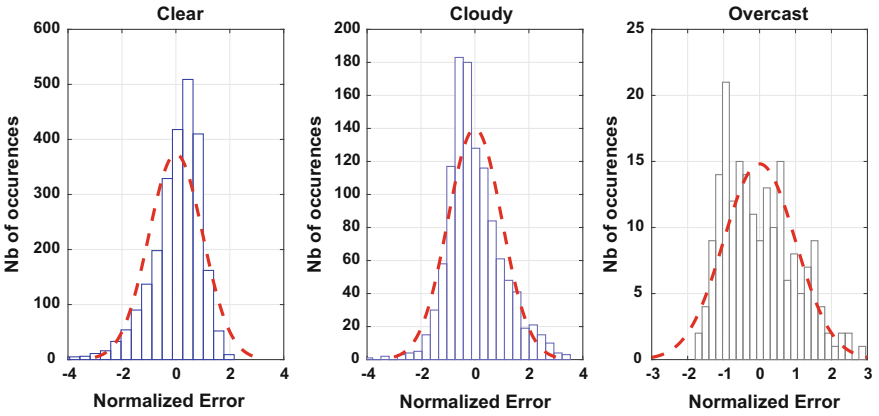


Fig. 3 PDF of the error of the hourly profile of day ahead forecasts provided by ECMWF-HRES for three different sky conditions and for the site of Saint-Pierre (21.34°S, 55.491°E), Reunion in 2012. Red dashed line represents the fitted normal PDF

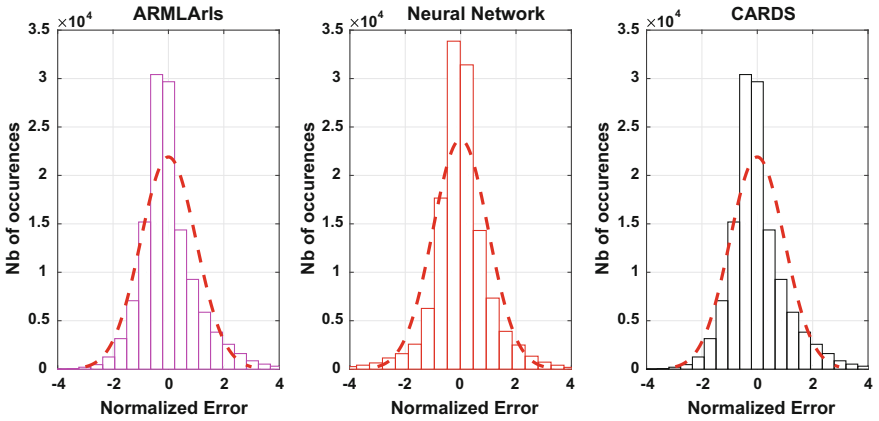


Fig. 4 PDF of the error of point forecasts generated by three-time series models (ARMArls, Neural Network and CARDS) and for very short-terms horizons (from 1 to 6 h). Red dashed line represents the fitted normal PDF. Illustrations taken from [15]

sive process that generates the main trend of the exchange market. As it assumes a normal distribution of the forecast errors, the GARCH model belongs to the class of parametric models. Its particularity comes from taking into account the heteroscedastic nature of a time series [18]. That is to say, the time series exhibits periods of low variability and periods of high variability. This behavior is also met for time series of clear sky indices, stable when the sky is clear or overcast and highly variable when the sky is partially cloudy.

In GARCH models, the conditional variance is a linear function of lagged squared error terms and also lagged conditional variance terms [53]. The general formulation of a GARCH(p, q) model, with p error terms, q conditional variance terms is given by the Eqs. 7 and 8 [55]. To clarify the equations, we used the following notation. y is the variable to predict, also called the predictand. x is the explanatory variable or a vector of explanatory variables. A variable with a hat (i.e., $\widehat{y}$) is a forecast or an estimate. h is the horizon of forecast also called the lead time. Finally, t denotes the time when the forecasts are generated.

$$\widehat{y}(t+h) = \varepsilon(t).\widehat{\sigma}(t+h), \tag{7}$$

where ε is a random variable uniformly distributed with a zero mean and a unitary variance.

$$\widehat{\sigma}^2(t+h) = \alpha_0 + \sum_{i=1}^{p} \alpha_i.\epsilon^2(t-i+1) + \sum_{j=1}^{q} \beta_i.\sigma^2(t-j+1) \tag{8}$$

with the error term $\epsilon(t) = \widehat{y}(t) - y(t)$.

An application of the econometric approach presented above to the solar field has been proposed by David et al. [16]. They implemented a combination of an AutoRegressive and Moving Average model (ARMA) with a GARCH model inside a recursive framework to generate probabilistic forecasts of the clear sky index. Their work highlights the strengths and the weaknesses of the use of parametric models.

As the GARCH model assumes a Gaussian distribution of the error of the underlying autoregressive process (i.e., an ARMA process in our study case [16]), we only need to forecast a mean and a standard deviation (8). Thus, this method requires only two parameters to estimate the whole PDF of a forecast. The main advantage of this parametric approach is based on its simplicity.

As shown in Figs. 3 and 4, PDFs of the error of different point forecasting methods do not seem to follow a normal law or any other laws of distribution. The distribution of the error depends on the model, on the sky conditions, and on various other parameters. As a consequence, a method based on a unique parametric law will not be able to reproduce a realistic distribution of the forecast errors. In the case studied in [16], the normal law systematically produces symmetrical prediction intervals around the median. Figure 3 shows an evidence while dealing with solar radiation. When the sky is clear, it is very uncommon to observe an even more clear sky and conversely with an overcast sky. Therefore, the distribution of the forecast errors of a clear sky or of an overcast sky cannot be symmetrical around the median. In addition, it is strongly different when the sky is partially cloudy. As a conclusion, parametric models suffer from a lack of calibration (see Sect. 5).

4 Nonparametric Models

A common way to produce probabilistic forecasts is to build models without imposing any parametric assumptions on the underlying distribution. Considering the issues highlighted by the parametric approach (see Sect. 3), most of the solar probabilistic forecasting systems are based on nonparametric models. In that way, they avoid the lack of calibration induced by a fixed law of distribution. However, their set up is generally more complex and their run is more computationally intensive. In this section, we will give an overview of some of the most common nonparametric methods in the field of solar forecasting.

4.1 Bootstrap

Bootstrap methods are widely used to estimate the statistical distribution of a quantity from a limited number of observations [17]. The following sentence, written by Singh and Xie [49], gives a good overview of the concept of bootstrap: “The idea behind bootstrap is to use the data of a sample study at hand as a ‘surrogate population’, for the purpose of approximating the sampling distribution of a statistic; i.e. to resample (with replacement) from the sample data at hand and create a large

number of 'phantom samples' known as bootstrap samples. The bootstrap method for a one-sample problem is summarized by the following three steps [17]:

1. Build the PDF $\widehat{F}$ of a sample X at hand with a finite number of n elements: $X = (x_1, x_2, ..., x_n)$.
2. Draw a "bootstrap sample" X^* with replacement of size n from the initial sample X. In other words, construct a new sample X^* by choosing randomly n elements inside the initial sample X. Thus, an element x_k of X can be repeated more than once in X^*.
3. Approximate the sampling distribution $\widehat{F}$ by the bootstrap distribution $\widehat{F}^*$.

With the current computational capacities, the bootstrap method presented above is frequently iterated thousands or tens of thousand times to approximate more accurately a PDF from a limited number of observations. Bootstrap methods are also commonly combined with time-series models to assess their prediction intervals [11]. In the field of solar forecasting, Grantham et al. [24] proposed to used the sieve bootstrap approach. This specific method of bootstrap was initially proposed by Bühlman [10] and extended to the forecasting of time series by Alonso et al. [3]. The sieve bootstrap enjoys a nice nonparametric property, being model-free in a class of linear processes. However, it must be stressed that the sieve bootstrap method only works in combination with linear time series models (AR, MA, ARMA, etc.). As an example of the application of the bootstrap to the solar forecasting, we reproduce here the methodology proposed by Grantham et al. [24].

Let us consider a time series $y(t)$. It can be described by an autoregressive (AR) process of order p. For a finite sample of $y(t)$, the residual $\varepsilon(t)$ of the process can be formulated as follows:

$$\varepsilon(t) = \sum_{k=0}^{p} -\alpha_k (y(t-k) - \overline{y}), \text{ with } \alpha_0 = -1. \tag{9}$$

$\overline{y}$ is the mean of $y(t)$, α_k is a vector of parameters and $\varepsilon(t)$ are i.i.d. random variables with zero mean and finite variance. Once the vector of parameters α_k is estimated, we compute the resulting centered residuals $\widetilde{\varepsilon}(t) = \varepsilon(t) - \overline{\varepsilon}$ and we construct their empirical distribution $\overline{F}_{\widetilde{\varepsilon}}(x)$ as follows:

$$\widehat{F}_{\widetilde{\varepsilon}}(x) = \frac{1}{n-p} \sum_{t=p+1}^{n} \mathbb{1}_{\{\widetilde{\varepsilon}(t) \leq x\}} \tag{10}$$

The indicator function $\mathbb{1}_{\{u\}}$ has the value of 1 if its argument u is true and 0 otherwise. We draw a resample $\varepsilon^*(t)$ of i.i.d. observations from the distribution $\widehat{F}_{\widetilde{\varepsilon}}$ of the centered residuals $\widetilde{\varepsilon}(t)$. Finally, we construct a sieve bootstrap sample of the forecasts $\widehat{y}^*(t)$ by the following recursion:

$$\widehat{y}^*(t) - \overline{y} = \sum_{k=1}^{p} \alpha_k (y^*(t-k) - \overline{y}) - \varepsilon^*(t) \tag{11}$$

In their method, Grantham et al. [24] also assume a conditional distribution of the forecast errors done by the AR model. Thus, they proposed to bin the centered residuals, according to two parameters of the solar path: sun elevation and hour angle. The resample is then done using only the centered residuals corresponding to the past observations, where the values of these two parameters are equivalent to the date and the time of the forecast.

4.2 Quantile Regression Methods

Quantile regression methods provide discrete quantile information about a variable to forecast, named predictand. The basic idea of these specific methods of regression is to estimate the parameters of functions that link the predictand quantiles with explanatory variables. Different types of functions have been previously tested within this framework: linear models [35, 36, 58], quantile regression Forest [43, 58], and neural networks [13, 52]. In this subsection, we will briefly present three of these methods: Linear Model in Quantile Regression (LMQR), Weighted Quantile Regression (WQR), and Quantile Regression Neural Network (NNQR).

A simple method based on linear functions (LMQR) was proposed by Koenker and Bassett [35]. This method provides, at time t, the τth quantiles $y_\tau(t)$ of the variable to forecast $y(t)$, establishing a linear relation with a vector of explanatory variables $x(t)$:

$$y_\tau(t) = \beta_\tau x(t) + \varepsilon, \tag{12}$$

where ε is a random error term.

In classical regression, the vector of parameters β_τ is estimated by minimizing the sum of squared errors. In quantile regression, conditional quantiles are estimated applying asymmetric weights to the mean absolute error. Koenker and Bassett [35] described a loss function ρ_τ to achieve this goal:

$$\rho_\tau(u) = \begin{cases} \tau u, \text{ if } u \geq 0 \\ (\tau - 1)u, \text{ if } u < 0 \end{cases} \tag{13}$$

Then the τth sample quantile of $y(t)$ may be defined thanks to any solution of the following minimization problem:

$$\widehat{\beta_\tau} = argmin_{(\beta)} \left[\sum_{i=1}^{N} \rho_\tau(y(i) - \beta x(i)) \right], \tag{14}$$

where N is the size of the considered sample, i.e., the number of pairs of explanatory variables x and of corresponding observed predictand y.

$\widehat{y_\tau}(t)$ is the τth quantile estimate, forecasted by the regression function $\widehat{\beta_\tau} x(t)$. The quantile regression provides the quantile distribution estimating each quantile separately. Calculating each quantile independently, could lead one to obtain quantile curves intersect, that is to say crossing quantiles ($\widehat{y_{\tau_1}}(t) > \widehat{y_{\tau_2}}(t)$ when $\tau_1 < \tau_2$). A CDF should be a monotonically increasing function with τ, so the quantile crossing problem should be corrected. For avoiding this crossing quantiles a simple rearrangement could be applied to the estimated quantiles. For example, Chernozhukov et al. [14] proposed to rank in ascending order the estimated quantiles and to permute the corresponding probability, τ if necessary.

Using exactly the same framework, a research team of the Danmark Tekniske Universitet (DTU) proposed a variant of the classical quantile regression named Weighted Quantile Regression (WQR) [4, 44]. This is a nonparametric method based on a Gaussian kernel smoother. Considering a sample of an explanatory variable $x_i (i = 1, ..., N)$ the kernel smoother k is defined as follows:

$$k(x_t, x_i) = \frac{e^{-\frac{1}{2}(\frac{x_t - x_i}{S_w})^2}}{\sum_{i=1}^{N} e^{-\frac{1}{2}(\frac{x_t - x_i}{S_w})^2}}, \text{ with the screen width } S_w \tag{15}$$

The problem is reduced to estimating the τth quantile $\widehat{y_\tau}(t)$ as a local constant for each possible value of y. The quantile $\widehat{y_\tau}(t)$ is then found by applying the classical quantile regression approach to the previous kernel smoother as follows:

$$\widehat{y_\tau}(t) = argmin_{(\gamma)} \left[\sum_{i=1}^{N} k(x_t, x_i) \cdot \rho_\tau (y(i) - \gamma) \right] \tag{16}$$

x_t is an explanatory variable that can be the point forecast itself. The screen width S_w is a tuning parameter that defines the smoothness of the Gaussian kernel. A suitable value should provide a good compromise between smoothness and overfitting of the Gaussian kernel. Unfortunately, there is no method to directly estimate this parameter and a cross-validation is thus necessary. This method is able to reduce drastically the computation cost in comparison with the other quantile regression methods. Indeed, in specific cases, it is sufficient to compute and store the conditional quantiles for a binned set of the explanatory variable.

Another alternative method to the classical quantile regression is the Quantile Regression Neural Networks (QRNN) based on the work of Taylor [52] and described by Cannon [13] for precipitation data. This method establishes a nonlinear relation between input–output pairs of data, without any previous model specifications. Based on a multilayer perceptron, the method can be implemented with the qrnn R-package [13]. The τth quantile forecasted by the regression is computed as follows:

$$\widehat{y_\tau}(t) = f \left(\sum_{j=1}^{m} \left[w_j \cdot g \left(\sum_{i=1}^{d} w_{ji} \cdot x_t + b_j \right) \right] + b^0 \right) \tag{17}$$

b_j and b^o are the input and output layer bias, respectively, $g()$ is the transfer function related to each neuron and $f()$ is the output transfer function. The output layer commonly uses a ramp activation function and, as solar radiation or clear sky index could not be a negative value, this function must have a left censored limit to zero. The weights of the NN are optimized during the training phase by minimizing a cost function [5]. To obtain the conditional quantiles, QRNN combines a nonlinear relation between input–output data and the quantile regression loss function (see Eq. 13). As the loss function used for the quantile regression is not differentiable in zero, qrnn package uses Huber norm to construct smooth approximations.

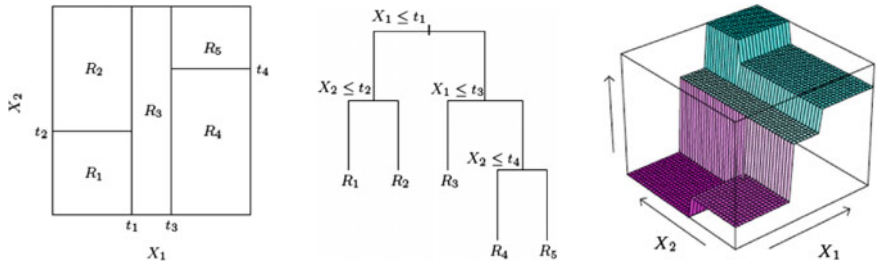
A key issue of the NN is the control of their complexity, i.e., the number of neurons of the hidden layer. Indeed, the number of hidden units affects the NN accuracy. If a NN obtains very good results for the training set but poor results for the testing data set, this NN is not establishing a generalized relation between input and output data. This problem is called overfitting and it is usually avoided by using regularization techniques for controlling NN number of units [27]. The regularization technique uses a new parameter called weight decay term. This parameter controls the contribution of the regularization term and penalizes large weights. To estimate the optimal weight decay parameter and the number of hidden units, cross-validation technique has to be used for each quantile. The training set must be divided at least in N-folds. For each training phase N-1 folds are used for training and the last fold for validation. Cross-validated errors were calculated for different models using an increasing number of hidden units. Once the optimal number of hidden units and weight decay value were selected, the model is trained again with the whole training dataset and finally used to forecast conditional quantiles of the testing dataset.

4.3 Quantile Regression Forest (QRF)

Quantile Regression Forest (QRF) [43] is an extension of random forest (RF) [27] to quantile regression. For regression problems, RF provides an estimation of the conditional mean of a response variable while QRF leads to the estimation of the conditional distribution function of Y, given $X = x$, i.e., $F(y|X = x)$. RF (and by extension QRF) is constructed by growing many binary decision trees called classification and regression trees (CART) [27].

4.3.1 Regression Tree

Regression trees are simple models that divide the input (or feature) space into a set of rectangular regions and then associate to each region a constant value (also called terminal-node mean value) corresponding to the average of the targets Y_i of the training set that fall in this region. A recursive greedy top-down algorithm is used to partition the input space into regions. The interested reader can refer to [27] for details regarding the stratification of the input space. Following the notation given



(a) Recursive partitioning of the feature space (b) Tree example (c) Tree perspective

Fig. 5 **a** Recursive binary splitting of the feature space. **b** Tree corresponding to the previous partition. **c** A perspective plot of the prediction surface. In this example, there is $d = 4$ splits corresponding to five regions. Illustrations taken from [27]

by [8], a tree is denoted by $T(\theta)$. θ is a random parameter vector that determines how a tree is grown, i.e., how the input space is stratified in terms of split variables, split locations, and terminal-node mean values. $T(\theta)$ partitions the input space into L distinct regions $\{R_l\}_{l=1}^{L}$. Using the N independent observations of the training set $(X_i, Y_i), i = 1, 2, \cdots, N$, the prediction $\hat{y}(x)$ (or equivalently the conditional mean of Y) of a single tree $T(\theta)$ for a new data point $X = x$ is obtained by averaging over the training observations that fall in the region $R_l(x, \theta)$ of the input space to which x belongs:

$$\hat{E}(Y|X = x) = \hat{y}(x) = \sum_{i=1}^{N} \omega_i(x, \theta) Y_i, \tag{18}$$

where the weights $\omega_i(x, \theta)$ are given by: $\omega_i(x, \theta) = \frac{\mathbb{1}_{\{X_i \in R_{l(x,\theta)}\}}}{\#\{j : X_j \in R_{l(x,\theta)}\}}$. The indicator function $\mathbb{1}_{\{u\}}$ has the value of 1, if its argument u is true and 0 otherwise. Figure 5 illustrates the construction of a regression tree to model a function of two variables (also called covariates) $y = f(X_1, X_2)$.

4.3.2 Random Forest

RF is a collection or committee of K trees that are built on bootstrapped training subsets. In addition, a random sample of input variables is considered at each split of the tree. Using RF, the conditional mean $E(Y|X = x)$ is approximated by the averaged prediction of K single trees:

$$\hat{E}(Y|X = x) = \hat{y}(x) = \sum_{i=1}^{N} \omega_i(x) Y_i, \tag{19}$$

where $\omega_i(x) = \frac{1}{K} \sum_{t=1}^{K} \omega_i(x, \theta_t)$.

4.3.3 Estimation of the CDF by the QRF Method

Just as $E(Y|X = x)$ is approximated by a weighted mean over the observations of Y, the CDF can be estimated by the weighted mean over the observations of $\mathbb{1}_{\{Y \leq y\}}$. Using the same weights as for random forests $\omega_i(x)$ defined above, the CDF can be approximated by

$$\hat{F}(y|X = x) = \sum_{i=1}^{N} \omega_i(x) \mathbb{1}_{\{Y_i \leq y\}}. \tag{20}$$

Finally, estimates $\hat{Y}_\tau$ of the conditional quantiles are obtained from the estimated distribution $\hat{F}(y|X = x)$. Unlike the LMQR method, the QRF technique does not suffer from the problem of crossing quantiles. However, on the other hand, QRF is not capable of estimating a quantile higher than the maximum clear sky index in the training sample.

4.4 The Gradient Boosting (GB) Technique

Boosting is a general approach that can be applied to many statistical learning methods for regression or classification [19, 27]. For regression problems, given a training data set, the goal is to find a function $f(x)$ such that a specified loss function is minimized. Boosting approximates $f(x)$ by an additive expansion of the form

$$\hat{f}(x) = \sum_{m=0}^{M} \beta_m h(x, \theta_m), \tag{21}$$

where the functions $h(x, \theta_m)$ are simply functions of x parameterized by θ_m. $h(x, \theta_m)$ are called "base learners" or "weak learners" [19]. The expansion coefficients β_m and the parameters θ_m are fit to the training data in a forward "stagewise" manner (i.e., without adjusting the previous expansion coefficients and parameters of the base learners that have already been added). Here, we restrict the application of boosting to the context of regression trees (i.e., the base learner $h(x, \theta)$ is a tree $T(\theta)$). For that purpose, boosting builds an ensemble of trees iteratively in order to optimize a loss function Ψ.

The generic gradient tree boosting algorithm [27] is depicted below. The training set contains N samples (X_i, Y_i), $i = 1, 2, \cdots, N$ and Ψ is the loss function to minimize. To compute the prediction for a point $X = x$, the GB algorithm follows the steps:

1. Initialize $f(x)$ to be a constant, $\hat{f}_0(x) = argmin_\gamma [\sum_{i=i}^{N} \Psi(Y_i, \gamma)]$
2. FOR $m = 1$ to M:

(a) Compute the negative gradient of the loss function Ψ (also called the pseudo-residuals $\tilde{y}_{im}$):

$$\tilde{y}_{im} = -\left[\frac{\partial \Psi(Y_i, f(X_i))}{\partial f(X_i)}\right]_{f(x)=\hat{f}_{m-1}(x)} \tag{22}$$

(b) Fit a tree $T(\theta)$ with d splits predicting the pseudo-residuals $\tilde{y}_{im}$ from covariates x_i. $T(\theta)$ partitions the input space into $L = d + 1$ distinct regions R_{lm}.

(c) Compute the optimal node predictions, $l = 1, 2, \cdots, L$

$$\gamma_{lm} = argmin_{\gamma}\left[\sum_{X_i \in R_{lm}} \Psi(Y_i, f_{m-1}(X_i) + \gamma)\right] \tag{23}$$

(d) Update the estimates of $f(x)$:

$$\hat{f}_m(x) = \hat{f}_{m-1}(x) + \nu\gamma_{lm}\mathbb{1}_{\{x \in R_{lm}\}} \tag{24}$$

3. ENDFOR

In Eq. 24, ν is called the shrinkage parameter. This parameter controls the rate at which boosting learns. Gradient boosting models must be finely tuned to prevent overfitting. Some tuning parameters or hyperparameters are adjustable by the users to control the model's complexity, including M, ν, d:

- The number of trees (or iterations) M. Boosting can overfit if M is too large.
- The shrinkage parameter ν. Typical values are 0.01 or 0.001, and the right choice can depend on the problem. Smaller values of ν require larger numbers of iterations (M) to converge.
- The number d of splits (also called the interaction depth) in each tree, which controls the complexity of the boosted ensemble.

The Stochastic gradient boosting [19] is a variant of the *generic* algorithm described above. It consists in taking, at each iteration of the process, a subsample of the data drawn at random (without replacement) from the full training database. Line 2b (fitting of the tree) and Eq. 24 (model update) of the GB algorithm make use of this subsample instead of the full training data set. [19] showed that the introduction of this randomization step improves the accuracy of the GB algorithm. Estimates of the conditional quantiles are obtained by plugging the quantile loss function ρ_τ in the preceding algorithm (i.e. $\Psi = \rho_\tau$).

5 Assessment of Probabilistic Forecasts

Three main properties characterize the quality of a probabilistic forecasting system namely reliability, sharpness, and resolution.

- **Reliability** or calibration refers to the statistical consistency between the forecasts and the observations. As noted by [47], reliability is seen as a primary requirement when verifying probabilistic forecasts, since a lack of reliability would introduce a systematic bias in subsequent decision-making.
- In the atmospheric sciences literature [32, 57], the **sharpness** property characterizes the ability of a forecasting system to produce forecasts that are able to deviate from the climatological mean probabilities of the variable to predict (called predictand). However, some authors like [20, 47] proposed a pragmatic approach, where the sharpness property refers to the concentration of the predictive distributions. In the following subsections, we used this second meaning of the sharpness.
- **Resolution** is related to the capacity of a forecasting system to generate different predictive distributions. The meaning of the resolution attribute consists of evaluating the ability of the forecast system to issue different predictive distributions not only in relation with the level of the predict and but also in relation with the external forecast conditions [47].

For instance, in the case of GHI, the level of uncertainty may vary according to the sun's position in the sky (see, for instance, the work of [24]). Notice also that reliability can be improved by statistical techniques called calibration techniques [22], whereas this is not possible for sharpness and resolution.

In the weather verification community, several diagnostic tools are used to evaluate these required properties of reliability, resolution, and sharpness. One can cite among others the reliability diagram [46, 57] and the rank histogram [25, 57] for assessing the reliability property.

The sharpness attribute can be evaluated through the use of sharpness diagrams. For instance, [20] proposed to summarize the distribution of the width of the interval forecasts with box-plots while others like [34, 47] summarized sharpness with only the mean of the width of the prediction interval. More precisely, [47] proposed $\delta-$diagrams that plot the mean of the width of the prediction interval in relation to the nominal coverage rate. Similarly, [34, 47] used a metric called PINAW for prediction interval normalized average width. Care must be taken when one analyzes the sharpness diagrams. This visual inspection should be done in relation to the reliability analysis. Indeed, as mentioned by [57], anyone can produce sharp forecasts, but the difficult task is to ensure that these forecasts correspond well to the subsequent observations. Put differently, [20] stressed that a consistent evaluation framework should be based on the paradigm of maximizing the sharpness of the predictive distributions subject to calibration. It must be stressed also that no forecast system can reduce to zero the uncertainty due to the chaotic nature of the atmosphere.

In addition to these tools that permit to visually assess the quality of a forecasting system, a metric called continuous ranked probability score (CRPS) [28] is used to objectively quantify the quality of the probabilistic forecasts. The CRPS is a proper score [21], which means that this metric ensures that the best forecasts are given the best score value. Furthermore, this type of scoring metric is attractive as it provides a simultaneous assessment of reliability and sharpness. In other words, the CRPS provides an evaluation of the global skill of the probabilistic models.

Finally, a comparison with a naive reference model can give a useful added value about the efficiency of a model. The persistence ensemble model can be considered as a suitable benchmark model for probabilistic solar forecasts evaluation because it has already been used inside various works [2, 16, 23, 24, 36]. The members of the persistence ensemble are derived from the $N + 1$ last measurements. Thus, the PDF of the forecasted variable $\widehat{y}_{t+h}$ at time t with a lead time h is

$$PDF(\widehat{y}_{t+h}) = PDF(y_t, y_{t-1}, ..., y_{t-N}) \tag{25}$$

5.1 Reliability Diagrams

The reliability diagram is a graphical verification display used to verify the reliability component of a probabilistic forecast system. For instance, a probabilistic forecasting system based on quantile forecasts is reliable if, statistically, the nominal proportions of the quantile forecasts are equal to the proportions of the observed value. In other words, over a testing set of significant size, the difference between observed and nominal probabilities should be as small as possible [46]. This representation is appealing because the deviations from perfect reliability (the diagonal) can be visually assessed [46]. However, due to the finite number of pairs of observation/forecast and also due to possible serial correlation in the sequence of forecast-verification pairs, it is not expected that observed proportions lie exactly along the diagonal, even if the density forecasts are perfectly reliable. [46] proposed a method to add consistency bars to the reliability diagram. This adding of consistency bars to the reliability diagrams may help the user to have more credibility in his (possibly subjective) judgment regarding reliability of the different models. Figure 6 shows an example of reliability diagram with consistency bars. In this example, the forecasts cannot be considered as reliable because the line corresponding to the forecasts falls out of the consistency bars.

5.2 Rank Histograms

The rank histogram [57] is another graphical tool for evaluating ensemble forecasts consistency. Rank histograms are useful for determining the statistical consistency of the ensemble, that is, if the observation being predicted looks statistically just like another member of the forecast ensemble [57]. A necessary condition for ensemble consistency is an appropriate degree of ensemble dispersion leading to a flat rank histogram [57]. In other words, a flat rank histogram shows that the members of an ensemble system are statistically indistinguishable from the observations. If the ensemble dispersion is consistently too small (underdispersed ensemble), then the observation (also called the verification sample) will often be an outlier in the distribution of ensemble members. This will result in a rank histogram with a U-shape.

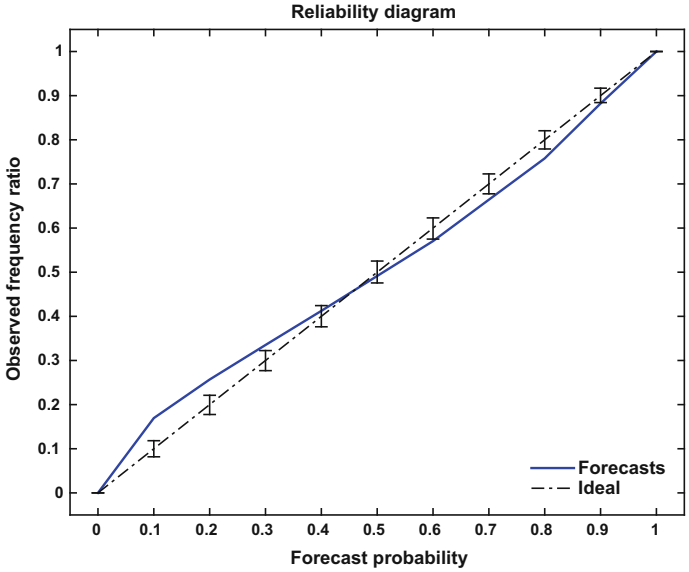


Fig. 6 Example of reliability diagram with consistency bars

Conversely, if the ensemble dispersion is consistently too large (overdispersed ensemble) then the observation may too often be in the middle of the ensemble distribution. This will give a rank histogram with a hump shape.

In addition, asymmetric rank histograms may suggest that ensemble may possess some unconditional biases. Ensemble bias can be detected from overpopulation of either the smallest ranks, or the largest ranks, in the rank histogram (Fig. 7). An overforecasting bias will correspond to an overpopulation of the smallest ranks while an underforecasting bias will overpopulate the highest ranks. As a consequence, rank histograms can also reveal deficiencies in ensemble calibration or reliability [57].

Again, care must be taken when analyzing rank histograms when the number of verification samples is limited. In addition, as demonstrated by [25], a perfect rank histogram does not mean that the corresponding EPS is reliable.

To obtain a verification rank histogram, one needs to find the rank of the observation when pooled within the ordered members of the ensemble. Then, a histogram shows the frequency of apparition of the rank of the observation. Thus, for a number of members M, the number of ranks of the histogram is $M + 1$. If the consistency condition is met, this histogram of verification ranks will be uniform with theoretical relative frequency of $\frac{1}{M+1}$. Figure 8 plots the rank histograms of the NWP models gathered in the TIGGE (see Sect. 2). As shown by Fig. 8 the NWPs lead to under-dispersed ensemble (i.e., over-confident models).

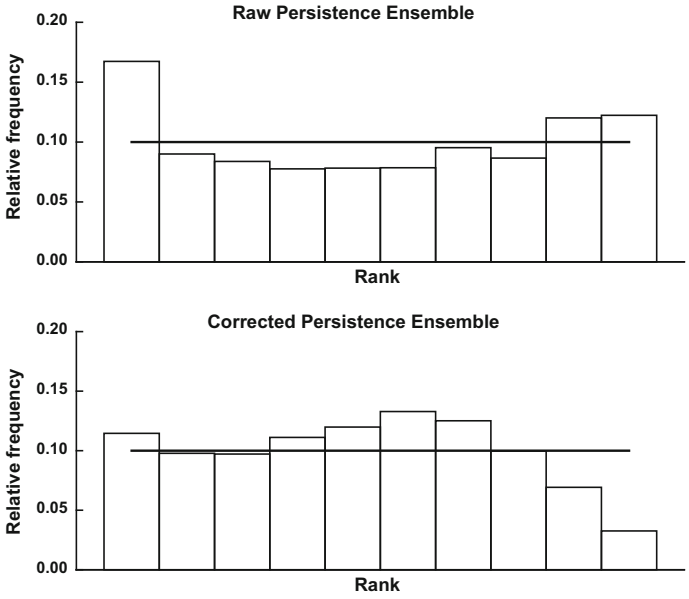
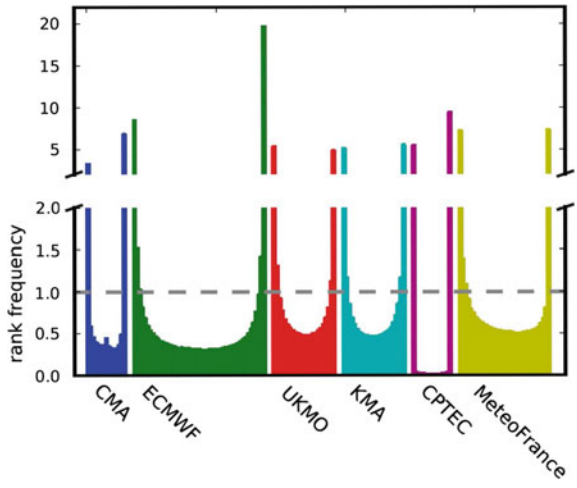


Fig. 7 Example of rank histograms: persistence ensemble (9 members) and persistence ensemble corrected by the Variance Deficit method (see Sect. 2.2). Forecasts of hourly GHI with a 6 h lead time for Oahu, Hawaii [48]

Fig. 8 Rank histograms of ensemble solar forecasts provided by six global NWPs and available in the TIGGE. Illustration taken from [54]



5.3 CRPS

The CRPS measures the difference between the predicted and observed cumulative distributions functions (CDF) [28]. The formulation of the CRPS is

$$\text{CRPS} = \frac{1}{N}\sum_{i=1}^{N}\int_{-\infty}^{+\infty}[\hat{P}^{i}_{fcst}(x) - P^{i}_{x_0}(x)]^2 dx, \tag{26}$$

where $\hat{P}^{i}_{fcst}(x)$ is the predictive CDF of the variable of interest x (here GHI) and $P^{i}_{x_0}(x)$ is a cumulative-probability step function that jumps from 0 to 1 at the point where the forecast variable x equals the observation x_0 (i.e., $P_{x_0}(x) = \mathbb{1}_{\{x \geq x_0\}}$). The squared difference between the two CDFs is averaged over the N ensemble forecast/observation pairs. The CRPS has the same dimension as the forecasted variable. The CRPS is negatively oriented (smaller values are better), and it rewards concentration of probability around the step function located at the observed value [57]. Thus, the CRPS penalizes lack of sharpness of the predictive distributions as well as biased forecasts (Fig. 9).

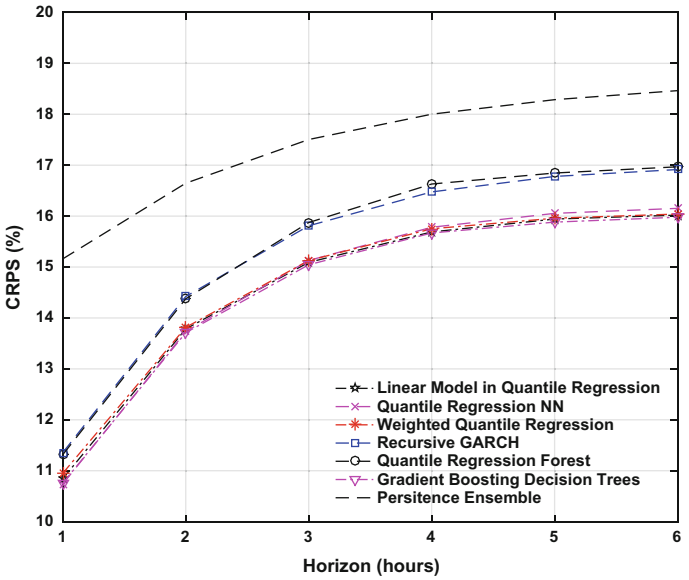


Fig. 9 Relative CRPS as a function of the lead time. This plot presents a comparison of eight models that provide probabilistic forecasts of the GHI at six locations with different climates (Two sites in continental US, two sites in La Reunion, Hawaii, and Guadeloupe). Illustration taken from [15]

6 Conclusions

Considering an increasing rate of solar renewables in the energy mix, probabilistic forecasts of the solar irradiance are useful for grid operations and storages management. Indeed, they allow assessing the risk associated to future decisions. Depending on the forecast horizon, different methods to generate forecasts with prediction intervals are already available. For short-term (from 6 h and up to several days ahead), the methods based on the post-processing of Ensemble Prediction Systems (EPS) of Numerical Weather Prediction models (NWP) are the most suitable. Regarding very short-term (from several minutes to several hours ahead), statistical approaches combined with point forecasts are the most promising. When dealing with complex terrains, a fine spatial resolution is needed to take into account the high variability of the climate conditions. Therefore, downscaling methods or statistical refinement are necessary to achieve this goal.

Finally, to ensure the quality of probabilistic forecasts, it is of great importance to assess their different properties (i.e., reliability, resolution, and sharpness). The first aspect to check is the reliability. Indeed, even if one prefers to have sharp prediction intervals, it is more important that the distribution of the forecasts corresponds well to the observations. The reliability diagram and the rank histogram are designed for this purpose. In addition to these graphical tools, the CRPS provides an evaluation of the global skill of probabilistic models.

References

1. Meteoblue, multimodel ensemble forecast for Boston (2017). https://www.meteoblue.com/
2. Alessandrini S, Delle Monache L, Sperati S, Cervone G (2015) An analog ensemble for short-term probabilistic solar power forecast. Appl Energy 157:95–110. https://doi.org/10.1016/j.apenergy.2015.08.011
3. Alonso AM, Pea D, Romo J (2002) Forecasting time series with sieve bootstrap. J Stat Plan Inference 100(1):1–11. https://doi.org/10.1016/S0378-3758(01)00092-1
4. Bacher P, Madsen H, Nielsen HA (2009) Online short-term solar power forecasting. Sol Energy 83(10):1772–1783. https://doi.org/10.1016/j.solener.2009.05.016
5. Bishop CM (1995) Neural networks for pattern recognition. Oxford university press edn, Oxford
6. Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. J Econom 31(3):307–327. https://doi.org/10.1016/0304-4076(86)90063-1
7. Bougeault P, Toth Z, Bishop C, Brown B, Burridge D, Chen DH, Ebert B, Fuentes M, Hamill TM, Mylne K, Nicolau J, Paccagnella T, Park YY, Parsons D, Raoult B, Schuster D, Dias PS, Swinbank R, Takeuchi Y, Tennant W, Wilson L, Worley S (2010) The THORPEX interactive grand global ensemble. Bull Am Meteorol Soc 91(8):1059–1072
8. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
9. Brooks MJ, Du Clou S, Niekerk VLW, Gauche P, Leonard C, Mouzouris MJ, Meyer R, Van der Westhuizen N, Dyk VEE, Vorster FJ (2015) SAURAN: a new resource for solar radiometric data in Southern Africa. J Energy South Afr **26**(1)
10. Bühlmann P (1997) Sieve bootstrap for time series. Bernoulli 3(2):123–148. https://doi.org/10.2307/3318584

11. Bühlmann P (2002) Bootstraps for time series. Stat Sci 17(1):52–72. https://doi.org/10.1214/ss/1023798998
12. Buizza R, Richardson DS, Palmer TN (2003) Benefits of increased resolution in the ecmwf ensemble system and comparison with poor-man's ensembles. Q J R Meteorol Soc 129(589):1269–1288. https://doi.org/10.1256/qj.02.92
13. Cannon AJ (2011) Quantile regression neural networks: implementation in R and application to precipitation downscaling. Comput Geosci 37(9):1277–1284. https://doi.org/10.1016/j.cageo.2010.07.005
14. Chernozhukov V, Fernandez-Val I, Galichon A (2010) Quantile and probability curves without crossing. Econometrica 78(3):1093–1125. https://doi.org/10.3982/ECTA7880
15. David M, Mazorra L, Lauret P (in press) Comparison of intraday probabilistic forecasting of solar irradiance using only endogenous data. Int J Forecast 19(4):299–311
16. David M, Ramahatana F, Trombe P, Lauret P (2016) Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models. Sol Energy 133:55–72. https://doi.org/10.1016/j.solener.2016.03.064
17. Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7(1):1–26. https://doi.org/10.1214/aos/1176344552
18. Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica 50(4):987–1007
19. Friedman JH (2002) Stochastic gradient boosting. Comput Stat Data Anal 38(4):367–378
20. Gneiting T, Balabdaoui F, Raftery AE (2007) Probabilistic forecasts, calibration and sharpness. J R Stat Soc Ser B (Stat Methodol) 69(2):243–268
21. Gneiting T, Raftery AE (2007) Strictly proper scoring rules, prediction, and estimation. J Am Stat Assoc 102(477):359–378
22. Gneiting T, Raftery AE III, Westveld III AH, Goldman T (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. Mon Weather Rev 133(5):1098–1118. https://doi.org/10.1175/MWR2904.1
23. Golestaneh F, Pinson P, Gooi HB (2016) Very short-term nonparametric probabilistic forecasting of renewable energy generation with application to solar energy. IEEE Trans Power Syst 31(5):3850–3863. https://doi.org/10.1109/TPWRS.2015.2502423
24. Grantham A, Gel YR, Boland J (2016) Nonparametric short-term probabilistic forecasting for solar radiation. Sol Energy 133:465–475. https://doi.org/10.1016/j.solener.2016.04.011
25. Hamill TM (2001) Interpretation of rank histograms for verifying ensemble forecasts. Mon Weather Rev 129(3):550–560
26. Hamill TM, Colucci SJ (1998) Evaluation of EtaRSM ensemble probabilistic precipitation forecasts. Mon Weather Rev 126(3):711–724. https://doi.org/10.1175/1520-0493(1998)126<0711:EOEREP>2.0.CO;2
27. Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. Springer series in statistics. Springer, New York
28. Hersbach H (2000) Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather Forecast 15(5):559–570
29. Hoffman RN, Kalnay E (1983) Lagged average forecasting, an alternative to monte carlo forecasting. Tellus A 35A(2):100–118. https://doi.org/10.1111/j.1600-0870.1983.tb00189.x
30. Hong T, Pinson P, Fan S, Zareipour H, Troccoli A, Hyndman RJ (2016) Probabilistic energy forecasting: global energy forecasting competition 2014 and beyond. Int J Forecast 32(3):896–913. https://doi.org/10.1016/j.ijforecast.2016.02.001
31. Iqbal M (1983) Chapter 6 - Solar spectral radiation under coudless skies. In: An Introduction to solar radiation. Academic Press, pp 107–168. https://doi.org/10.1016/B978-0-12-373750-2.50011-2
32. Jolliffe IT, Stephenson DB (2003) Forecast verification: a practitioner's guide in atmospheric science. Wiley, Chichester
33. Juban R, Ohlsson H, Maasoumy M, Poirier L, Kolter JZ (2016) A multiple quantile regression approach to the wind, solar, and price tracks of GEFCom2014. Int J Forecast 32(3):1094–1102. https://doi.org/10.1016/j.ijforecast.2015.12.002

34. Khosravi A, Nahavandi S, Creighton D (2013) Prediction intervals for short-term wind farm power generation forecasts. IEEE Trans Sustain Energy 4(3):602–610
35. Koenker R, Bassett G (1978) Regression quantiles. Econometrica 46(1):33–50. https://doi.org/10.2307/1913643
36. Lauret P, David M, Pedro H (2017) Probabilistic solar forecasting using quantile regression models. Energies 10(10):1591. https://doi.org/10.3390/en10101591
37. Lauret P, Lorenz E, David M (2016) Solar forecasting in a challenging insular context. Atmosphere 7(2):18. https://doi.org/10.3390/atmos7020018
38. Lauret P, Voyant C, Soubdhan T, David M, Poggi P (2015) A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol Energy 112:446–457. https://doi.org/10.1016/j.solener.2014.12.014
39. Leutbecher M, Palmer TN (2008) Ensemble forecasting. J Comput Phys 227(7):3515–3539. https://doi.org/10.1016/j.jcp.2007.02.014
40. Liu Y, Shimada S, Yoshino J, Kobayashi T, Miwa Y, Furuta K (2016) Ensemble forecasting of solar irradiance by applying a mesoscale meteorological model. Sol Energy **136**(Supplement C):597–605. https://doi.org/10.1016/j.solener.2016.07.043
41. Lorenz E, Hurka J, Heinemann D, Beyer HG (2009) Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. IEEE J Sel Top Appl Earth Obs Remote Sens 2(1):2–10. https://doi.org/10.1109/JSTARS.2009.2020300
42. Lorenz EN (1963) Deterministic nonperiodic flow. J Atmos Sci 20(2):130–141. https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2
43. Meinshausen N (2006) Quantile regression forests. J Mach Learn Res 7:983–999
44. Mller JK, Nielsen HA, Madsen H (2008) Time-adaptive quantile regression. Comput Stat Data Anal 52(3):1292–1303. https://doi.org/10.1016/j.csda.2007.06.027
45. Nagy GI, Barta G, Kazi S, Borbly G, Simon G (2016) GEFCom2014: probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. Int J Forecast 32(3):1087–1093. https://doi.org/10.1016/j.ijforecast.2015.11.013
46. Pinson P, McSharry P, Madsen H (2010) Reliability diagrams for non-parametric density forecasts of continuous variables: accounting for serial correlation. Q J R Meteorol Soc 136(646):77–90
47. Pinson P, Nielsen HA, Mller JK, Madsen H, Kariniotakis GN (2007) Non-parametric probabilistic forecasts of wind power: required properties and evaluation. Wind Energy 10(6):497–516. https://doi.org/10.1002/we.230
48. Sengupta M, Andreas A (2010) Oahu solar measurement grid (1-year archive): 1-second solar irradiance, Oahu (Data). https://doi.org/10.5439/1052451
49. Singh K, Xie M (2017) Bootstrap: a statistical method. http://www.stat.rutgers.edu/home/mxie/rcpapers/bootstrap.pdf
50. Skamarock C, Klemp B, Dudhia J, Gill O, Barker D, Duda G, Huang XY, Wang W, Powers G (2008) A description of the advanced research WRF version 3. https://doi.org/10.5065/D68S4MVH
51. Sperati S, Alessandrini S, Delle Monache L (2016) An application of the ECMWF Ensemble Prediction System for short-term solar power forecasting. Sol Energy 133:437–450. https://doi.org/10.1016/j.solener.2016.04.016
52. Taylor JW (2000) A quantile regression neural network approach to estimating the conditional density of multiperiod returns. J Forecast 19(4):299–311. https://doi.org/10.1002/1099-131X(200007)19:4<299::AID-FOR775>3.0.CO;2-V
53. Taylor JW (2004) Volatility forecasting with smooth transition exponential smoothing. Int J Forecast 20(2):273–286. https://doi.org/10.1016/j.ijforecast.2003.09.010
54. Thorey J, Mallet V, Chaussin C, Descamps L, Blanc P (2015) Ensemble forecast of solar radiation using TIGGE weather forecasts and HelioClim database. Sol Energy **120**(Supplement C):232–243. https://doi.org/10.1016/j.solener.2015.06.049
55. Tsay RS (2005) Analysis of financial time series, 2nd edn. Wiley, Hoboken
56. Weart SR (2008) The discovery of global warming, rev. and expanded edn. New histories of science, technology, and medicine. Harvard University Press, Cambridge. OCLC: 605622963

57. Wilks DS (2009) Extending logistic regression to provide full-probability-distribution mos forecasts. Meteorol Appl 16(3):361–368
58. Zamo M, Mestre O, Arbogast P, Pannekoucke O (2014) A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production. Part II: probabilistic forecast of daily production. Sol Energy 105:804–816. https://doi.org/10.1016/j.solener.2014.03.026

Solar Radiation Maps

Felipe Díaz, Gustavo Montero and Luis Mazorra-Aguiar

Abstract Solar maps are very interesting tools to describe the characteristics of a region from the solar radiation point of view, and can be applied in atmospheric sciences and for energy engineering. To make them possible, a solar radiation numerical model is proposed. This one allows us to estimate radiation values on any point on earth. The model takes into account the terrain surface conditions and the cast shadows. The procedure uses 2-D adaptive triangles meshes built refining according to surface and albedo characteristics. Solar irradiance values are obtained for clear sky conditions. Using clear sky index as a conversion factor, real sky values are computed in terms of irradiance or irradiation with a desired time step. Finally, the solar radiation maps are obtained for all the domain.

1 Introduction

Solar radiation is important in several disciplines like agronomy, meteorology, medicine, or engineering. Knowledge about solar radiation is essential for solar power generation and we all know the growing importance of these renewable energies. In literature, we can find mainly two kinds of spatial solar radiation models:

- Those obtained through satellite observation [1]
- Those obtained using geometrical, physical, and atmospheric conditions [2–6]

F. Díaz · G. Montero · L. Mazorra-Aguiar (✉)
University Institute for Intelligent Systems and Numerical Applications in Engineering, University of Las Palmas de Gran Canaria, Edificio Central del Parque Tecnológico, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain
e-mail: luis.mazorra@ulpgc.es

F. Díaz
e-mail: felipe.diaz@ulpgc.es

G. Montero
e-mail: gustavo.montero@ulpgc.es

R. Perez (ed.), *Wind Field and Solar Radiation Characterization and Forecasting*, Green Energy and Technology, https://doi.org/10.1007/978-3-319-76876-2_10

For the purpose of this chapter, we will focus on the second ones. These models take into account the interaction between radiation, Earth's atmosphere and terrain, in other words [5, 6]:

1. Geometry of Earth (latitude, solar hour angle,...)
2. Terrain characteristics (surface inclination and orientation, elevation, albedo)
3. Atmospheric attenuation caused by gases, particles, and clouds

The model with no clouds attenuation gives clear sky irradiance values. On the contrary, when attenuation is considered, computed values will be those of real sky irradiance. Some GIS-based models [5, 6] are being used with an important computational cost when dealing with high accuracy. Starting from these ones, other numerical models have been developed with a lower computational cost [7, 8]. To reach this goal, the geographical domain is discretized using a 2-D adaptive mesh as described in Chap. 3 having as a reference the elevation, the inclination, and the albedo of the terrain.

Another important issue to be considered is the shadowing effect that appears over a surface. Many researchers like Zakšek et al. [9] or Niewienda et al. [10], who propose a geometrical shading coefficient to compute the effect of cast shadows, have studied this geometrical problem. As a result of the use of a regular grid, a high computational cost is obtained. Other researchers [11] do not consider solid surfaces so that higher density of sample points is needed. Recently, new approaches [12] for computing shadows have been developed.

2 Terrain and Shadows Modeling

2.1 Terrain Mesh

The problem of computing solar radiation over the earth's surface begins with the discretization of the terrain. So, first, we need to define a geographical domain where computations will be done. As said above in this book, we have faced the problem on complex terrains. In Chap. 3, the process to build the meshes for the terrain has been explained. An adaptive method for mesh refinement and derefinement can be applied using two different parameters, one for orography and another for albedo. Different meshing strategies may be used as can be seen in Chap. 3.

As grid examples, we present several domains where we have applied some computations explained in this chapter. In Fig. 1, an adaptive mesh for Gran Canaria Island is shown. On the other hand, Fig. 2 shows a mesh for the Tatra Mountains between Poland and Slovakia. Both are fine meshes that describe quite well the orography. However, for solar radiation purposes, coarser meshes can be used. In fact, for Gran Canaria Island (1560 km^2), good results are obtained with a 5866 nodes and 11683 triangles mesh.

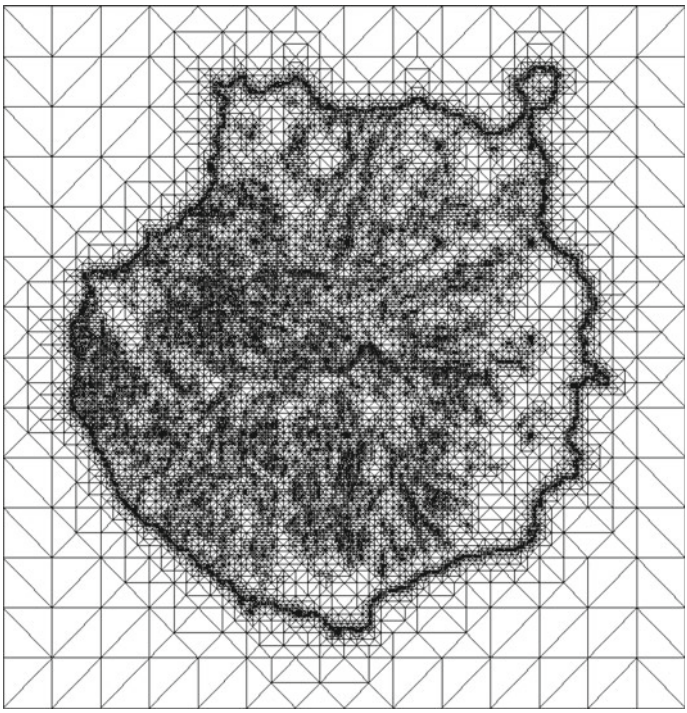


Fig. 1 Adaptive mesh for Gran Canaria Island

On the other hand, the shadowing problem is a geometrical one. The approach to this issue should take into account the cast shadows by the complex orography in each time step, and the shadows cast by the clouds on the terrain.

2.2 Shadowing by Orography

It is easy to think of a general technique to detect which triangles in the mesh are shadowed in an instant of time [7, 8]. A triangle will be under a cast shadow when, looking at the mesh from the Sun, we can find a triangle that covers it. To do that, a reference system with the East in the x-axis direction, and the North in the y-axis direction (see Fig. 3) is used. To determine the triangles that are behind another one, seen from the Sun, the reference system is changed to that of x', y', and z', with z' in the direction of the beam radiation. Now the domain mesh has to be projected on the $x'y'$ plane to find out which triangles are overlapped (Fig. 4).

To build this reference system transformation, we need to know the position of the Sun respect to the Earth, for each time step. This function is named the solar vector, and has been studied by many authors [13, 14]. For every single time step,

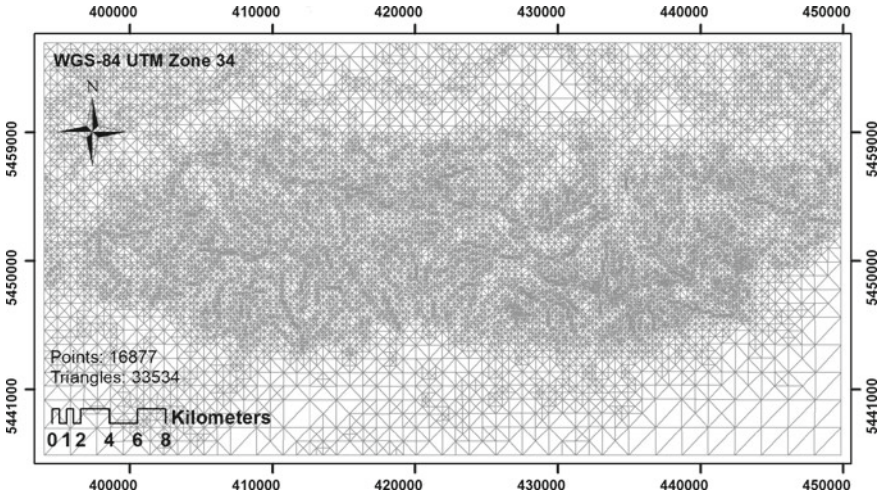


Fig. 2 Adaptive mesh for Tatra Mountains

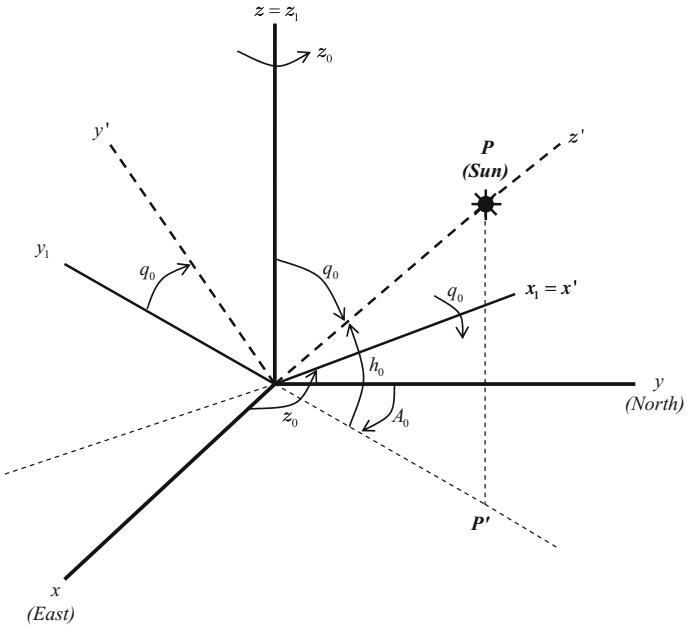


Fig. 3 Reference system

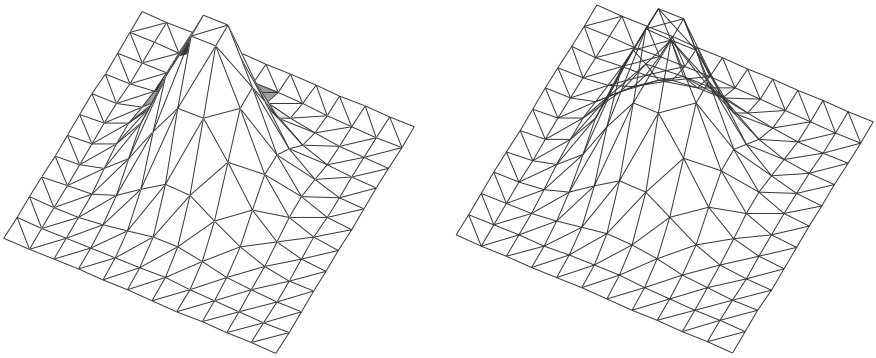


Fig. 4 Mesh from the Sun and projected one

all the triangles need to be analyzed and assigned a shading level which is called the light factor, L_f, and that allows the user to diminish the solar radiation values in the shadowed triangle, according to the triangle proportion that is shadowed by other mesh elements. A triangle Δ has a light factor computed as

$$L_f = \frac{n_{wp} - i}{n_{wp}} \tag{1}$$

with $i = 0, 1, \ldots, n_{wp}$ being the number of warning points inside other triangles that are in front of Δ. These *warning points* are the ones used to check which parts of a triangle are shadowed. The factor will be applied to the estimation of diffuse and beam irradiance. To implement this procedure, a 4-T Rivara's refinement algorithm [15] is applied every time step to the intersected triangles (see Fig. 5). The number of *warning points* depends upon the need of accuracy in the shadows determination process on the terrain.

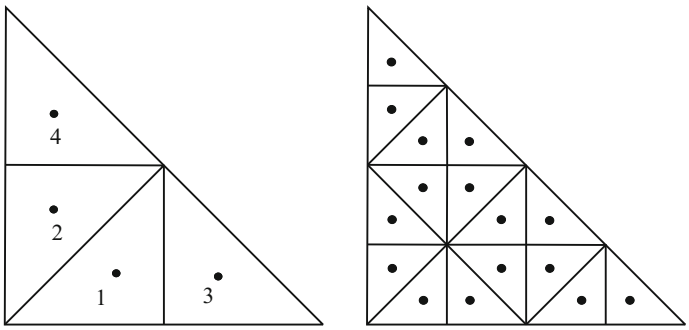


Fig. 5 Warning points

The presented procedure is simple and easy but, when the analyzed domain is big, the computational cost becomes unacceptable.

2.2.1 Triangles Filtering Process

Because of the abovementioned high computational cost, the need of a filtering process arises in order to choose the triangles which are likely to be shadowed and avoid the analysis on triangles that are impossible to be shaded. A triangle may be under shadows due to its own or self-shadows, or due to the shadows cast by another one. When a surface is back to the Sun, it will be under its own or self-shadows as can be observed in Fig. 6. A good way to compute this possibility is making use of the incidence angle[1] (δ_{exp}) as its value should be greater than $\pi/2$.

This way, a *Self-Shadows Light Factor* (L_{fss}) can be defined:

$$0 \quad \text{IF} \quad \left|\delta_{exp}\right| > (\pi/2) \tag{2}$$

$$1 \quad \text{IF} \quad (\pi/2) \geq \left|\delta_{exp}\right| \geq 0 \tag{3}$$

Now we are ready for the **first filtering** process, which means that the analysis on all *back to Sun* triangles can be avoided. A first conclusion can be done: only self-shadowed triangles are able to cast shadows (see Fig. 7), so they will be called *potential 1* triangles. Now the **second filtering** process begins projecting the mesh, referenced on the new plane $x' - y'$, towards z', forming parallelepipeds that we have called *cubes* (see Fig. 8). Cubes depend on the studied time step, this is, the lower the Sun is, the more triangles will be included in one cube. At this stage, we need

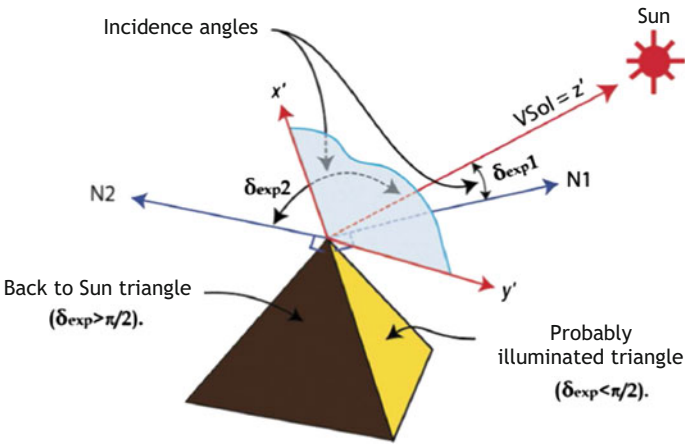


Fig. 6 Incidence angle and self-shadows

[1] Angle between the solar vector and the normal to a triangle surface.

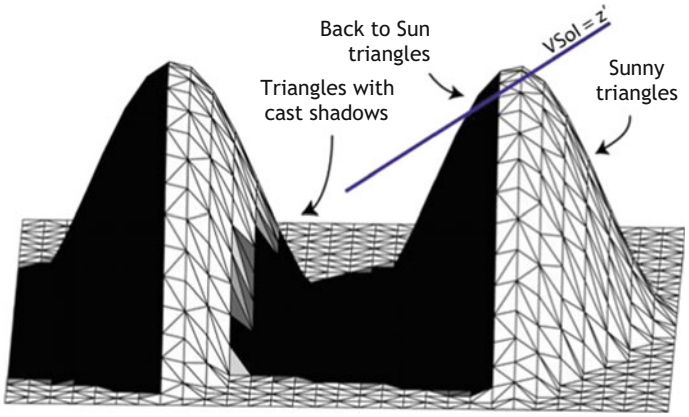


Fig. 7 Cast and self-shadows

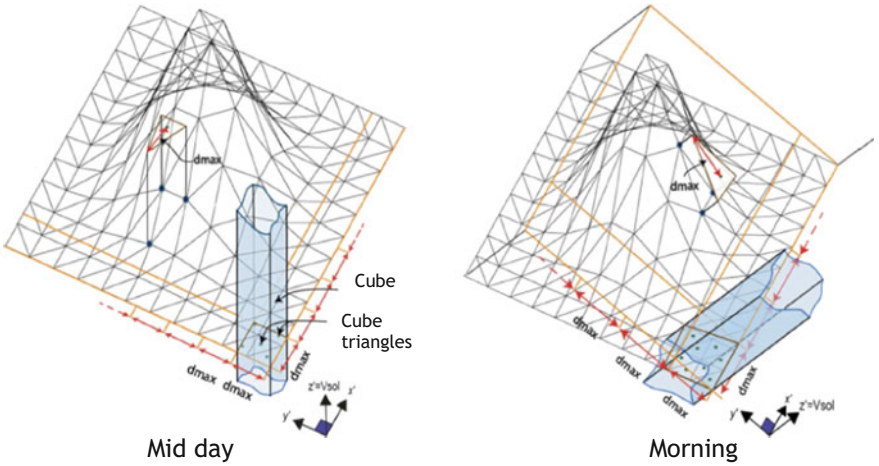


Fig. 8 2nd filtering: cubes

to remember that each triangle Δ needs to be analyzed to find other triangle(s) Δ' that intersects Δ and is before it respect to the Sun (coordinate z' of the intersection point is bigger for Δ' than for Δ).

Cubes are built in the plane $x' - y'$ in a way that the sides of their bases have to be equal or greater than the maximum value of the distance from the center to any of its nodes for all the grid triangles. So, the nodes of the mesh largest triangle do not extend beyond the eight parallelepipeds surrounding the cube which contains its center. Figure 9 shows the analysis of a time instant. Every *potential 1* triangle is associated to a cube containing its gravity center. When studying any of the illuminated triangles, the analysis is now limited to those *back to Sun* triangles placed into those cubes

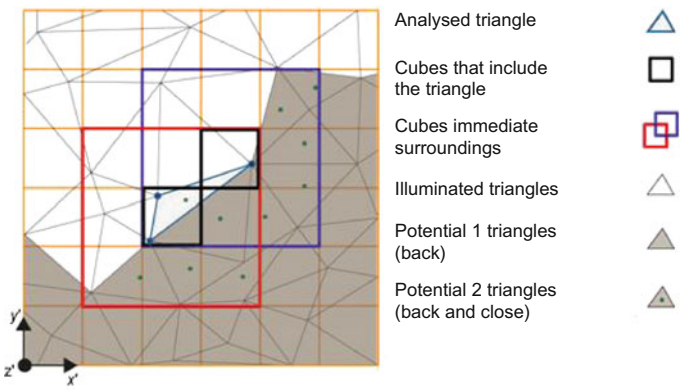


Fig. 9 Envelope and *potential 2* triangles

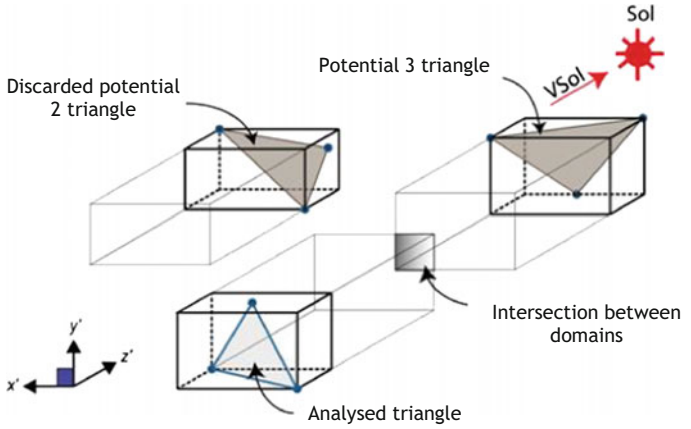


Fig. 10 Cubic domains

which are around[2] them. At this stage, the *potential 2* triangles are defined as in Fig. 9. So, now and for an illuminated triangle Δ, *potential 1* triangles far from our Δ, are discarded. What happens to *potential 2* triangles?. These ones and the analyzed illuminated triangle are embedded into cubic domains as shown in Fig. 10. Now, we can find *potential 2* triangles with cubic domains intersected with the domain of Δ, what will be called *potential 3* triangles (back to Sun, in close cubes, and with intersected domains), and other *potential 2* triangles, not intersected. These last ones are discarded too, making the **third filtering** process.

[2]The eight nearer cubes.

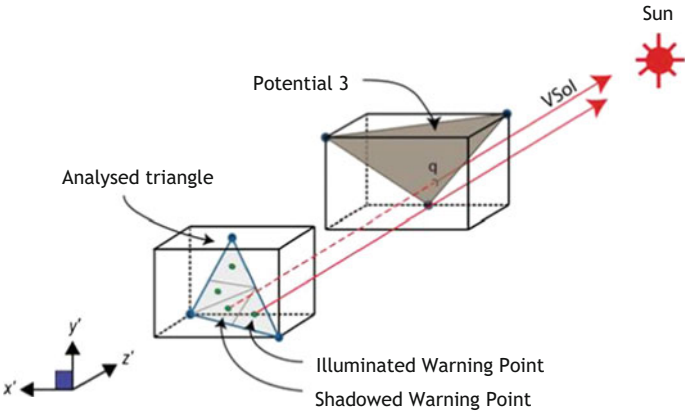


Fig. 11 Warning points analysis

Now it is time to face the final decision on the quantity of shadowing that our illuminated triangle, Δ, has. As said above, the process is finished checking the projection of shadows on a small set of points, *warning points (wp)*. A *warning point* is shaded when (Fig. 11):

1. The *warning point*, projected on the plane $x' - y'$, is into the projection of any *potential 3* triangle.
2. The *warning point* is farther from the Sun than point q in the potential triangle.

Cast shadows *light factor*, L_{fcs} will be computed as above:

$$L_{fcs} = \frac{n_{wp} - i}{n_{wp}} \tag{4}$$

where $i = 0, 1, \ldots, n_{wp}$ is the number of *warning points* which are shaded, and n_{wp} the total number of *warning points*. Including the possibility of considering self-shadows, the final *light factor*, L_f, is

$$L_f = L_{fcs} \cdot L_{fss} \tag{5}$$

As examples, simulations on orographic shadows have been done on Gran Canaria and Tenerife islands (Canary Islands–Spain). Vectors (red or blue) in the Figs. 12, 13 and 14 represent the solar vector direction for time and date of the simulation. As said above, the quantity of shadow is represented in a gray scale, from 0 to 1, according to the number of shaded *warning points* (L_f). See Figs. 12, 13, and 14.

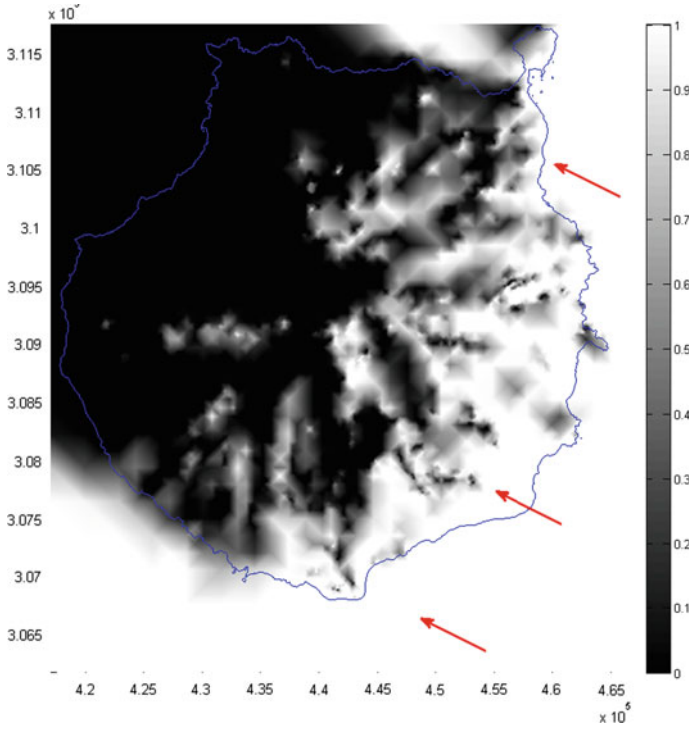


Fig. 12 Orographic shadows on Gran Canaria Island. January, 1st, 8:00 UTC

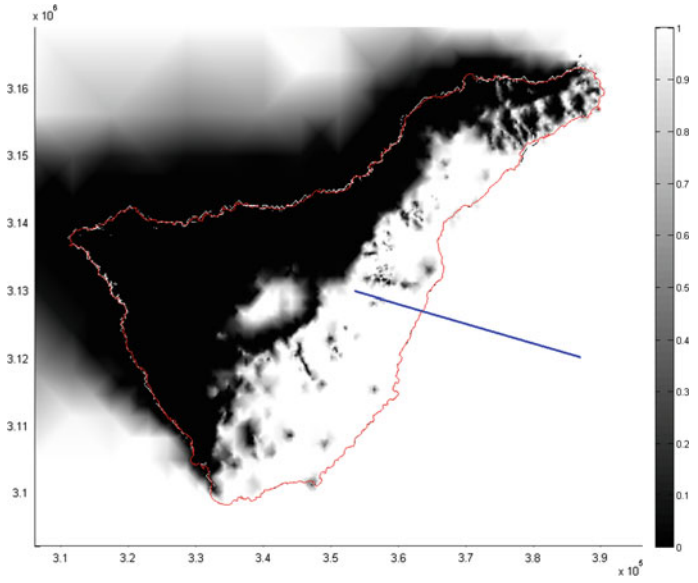


Fig. 13 Orographic shadows on Tenerife. February, 15th, 8:00 UTC

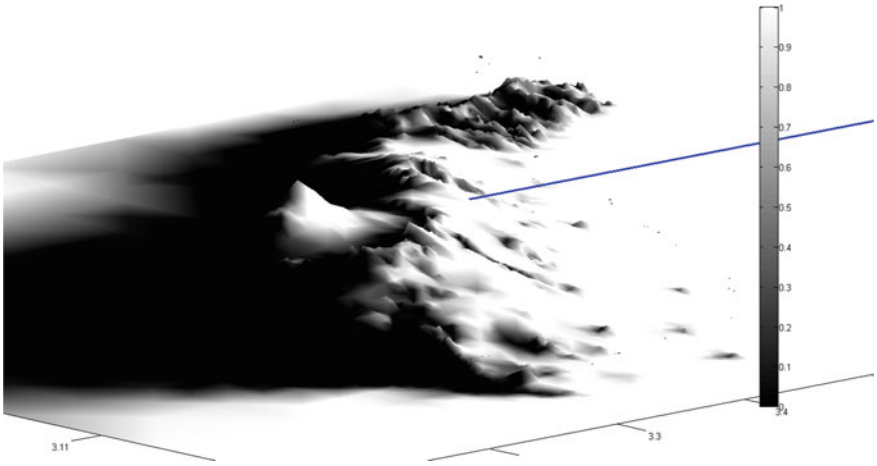


Fig. 14 3D orographic shadows on Tenerife. February, 15th, 8:00 UTC

2.3 Shadowing by Clouds

As we all know, clouds can cast shadows over the terrain and, considering their influence on the terrain is not an easy problem to solve because we do not know where and when clouds will be at any moment. Because of this, pictures of the actual clouds are interesting, what means that satellite images are a powerful tool for it. Statella and da Silva [16] or Fisher [17] are authors who have treated this issue. The Landsat dataset [18] is used in this chapter. Images come from sensors OLI and TIRS:

- Operational Land Imager (OLI)
- Thermal Infrared Sensor (TIRS) (bands 10 and 11)

OLI and TIRS sensors of the LANDSAT-8 images include 30 meters resolution nine spectral bands for bands 1–7 and 9. The panchromatic band (8) has a resolution of 15 meters. On the other hand, both thermal bands (10 and 11) are very interesting for surface temperatures, and have a resolution of 100 m. The Quality Control Band offers additional data, useful for determining the presence of snow, water, or ice. Cirrus and clouds can be determined too with different levels of confidence. Starting from an image like the one in Fig. 15, we will develop a cloud detection strategy based on the use of the satellite quality bands [19]. Assessing the shadows that come from clouds needs the estimate of radiance values from the image and the conversion of these ones into brightness temperature for, finally, computing the actual temperature that is used to assess clouds altitude and the shadow length (see Fig. 16).

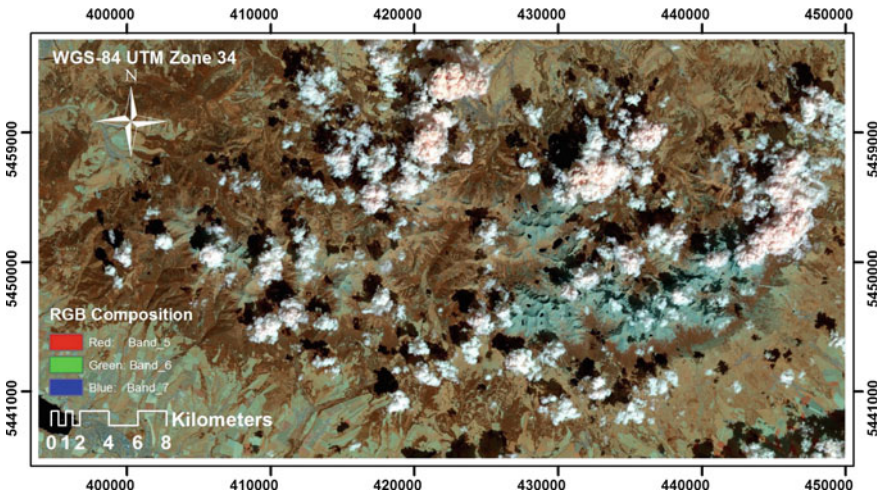


Fig. 15 RGB composition on Tatra Mountains image. Bands 5, 6, 7

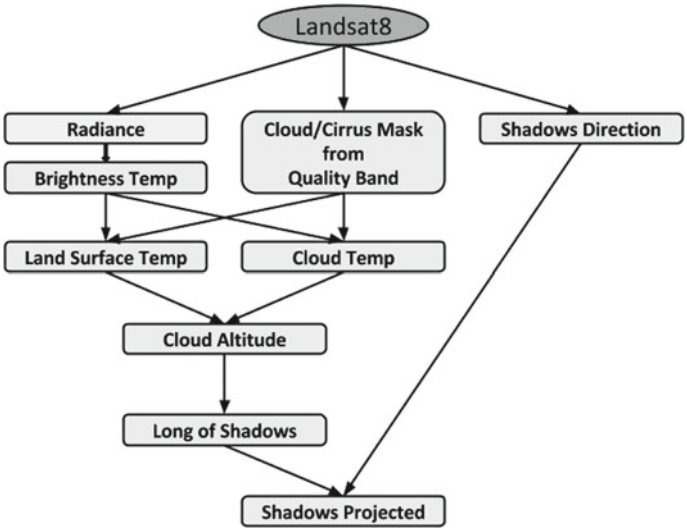


Fig. 16 Shadows from clouds flowchart

According to [19], data from OLI and TIRS can be rescaled to the top of the atmosphere radiance values, using Eq. 6.

$$L_\lambda = M_L \cdot Q_{cal} + A_L \tag{6}$$

where L_λ is the top of the atmosphere (TOA) spectral radiance ($W/(m^2 \cdot srad \cdot um)$), M_L is the multiplicative scaling factor, A_L is the additive scaling factor, and Q_{cal} is the standard product quantified and calibrated pixel values.

Brightness temperature (K) [19] is calculated as

$$T_b = \frac{K_2}{\ln\left(\frac{K_1}{L_\lambda} + 1\right)} \tag{7}$$

T_b is the apparent brightness temperature (K), and K_1 and K_2 are specific conversion constants for bands 10 and 11. Real Land Surface Temperature is assessed through Eq. 8:

$$T_s = c_0 + T_i + c_1 \cdot (T_i - T_j) + 2c_2 \cdot (T_i - T_j) + (c_3 + wc_4) \cdot (1 - \varepsilon) + (c_5 + wc_6) \cdot \Delta\varepsilon \tag{8}$$

with the coefficients computed according to [20, 21],

As indicated in [22, 23], Brightness Cloud Temperature obtained from thermal bands is assumed to be the Real Cloud Temperature. This way, the pixels group that makes a cloud, is assigned an average temperature to estimate its altitude, computed through Henning's formula (Eq. 9) [24].

$$H(m) = 125 \cdot (T - T_d) \tag{9}$$

where T is surface temperature and T_d is cloud temperature. Computing the shadows length (L) is easy [25, 26]:

$$L = \frac{H}{\tan h_0} \tag{10}$$

where h_0 is the elevation angle (see Fig. 3).

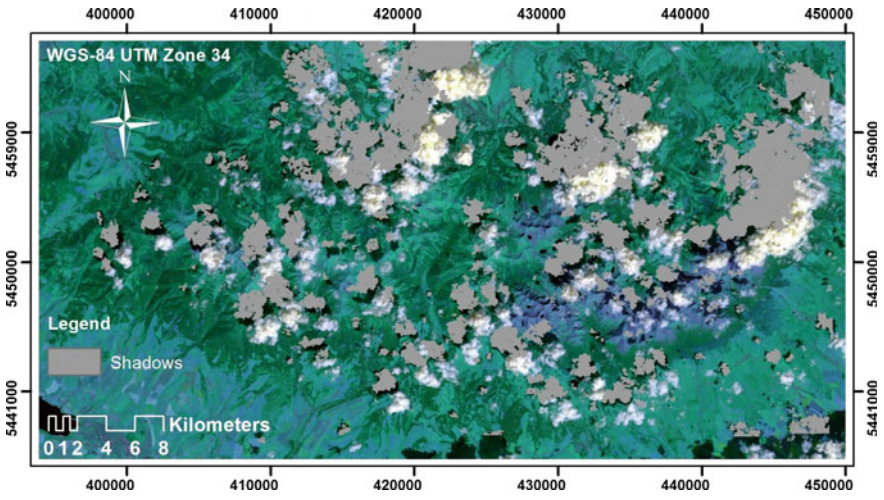


Fig. 17 Shadows mask. Band 10

The case of the Tatra Mountains (see Fig. 15) is presented as an example of finding shadows on the terrain caused by clouds. We have used two shadow masks bands: number 10 and number 11, and finally chose number 10 because of its better behavior (see Fig. 17).

Now it is time to move this mask into the mesh, and find what triangles are affected by shadows from clouds. Figure 18 shows our mesh with the new shadows. Light factor for each triangle is computed using the above mentioned strategy of the *warning points*.

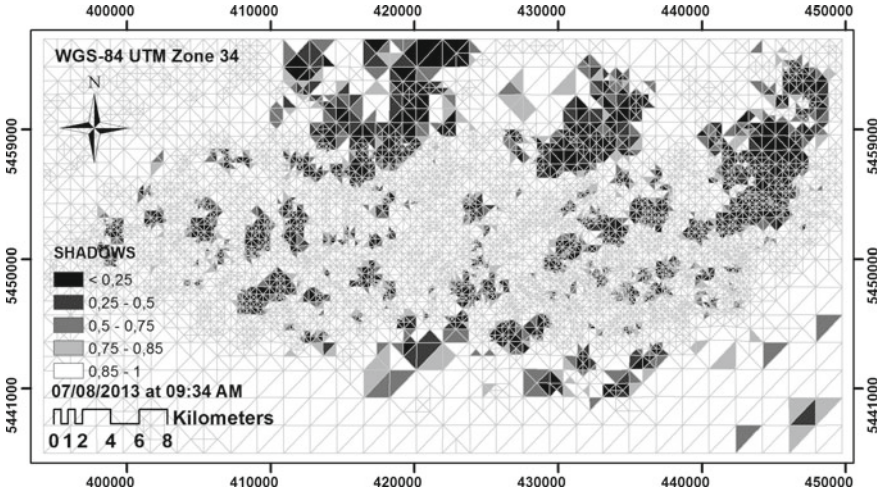


Fig. 18 Shadows estimation onto mesh

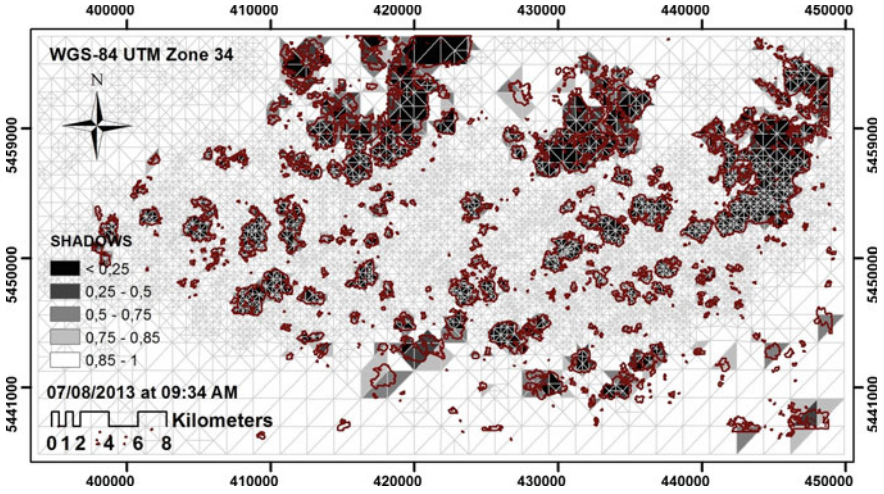


Fig. 19 Shadows borderline

Drawing the actual shadows borderline with the estimation on the mesh (Fig. 19) serves as a proof of how good is this approach that gives a good estimation for each time step with an affordable computational cost.

3 Solar Radiation Model

In this approach, as we have the domain discretized through a triangles mesh, the solar radiation model constructed on the terrain grid needs to compute the solar radiation values on each triangle, for every time step. In short, the model explained here will be built on a clear sky radiation[3] model whose results will be corrected using a *clear sky index* (K_t^*) (Eq. 11) which takes into account all the measurements available for the analyzed domain. This way real sky radiation values can be obtained.

$$G = K_t^* \cdot G_c \tag{11}$$

where G is real sky radiation and G_c is clear sky radiation.

As we have some mathematical clear sky radiation models available (see Chap. 8), the actual problem to be solved is the computation of the conversion factor, this is, K_t^*. To get the radiation map of the studied domain, we need to do the following:

1. TMY calculation in the available measurement stations
2. Solar radiation calculation for all the mesh, assuming clear sky conditions
3. Calculation of K_t^* for all the mesh
4. Correction of clear sky solar radiation values using K_t^* to obtain real sky radiation

Of course, all steps are computed for each time step except step one which is done just once. Computation of K_t^* can be done for each time step, or for a considered time interval, let us say, a day, a month, etc.

In step one, a Typical Meteorological Year (TMY) is computed for all the available measurement stations in the domain. There are different approaches for determining this TMY, and the chosen one appears in [7, 27].

3.1 Clear Sky Solar Radiation Model

In this chapter, the ESRA[4] clear sky model is used [5, 6]. This model gives solar irradiance values in terms of their three components: *beam*, *diffuse*, and *reflected* irradiances, being the global irradiance the addition of the three components. The beam irradiance is not scattered and reaches any surface directly with an associated

[3] Omitting cloud attenuation, clear sky radiation values are obtained.

[4] European Solar Radiation Atlas.

direction. Diffuse irradiance is the scattered one. The reflected irradiance is the one reflected from one surface onto another. It depends on the ground albedo.

3.1.1 Beam Radiation

The first step is computing the extraterrestrial irradiance G_0:

$$G_0 = I_0 \epsilon \tag{12}$$

where ϵ depends on the day angle, and the solar constant I_0 is 1367 (W/m^2) [28] The direct normal irradiance on earth surface in clear sky conditions, DNI_c (W/m^2), is attenuated by the cloudless atmosphere, and computed as

$$DNI_c = G_0 \, exp\{-0.8662 T_{LK} m \delta_R(m)\} \tag{13}$$

The term $0.8662 T_{LK}$ is the corrected [29] dimensionless Linke atmospheric turbidity factor. The parameter m is the relative optical air mass [30], and $\delta_R(m)$ is the Rayleigh optical thickness at air mass m.

On a horizontal surface, the beam irradiance, of course for clear sky conditions $G_{bc}(0)$, is

$$G_{bc}(0) = (DNI_c) \cdot L_f \cdot \sin h_0 \tag{14}$$

where L_f is the light factor that takes into account the shadows (Eq. 5), and h_0 is the solar altitude angle. On an inclined surface, $G_{bc}(\beta)$ is

$$G_{bc}(\beta) = (DNI_c) \cdot L_f \cdot \cos \delta_{exp} \tag{15}$$

where δ_{exp} is the *solar incidence angle* measured between the characteristic vector of an inclined surface and the sunbeam direction, and β is the angle between the inclined surface and the horizontal.

3.1.2 Diffuse Radiation

On horizontal surfaces, this component is estimated with

$$DHI_c = G_{dc}(0) = G_0 T_n(T_{LK}) F_d(h_0) \tag{16}$$

It depends on the diffuse transmission T_n which depends on the Linke turbidity factor T_{LK}. Function F_d, in turn, depends on the solar altitude h_0 [31]. On a inclined surface, $G_{dc}(\beta)$, both, sunlit and shadowed surfaces have to be considered as in [32].

3.1.3 Reflected Radiation

According to [33], the ground reflected irradiance $G_r(\beta)$ is computed as

$$G_r(\beta) = \rho_g GHI_c r_g(\beta) \tag{17}$$

where

$$r_g(\beta) = \frac{(1 - \cos\beta)}{2} \tag{18}$$

$$GHIc = G_{bc}(0) + DHIc \tag{19}$$

being ρ_g the mean ground albedo and $r_g(\beta)$ the fraction of ground viewed by an inclined surface.

3.2 *Computing the Clear Sky Index. Real Sky Solar Radiation Model*

At this stage, irradiance values in clear sky conditions have been obtained. However, it is obvious that real sky conditions, considering the clouds, are the actual goal. To reach these real sky radiation values, the clear sky index (see Eq. 11), is very useful. Now it depends upon the available measurement data to compute the K_t^*. If we have instantaneous irradiance measurements in any station, we will have instantaneous clear sky indexes for that station (s) at any time step (i) as seen in Eq. 20.

$$K^*_{t_{s,i}} = \frac{G_s}{G_{c,s}} \tag{20}$$

But sometimes the available data in a measurement station is irradiation (H_s) along a time period. In this case, clear sky index for a station (s) and for that time period is obtained this way:

$$K^*_{t_s} = \frac{H_s}{H_{c,s}} = \frac{H_s(0)}{H_{c,s}(0)} \tag{21}$$

where

$$H_{c,s}(0) = \int_0^T G_{c,s}(0)\, dt \tag{22}$$

As measurement equipments usually work on a horizontal plane ($\beta = 0$), computing $K^*_{t_s}$ will be as in Eq. 21. All these clear indexes are those computed for any of the available measurement stations in the domain. Of course, any computed $K^*_{t_s}$ is for the triangle where the measures were taken. But in the domain we will have

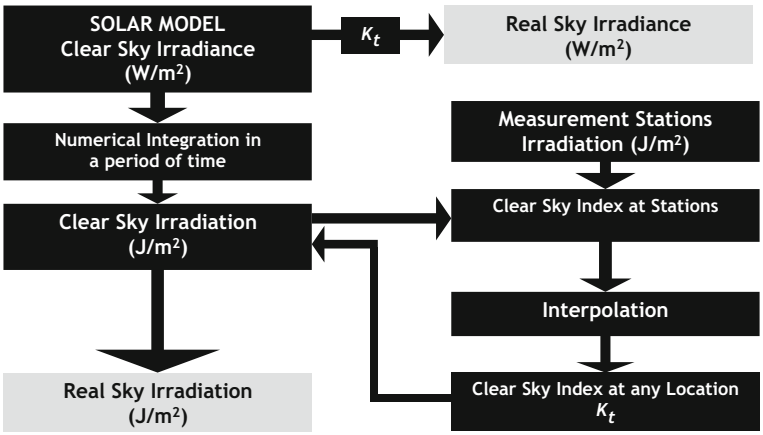


Fig. 20 Solar radiation model

hundreds of triangles with no measures. In these ones, an interpolation procedure is implemented to estimate the value ok K_t^*. A formula for complex orography [34] is applied,

$$K_t^* = \varepsilon \frac{\sum_{n=1}^{N} \frac{K_{tn}^*}{d_n^2}}{\sum_{n=1}^{N} \frac{1}{d_n^2}} + (1-\varepsilon) \frac{\sum_{n=1}^{N} \frac{K_{tn}^*}{|\Delta h_n|}}{\sum_{n=1}^{N} \frac{1}{|\Delta h_n|}} \tag{23}$$

where K_t^* is the clear sky index at each mesh triangle, K_{tn}^* is the clear sky index at measurement stations, d_n is the horizontal distance and $|\Delta h_n|$ is the altitude difference between station n and the studied point, N is the number of stations in the interpolation, and ε is a fitting parameter between 0 and 1. Of course, Eq. (23) continuity is guaranteed assuming the measured values at the stations.

Now, for any triangle of the mesh, irradiance in real sky conditions is easily computed as follows:

$$GHI = G(0) = K_t^* \cdot GHI_c \tag{24}$$

As a summary, an outline of the process is presented in Fig. 20

4 Numerical Experiments. Solar Maps

Talking about solar radiation maps is usually talking about irradiation distribution on a terrain. As an example, several numerical experiments in some of the Canary Islands are presented. The Canary Islands is an archipelago of seven islands placed in the northwestern part of Africa, in a latitude around 28, with a subtropical climate. In fact, we have already seen some shadowing examples for Tenerife and Gran

Canaria Island. Figure 21 shows the Gran Canaria Island elevation map, with all the measurements stations available (see characteristics in Table 1).

First, the correct determination of the ε parameter in Eq. 23 is sought. For that, it is necessary a good knowledge about the climate conditions of the domain, this is, of the zone under study. In this case, the Canary Islands are under the influence of the Trade Winds which come from the Northeast full of humidity, and that get stopped by the mountains of the islands producing a different behavior between the northeast and the southwest of the isles. Varying ε, we obtain a K_t^* distribution map for January and July (see Figs. 22 and 23).

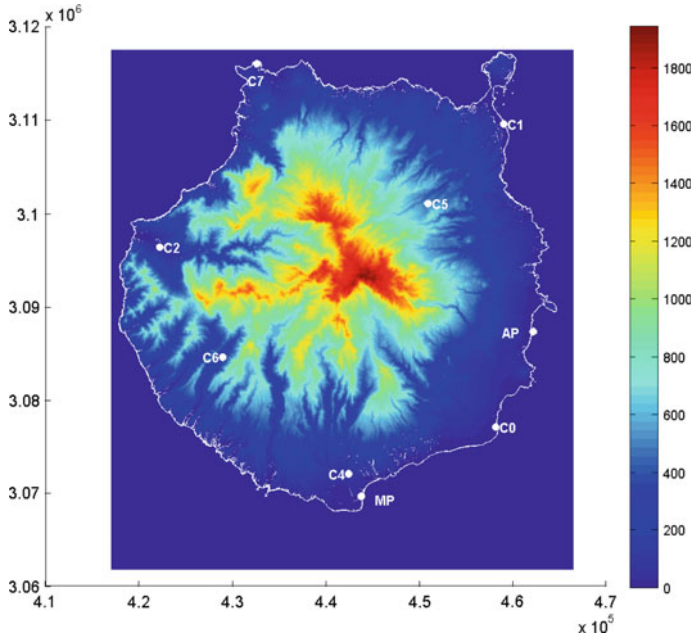


Fig. 21 Gran Canaria Island elevation map

Table 1 Available measurement stations in Gran Canaria Island

Station	Label	Latitude (N)	Longitude (W)	Altitude
Pozo Izquierdo	C0	27.8175	15.4244	47
Las Palmas de G. C.	C1	28.1108	15.4169	17
La Aldea de San Nicolás	C2	27.9901	15.7907	197
San Fernando de M.	C4	27.7716	15.5841	265
Santa Brígida	C5	28.0337	15.4991	525
Mogán (village)	C6	27.8839	15.7216	300
Sardina de Gáldar	C7	28.1681	15.6865	40

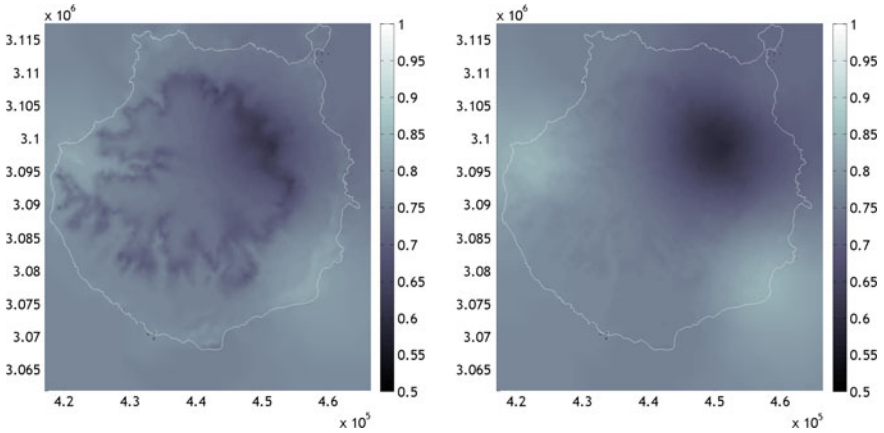


Fig. 22 K_t^* distribution for January. $\varepsilon = 0.5$ and $\varepsilon = 0.9$

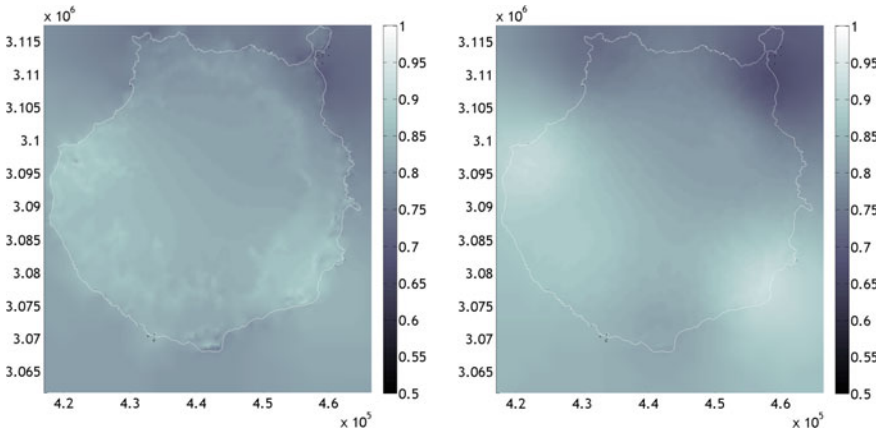


Fig. 23 K_t^* distribution for July. $\varepsilon = 0.5$ and $\varepsilon = 0.9$

Both figures show a false behavior of K_t^* with $\varepsilon = 0.5$, while with that parameter equal to 0.9, the behavior of the clouds fits what is known about the Canary Islands climate. This means that clouds depend on the horizontal displacements on the domain, more than on the vertical ones. With this ε, we compute K_t^* in all the triangles, and obtain the real sky irradiance as a correction of the clear sky one. To complete the maps, it is only needed to integrate the irradiance in all the domain, along a time period, and represent it. As an example, Figs. 24 and 25 show a monthly average irradiation map for all the clear sky components in a TMY January. It is really interesting to watch the influence of K_t^* comparing pictures in Fig. 26. The first one presents the clear sky irradiation map for TMY July, and the second one is the real sky irradiation map for the same month. This is a month with an important

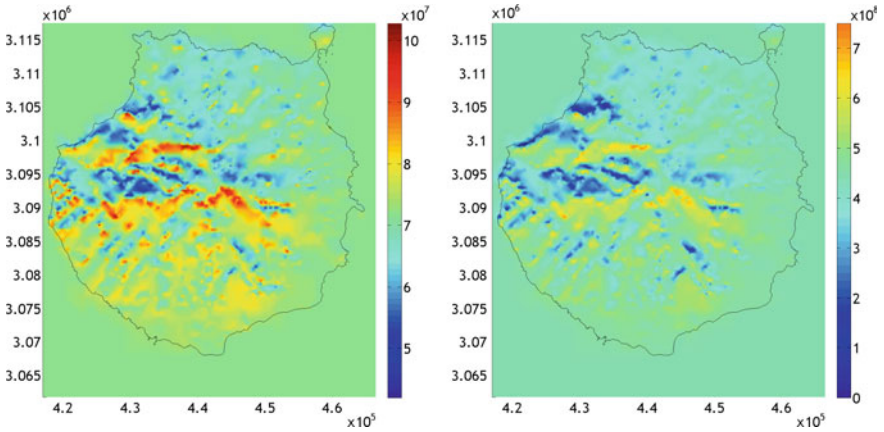


Fig. 24 Clear sky irradiation (J/m^2) maps for Gran Canaria in January. Diffuse and Beam

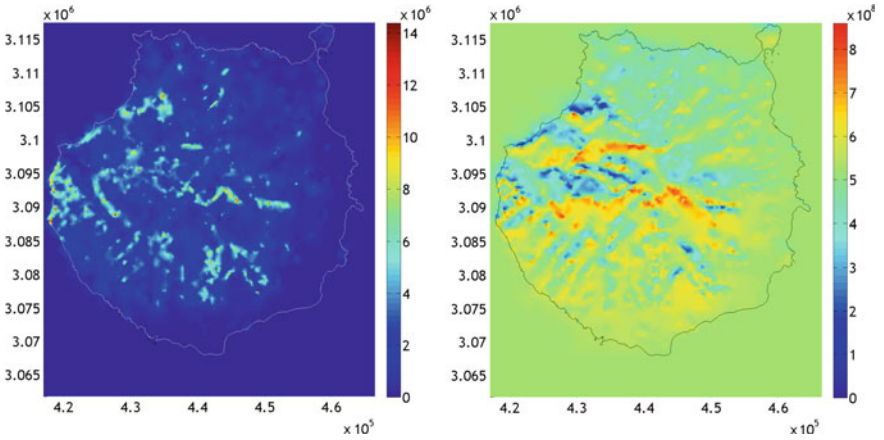


Fig. 25 Clear sky irradiation (J/m^2) maps for Gran Canaria in January. Reflected and Global

influence of the Trade Winds, that can be observed in the north and northeast side of the island, with lower irradiation values than on the south face.

It can be clearly observed that cloudiness and orographic shadows have been taken into account. To observe better this fact, Fig. 27 is presented, with a three-dimensional view of the island irradiation map. Please note that relation between vertical and horizontal scales has been exaggerated.

The morphology of this island (Gran Canaria) is very radial so, this is an advantage to apply Eq. 23. Moreover, it works perfectly when applying to other islands with more complex orography and with a morphology quite different from the radial one. As an example, numerical experiments in Tenerife Island are presented in Fig. 28.

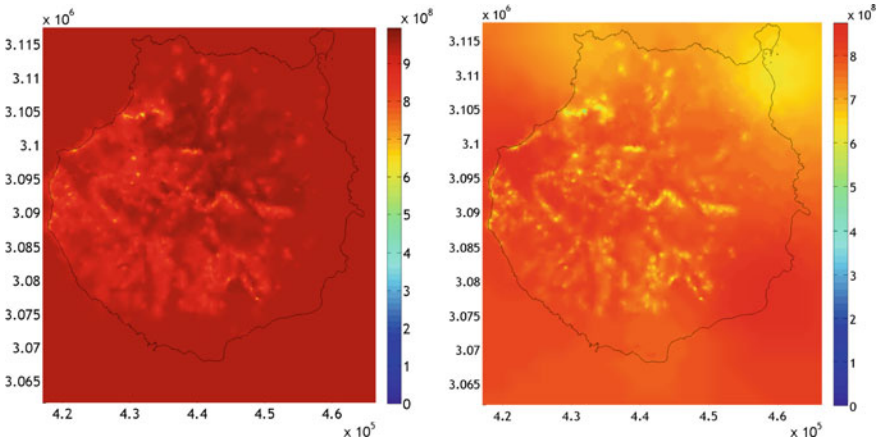


Fig. 26 Clear and real sky irradiation (J/m^2) maps for Gran Canaria in July

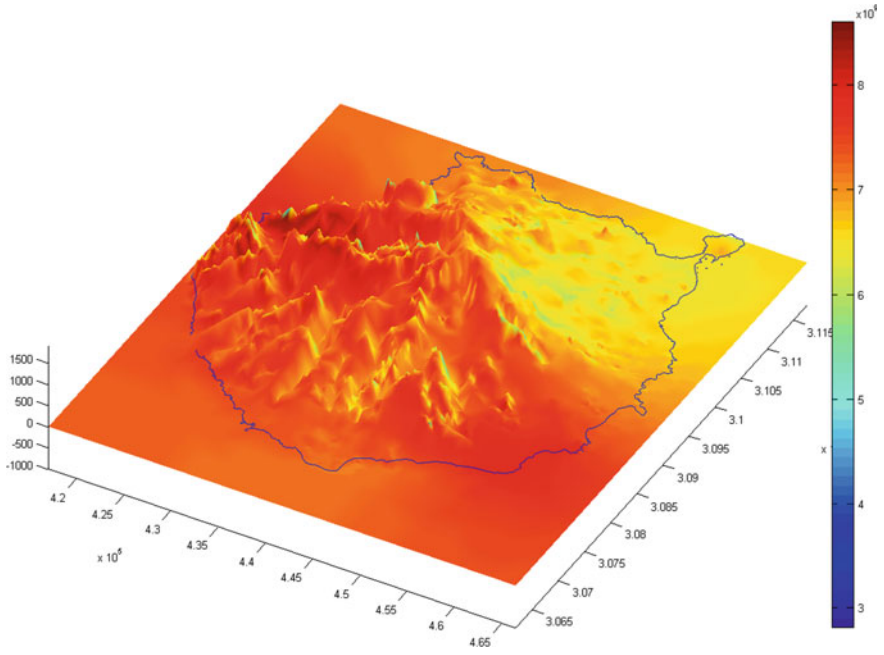


Fig. 27 3D real sky annual irradiation map for Gran Canaria Island

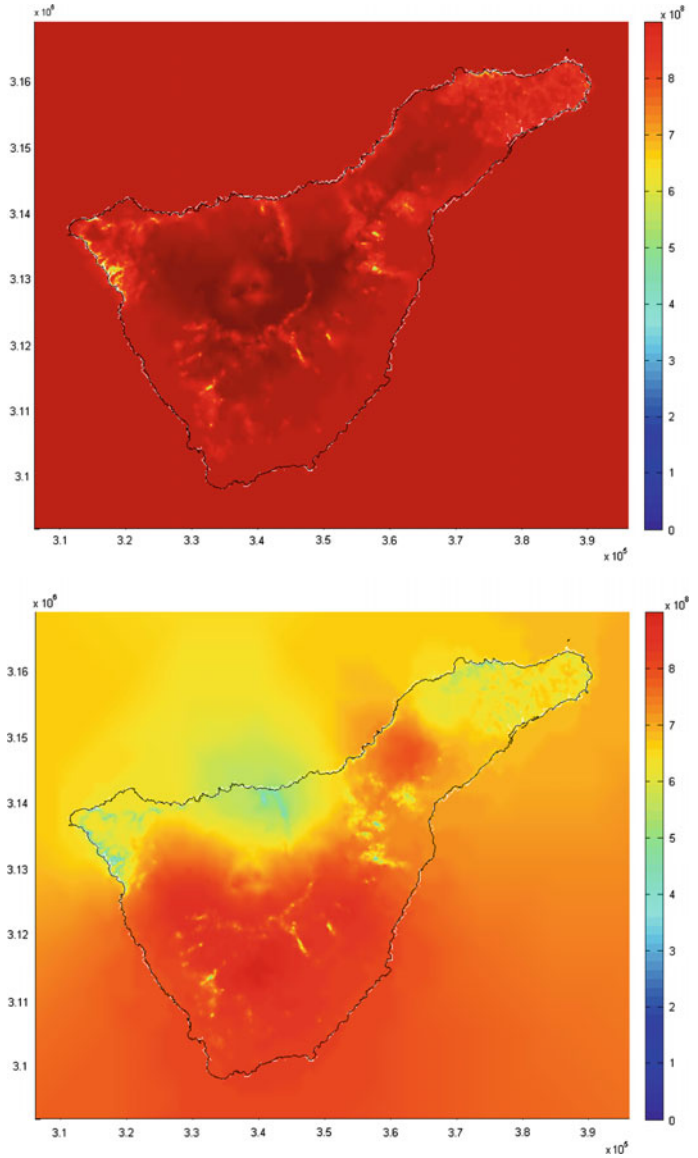


Fig. 28 Clear and real sky irradiation (J/m^2) maps for Tenerife Island in July

5 Constructing Predictive Solar Maps

In previous sections, how to build a solar map using punctual ground solar radiation measurements has been explained. This procedure results in a characterization of the domain (i.e., an island), which is very helpful to understand the solar radiation behavior for agronomic, biological, or engineering purposes. In this last aspect, electrical energy production highlights. Regarding to this, power systems stability and planning need a great knowledge about the future electrical production of the different generators involved in the grid. So, if there are big solar power facilities injecting energy, an efficient and accurate prediction tool for solar radiation is needed. Because of this, we have presented some solar radiation prediction tools in this book (see Chaps. 8 and 9).

In this section, we will convert our solar maps generation code in a prediction tool [35], making use of a Numerical Weather Prediction (NWP) model like Harmonie, ECMWF or other. In this example, model MM5[5] is used. The methodology applied is the one explained in this chapter but, using MM5 predicted data instead of the TMY measurement stations data. This NWP model assesses irradiances on a grid whose horizontal projection is not uniform. These data have to be transported to our adaptive mesh interpolating the results to the centers of the triangles.

We have to adapt our model according to the NWP used. In example, MM5 does not take into account the actual domain orography, nor the casting of shadows. On the other hand, as said above, determining the clear sky index for every node in the MM5 grid (transposed to the mesh) is needed. As the NWP grid changes their values depending on time, we will have to use a clear sky index for each node and time step ($K^*_{t_i}$).

$$K^*_{t_i} = \frac{G(0)}{G_{c_{ss}}(0)} \tag{25}$$

where $G_{c_{ss}}(0)$ is the horizontal clear sky irradiance, considering only self-shadows. Later, the final real sky predicted irradiance will be computed using this $K^*_{t_i}$ and $G_c(0)$ considering the self- and cast shadows. Computing the clear sky index includes these steps:

1. Make an MM5 grid triangulation where the nodes are those from the MM5 grid, and are assigned the predicted radiation for every time step

[5] National Center for Atmospheric Research (Pennsylvania State University -USA-) mesoscale fifth generation model.

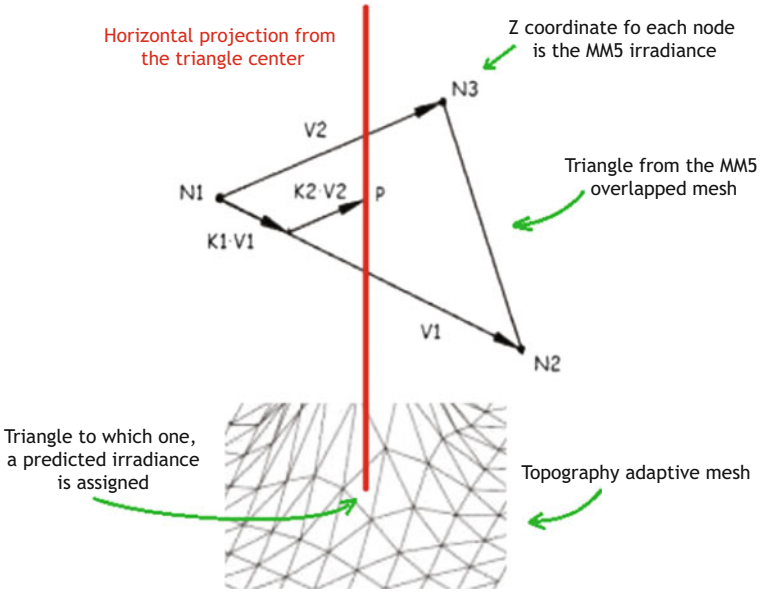


Fig. 29 Predicted irradiance transpose to mesh

2. The new MM5 triangulated grid overlaps the orography adaptive mesh. If we take a look at Fig. 29, we can see that every orography triangles are assigned the interpolated irradiance using Eq. (26)

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + K_1 \cdot \begin{bmatrix} V_{1_x} \\ V_{1_y} \\ V_{1_z} \end{bmatrix} + K_2 \cdot \begin{bmatrix} V_{2_x} \\ V_{2_y} \\ V_{2_z} \end{bmatrix} \tag{26}$$

   Three equations in three unknowns, K_1, K_2 and z, that is the predicted irradiance for the adaptive mesh triangle
3. Clear sky index ($K_{t_i}^*$) is computed for each triangle and time step, using self-shadows clear sky irradiance

Computing $K_{t_i}^*$ will be done with Eq. 23 for each and every triangle of the mesh, and for every time step. As a summary, Fig. 30 shows an outline of the procedure to follow, where measurements have been substituted by predictions in the determination of $K_{t_i}^*$.

A numerical simulation using this methodology is done in two locations of Gran Canaria: *Temisas*[6] and *Lomo Carbonero*.[7] As the clear sky index is computed for each time step, the daily cloudiness evolution can be addressed through this parameter

[6] 27.917°N 15.491°W.

[7] 28.033°N 15.533°W.

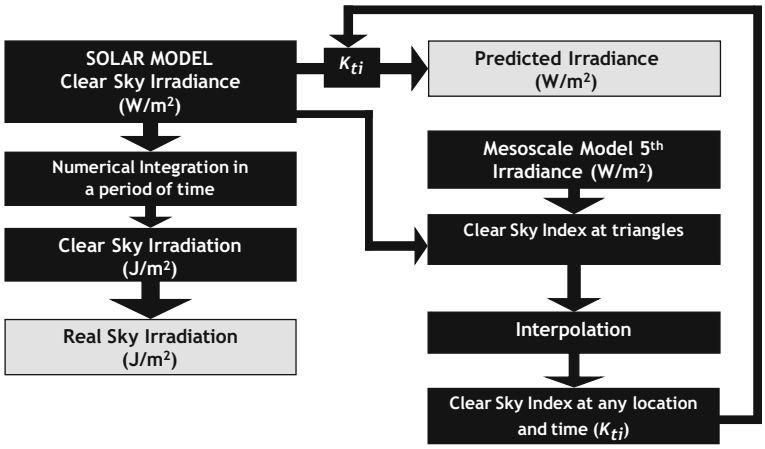


Fig. 30 Predictive solar radiation model

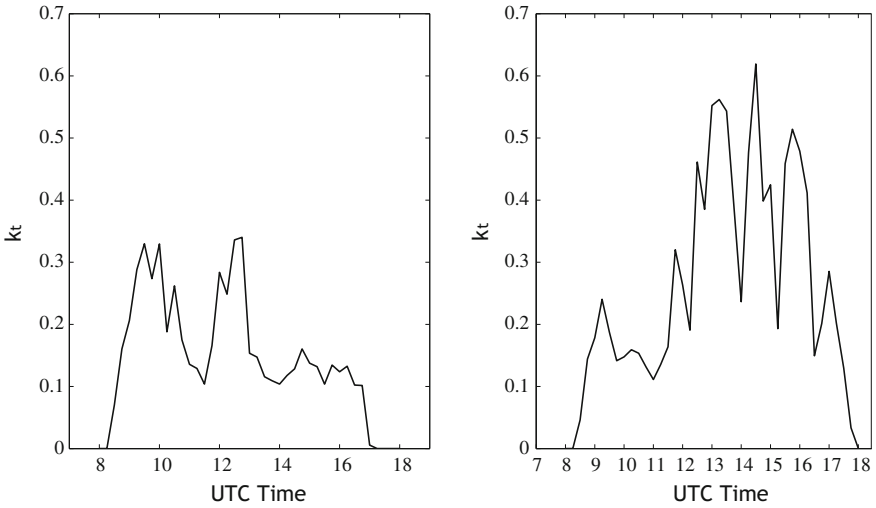


Fig. 31 Clear sky index prediction for *Temisas* and *Lomo Carbonero*

(see Fig. 31). Starting from this distribution, the predicted irradiance for every time step can be computed (see Fig. 32).

The example has been done for different locations in the domain. If we compute it for all the triangles and for one instant, we will obtain an irradiance predicted map and, for different time steps, we will get a family of maps with the predicted evolution of the irradiance. In case that daily (or monthly,...) irradiation values are needed, integration of the predicted irradiance will have to be previously performed.

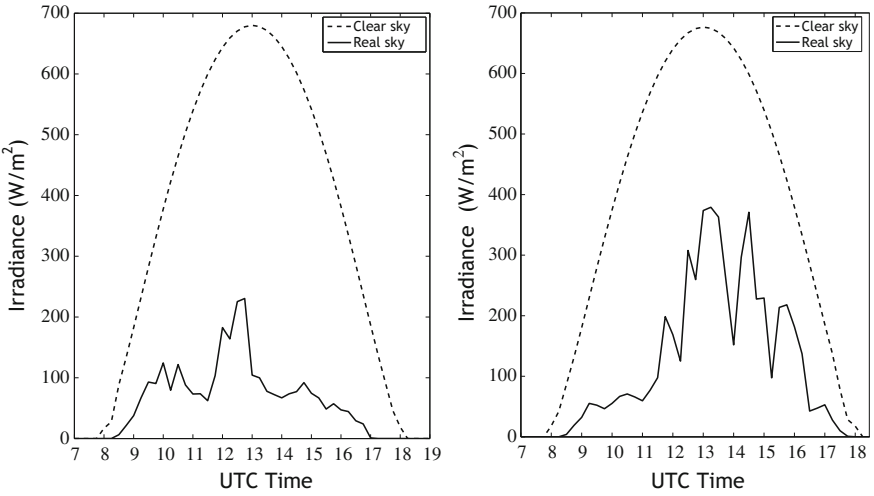


Fig. 32 *Temisas* and *Lomo Carbonero* horizontal predicted irradiance

6 Conclusions

This chapter is devoted to the explanation of the methodology used in making solar maps. These maps are usually done with monthly or annual irradiation values but sometimes it is interesting to have irradiance maps, especially when we are talking about solar radiation predictions. The discretization of the complex terrain orography has been done using adaptive triangular meshes which allow to build surfaces that cast shadows. Moreover, this method is more efficient from the computational point of view when compared to the use of digital elevation maps or uniform meshes.

Values for real sky conditions are obtained from those computed using a clear sky solar radiation model, affected by the clear sky index. This one is valid for a triangle and a time step. The key issue, therefore, is computing $K_{t_i}^*$. A TMY is very useful to get monthly or annual solar irradiation maps.

References

1. Cogliani E, Ricchiazzi P (2008) Generation of operational maps of global solar irradiation on horizontal plan and of direct normal irradiation from Meteosat imagery by using SOLARMET. Sol Energy 82:556–562
2. Ångstrom A (1924) Solar and terrestrial radiation. Mon Weather Rev 52:397
3. Liu BYH, Jordan RC (1963) The long-term average performance of flat-plate solar energy collectors. Sol Energy 7:53–74
4. Perez R, Seals R, Ineichen P, Stewart R, Menicucci D (1987) A new simplified version of the perez diffuse irradiance model for tilted surfaces. Sol Energy 39:221–231

5. Šúri M, Hofierka J (2004) A New GIS-based Solar Radiation Model and its application to photovoltaic assessments. Trans GIS 8:175–170
6. Šúri M, Hofierka J (2002) The solar radiation model for Open source GIS: implementation and applications. In: Benciolini B, Ciolli M, Zatelli P (eds) Proceedings of the Open source GIS-GRASS users conference, Trento, Italy, pp 1–19
7. Díaz F, Montero G, Escobar JM, Rodríguez E, Montenegro R (2012) An adaptive solar radiation numerical model. J Comput Appl Math 236:4611–4622
8. Montero G, Escobar JM, Rodríguez E, Montenegro R (2009) Solar radiation and shadow modelling with adaptive triangular meshes. Sol Energy 83:998–1012
9. Zakšek K, Podobnikar T, Oštir K (2005) Solar radiation modelling. Comput Geosci 31:233–170
10. Niewienda A, Heidt FD (1996) SOMBRERO: a PC-tool to calculate shadows on arbitrarily oriented sufaces. Sol Energy 58:253–263
11. Dozier J, Bruno J, Downey P (1981) A faster solution to the horizon problem. Comput Geosci 7:145–151
12. Díaz F, Montero H, Santana D, Montero G, Rodríguez E, Mazorra Aguiar L, Oliver A (2018) Improving shadows detection for solar radiation numerical models. Appl Math Comput 319:71–85
13. Blanco-Muriel M, Alarcón-Padilla DC, López-Moratalla T, Lara-Coira M (2001) Computing the solar vector. Sol Energy 70:431–441
14. Sproul AB (2007) Derivation of the solar geometric relationships using vector analysis. Renew Energy 32:1187–1205
15. Rivara MC (1987) A grid generator based on 4-triangles conforming: mesh-refinement algorithms. Int J Num Meth Engrg 24:1343–1354
16. Statella T, da Silva EA (2008) Shadows and clouds detection in high resolution images using mathematical morphology. In: Proceedings of Pecora 17 conference the future of land imaging...going operational, Denver, USA
17. Fisher A (2014) Cloud and cloud-shadow detection in SPOT5 HRG imagery with automated morphological feature extraction. Remote Sens 6:776–800
18. Zhu Z, Woodcock CE (2012) Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens Environ 118:83–94
19. Zanter K (ed) (2016) LANDSAT-8 (L8) data users handbook. Department of the Interior U.S. Geological Survey, Maryland-USA (2016)
20. Jiménez-Muñoz JC, Sobrino JA, Skokovic D, Mattat C, Cristóbal J (2014) Land surface temperature retrieval methods from Landsat-8 thermal infrared sensor data. IEEE Geosci Remote Sens Lett 11(10):1840–1843
21. Rajeshwari A, Mani ND (2014) Estimation of land surface temperature of Dindigul district using Landsat 8 data. Int J Res Eng Technol 3(5):122–126
22. Inoue T (1984) On the temperature and effective emissivity determination of semi-transparent cirrus clouds by Bi-Spectral measurements in the 10nm window region. J Meteorol Soc Jpn 63(1):88–99
23. Yost CR, Minnis P, Sun-Mack S, Chen Y, McGill M (2008) Use of active remote sensors to improve the accuracy of cloud top heights derived from thermal satellite observations. In: Procedings of the AMS symposium on recent developments in atmospheric applications of radar and lidar, New Orleans, USA (2008)
24. Chiyu T, Kon H, Magono C (1972) The cloud-base topography and formation condition of cumulus humilis clouds. J Fac Sci Hokkaido Univ Ser 7 Geophys **4**(1):43–57
25. Musial JP, Husler F, Sutterlin M, Neuhaus C, Wunderle S (2014) Probabilistic approach to cloud and snow detection on Advanced Very High Resolution Radiometer (AVHRR) imagery. Atmos Meas Tech 7:799–822
26. Simpson JJ, Stitt JR (1998) A procedure for the detection and removal of cloud shadow from AVHRR data over land. IEEE Trans Geosci Remote Sens 36(3):880–897
27. Mazorra L, Díaz F, Montero G, Montenegro R (2010) Typical meteorological year (TMY) evaluation for power generation in Gran Canaria Island, Spain. In: Proceedings of 25th European photovoltaic solar energy conference and exhibition, Valencia, Spain, pp 4726–4728

28. Page JK (ed) (1986) Prediction of solar radiation on inclined surfaces'. D. Reidel Publishing Co., Dordrecht
29. Kasten F (1996) The Linke turbidity factor based on improved values of the integral Rayleigh optical thickness. Sol Energy 56(3):239–244
30. Kasten F, Young AT (1989) Revised optical air mass tables and approximation formula. Appl Optics 28:4735–4738
31. Scharmer K, Greif J (2000) The European solar radiation atlas. Database and exploitation software, vol. 2. Les Presses de l'Ecole des Mines, Paris
32. Muneer T (1990) Solar radiation model for Europe. Build Serv Eng Res Technol 11:153–163
33. Muneer T (1997) Solar radiation and daylight models for energy efficient design of buildings. Architectural Press, Oxford
34. Montero G, Montenegro R, Escobar JM (1998) A 3-D diagnostic model for wind field adjustment. J Wind Engrg Ind Aer 74–76:249–7261
35. Díaz F, Montero G, Escobar JM, Rodríguez E, Montenegro R (2015) A new predictive solar radiation numerical model. Appl Math Comput 267:596–603