



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
ENERGÍA – FUENTES RENOVABLES

ANÁLISIS DE MODELOS GLOBALES Y MESOESCALA EN LA EVALUACIÓN
DEL RECURSO EÓLICO

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN INGENIERÍA

PRESENTA:
JOSÉ GUSTAVO HERNÁNDEZ YEPES

TUTOR PRINCIPAL
OSVALDO RODRÍGUEZ HERNÁNDEZ
INSTITUTO DE ENERGÍAS RENOVABLES

COTUTOR
ADOLFO VICENTE MAGALDI HERMOSILLO
CENTRO DE CIENCIAS DE LA ATMÓSFERA

TEMIXCO, MORELOS, ENERO 2020



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

JURADO ASIGNADO:

Presidente: DR. EDUARDO RAMOS MORA
Secretario: DR. ADOLFO VICENTE MAGALDI HERMOSILLO
1 er. Vocal: DR. OSVALDO RODRÍGUEZ HERNÁNDEZ
2 do. Vocal: DR. OSCAR MARTÍNEZ ALVARADO
3 er. Vocal: DR. DANIEL ROBERT DREW

Instituto de Energías Renovables, Temixco, Morelos.

TUTOR DE TESIS:

DR. OSVALDO RODRÍGUEZ HERNÁNDEZ

FIRMA

(Segunda hoja)

Preface

Este trabajo representa la culminación de una meta personal, la cual gracias al esfuerzo y dedicación de muchas personas he podido llevar a cabo. Sin el apoyo de quienes me han ofrecido su mano generosa, jamás hubiera podido lograrlo.

Doy gracias a Dios por darme una familia que ha estado para mí en todo momento. A mis padres y hermanos los llevo conmigo a donde quiera que voy.

Me siento agradecido por los excelentes profesores que he encontrado a lo largo de mis estudios. Agradezco al Dr. Osvaldo Rodríguez, quien me introdujo al mundo de la energía eólica con gran pasión y sin dudarle siempre me ha apoyado en todos sentidos. Al Dr. Adolfo Magaldi por su amable orientación en el ámbito de ciencias de la atmósfera y por su constante respaldo. Al Dr. Oscar Martínez por abrirme las puertas para una estancia de investigación y trabajar de la mano con el parte de esta tesis, gracias por esa gran experiencia. Al Dr. Daniel Drew por su colaboración para enriquecer el presente trabajo, appreciate it! Al Dr. Eduardo Ramos por aportar su esfuerzo y experiencia para que este proyecto fuera posible. A todos ellos gracias por compartir generosamente sus conocimientos y ayudarme a forjar el futuro.

Agradezco la oportunidad de formar parte del IER, ya que tenido una preparación de excelencia y he conocido a magníficas personas y amigos. Sería imposible mencionar a todos ellos, por lo que bien dicen que una sonrisa vale más que mil palabras.

Agradezco al proyecto 272063 "Fortalecimiento del Campo de Energía Eólica en el Programa de Doctorado en Ingeniería Campo de Conocimiento en Energía con sede en el Instituto de Energías Renovables de la Universidad Autónoma de México" Fondo de Sustentabilidad Energética SENER-CONACyT, por el apoyo de infraestructura computacional y de laboratorio para el adecuado desarrollo de este proyecto.

Agradezco al Ing. Juan Ramón López Samayoa y al parque eólico, por las facilidades otorgadas para el desarrollo del presente proyecto que ha detonado la colaboración con la Universidad de Reading y el Centro Nacional de Ciencias Atmosféricas, ambas instituciones en Reino Unido.

Finalmente, agradezco al CONACyT y a la Secretaría de Energía, por medio del fondo de Sustentabilidad Energética, por brindarme una beca para realizar mis estudios y a la UNAM, a través del Programa de Apoyo a los Estudios de Posgrado, por el apoyo para realizar la estancia de investigación en la Universidad de Reading.

SLPS,

José Gustavo Hernández Yepes

Abstract

Wind power resource assessment (WRA) is an essential tool to promote wind power penetration on current electrical grids. As wind is a variable resource, a reliable WRA to know potential locations to install wind turbines is crucial. Due to limitations on traditional WRA, the development of alternative methods to obtain accurate wind data in a cheaper and quicker way and in broader regions is convenient.

Using wind speed and wind power measurements from a 2 MW wind turbine, we assess the suitability of WRF mesoscale model and MERRA-2 reanalysis to determine the characteristics of wind speed in Chiapas, Mexico and their convenience to reproduce wind energy production. We analyse the wind characteristics at different spatiotemporal scales by varying the resolution of the spatial grids (equivalent to grid spacing of 50 km for MERRA-2 and 1, 3, 15 km for WRF) and temporal resolution of outputs (hourly and 3-hourly).

The numerical models show high correlations for bias corrected wind speeds: near 0.91 for WRF grids and 0.88 for MERRA-2. In addition, the simulated wind speeds along with an accurate power curve fit are able to reproduce the wind turbine capacity factor with an error of 5.54% (MERRA-2) and 4.6% (WRF-1km).

Finally, as part of this study, through a frequency analysis on wind speed and capacity factor, it is clear that the smaller the grid spacing, the more variance present in high frequencies (24-h or higher). Therefore, for synoptic or longer cycles, low spatial resolution grids (75km and MERRA-2) are useful, but to reproduce smaller cycles (diurnal or less), high resolution grids (1 and 3 km) are the optimum.

Contents

Preface	iii
Abstract	v
List of Figures	viii
List of Tables	x
1 Introduction	1
1.1 Objective	6
2 Wind Energy Theoretical Framework	8
2.1 Wind turbines	8
2.2 From gridded data to hub height values	9
2.3 Power Curve	10
2.4 Annual Energy Production	11
2.5 Capacity Factor	11
2.6 Frequency Analysis	12
3 Wind resource from models and observations	16
3.1 Observational Data	16
3.2 MERRA-2 Reanalysis	16
3.3 WRF model	18
3.3.1 Governing Equations	18
3.3.2 Program components of WRF	21
3.3.3 Configuration	22
4 Methodology	28
5 Results and Discussion	30
5.1 Power Curve Fitting	30

- 5.2 Overall Performance of Simulations 31
 - 5.2.1 Wind Speed 31
 - 5.2.2 Wind Power 35
- 5.3 Annual Capacity Factor 38
- 5.4 Influence of spatiotemporal resolution on wind speed and wind power simulations 39
 - 5.4.1 Effect of subsampling on the frequency analysis 43
- 6 Conclusions 48**
 - 6.1 Recommendations and future work 50
- References 51**

List of Figures

1.1	Wind energy in Mexico has growth constantly during the past decade . . .	2
1.2	Part of the equipment used for WRA	2
2.1	Main parts of a horizontal axis wind turbine	9
3.1	Surroundings of Arriaga, Chiapas	17
3.2	Eta Levels of WRF model	19
3.3	Data flow of WPS, and feeding of data to the ARW	22
3.4	Domains of WRF model	23
3.5	Nesting of WRF domains	24
4.1	Methodology	28
5.1	Empirical power curve and error function fit	30
5.2	Wind speed measurements of a wind turbine for 2016 in Arriaga, Chiapas	31
5.3	Wind rose for hourly observations of wind speed (<i>m/s</i>) in 2016	32
5.4	Wind speed measurements and simulations for 2016	32
5.5	Comparisons of wind speed observations against non bias-corrected and bias-corrected wind speeds from models	35
5.6	Comparisons of wind power observations against wind power derived from models	36
5.7	Comparison of observed wind power against calculated wind power de- rived from wind speed measurements	37
5.8	Annual capacity factor for observations, calculations and models	38
5.9	PSD of wind speed observations, hourly sampling	39
5.10	PSD of wind speed derived from models, hourly sampling	40
5.11	PSD of capacity factor observations, hourly sampling	41
5.12	PSD of capacity factor derived from models, hourly sampling	42
5.13	PSD of wind speed observations, 3-hourly sampling	43
5.14	PSD of wind speed derived from models, 3-hourly sampling	44

5.15 PSD of capacity factor observations, 3-hourly sampling	46
5.16 PSD of capacity factor derived from models, 3-hourly sampling	47

List of Tables

3.1	WRF parametrisation and scheme setup used in this work	24
-----	--	----

Introduction

The growing concern regarding global warming in the last few decades is driving a change in the energy systems worldwide[1]. As an example, Denmark and Germany have set the goal of energy systems based 100% on renewable energy towards 2050 [2, 3]. In the case of Mexico, this change is shown by the goal of 50% of clean energy by 2050 [4]. Also, Mexico has pledged to reduce 22% greenhouse gases emissions and 51% black carbon emissions by 2030 [5]. The Ministry of Energy seeks to reduce those emissions through the decarbonization of the energy system and has set the goal to include at least 35% of clean energy sources by 2024 [6].

Within clean energies, wind energy was the second most rapidly growing worldwide, with 23.4% average annual growth rate from 1990 to 2017 [7]. In Mexico, wind power is the second most important source of renewable energy, only behind of hydropower, with 5.74% of the total installed capacity and an average annual growth rate of 10.78% [8]. By 2017, Mexico had 46 wind farms located in the states of Baja California, Zacatecas, Chiapas, Jalisco, Nuevo León, Oaxaca, San Luis Potosi, Tamaulipas and Puebla [9]. Besides, according to the Mexican Wind Energy Association, in 2018 the wind energy installed capacity in Mexico was 4,935 MW (fig. 1.1) and by 2024 is planned to be 14,558 MW [10].

In order to continue the growing trend and establish a solid wind power industry in Mexico, a number of economic, social and technical factors need to be considered [5, 13, 14, 15]. As wind is a variable resource, one of the main technical elements to consider is wind resource assessment (WRA). Moreover, a precise WRA is essential, due to slight variations on wind speed can cause a large deviation on power production[16].

Typically, WRA uses meteorological stations to measure wind speed, wind direction and other atmospheric variables (fig. 1.2). However, the installation, operation and maintenance of stations makes this method expensive, and the need for continuous measurements for extended periods makes the method slow [17, 18]. In addition, the mea-

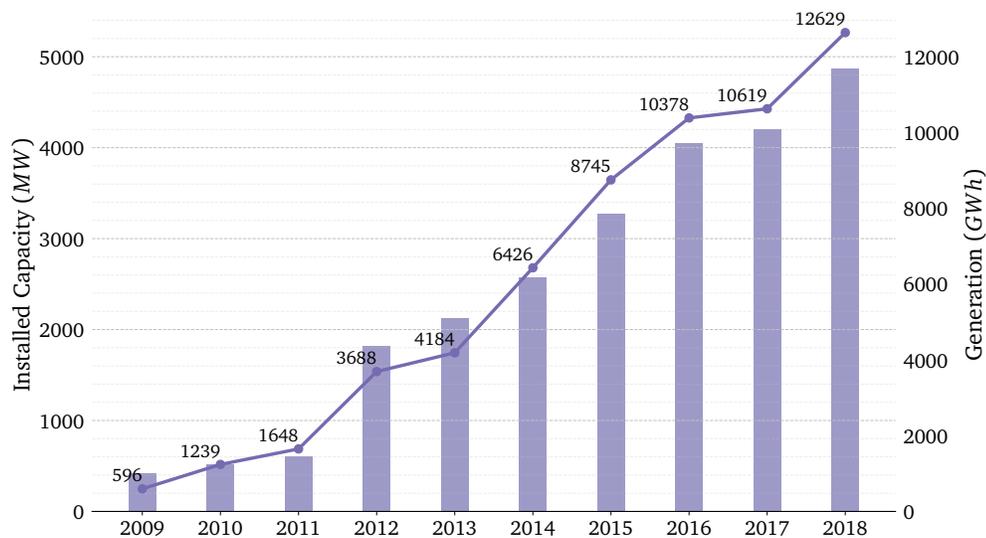


Figure 1.1: Wind energy in Mexico has growth constantly during the past decade [11, 12]

measurements are valid for the location of measurement, but its usefulness to characterise a larger region is less clear, in particular when the topography is not homogeneous (a review of WRA can be found in [16] and [19]). Therefore, the development of alternative methods to obtain accurate wind data in a cheaper and quicker way and in broader regions would be convenient to perform WRA optimally.

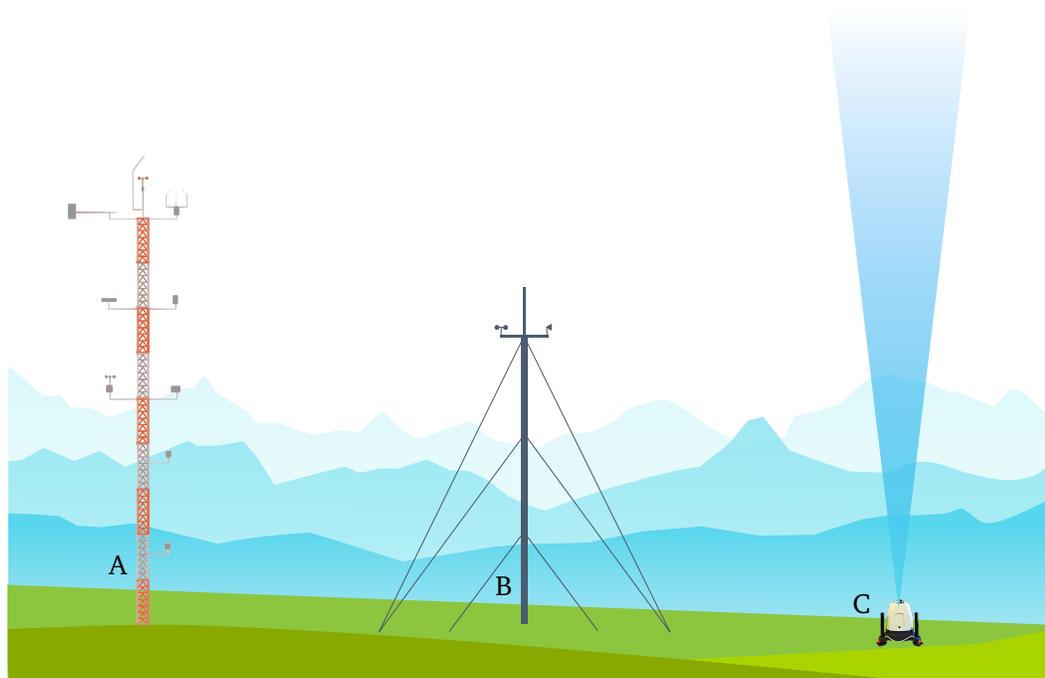


Figure 1.2: Part of the equipment used for WRA: weather stations (A), anemometers (B) and lidars (C)

A popular option to overcome the constraints of traditional WRA, like resolution of

wind measurements, the availability and the accuracy of data is to use Numerical Weather Prediction (NWP) models [17]. NWP models use numerical methods to solve the Navier-Stokes equations, relating atmospheric variables with chemical and physical processes, to predict the state of atmosphere at certain point of time and space.

Naturally, NWP models form the basis for modern weather and climate forecasting at various spatial scales, from global to local, and for various time scales, from hours to centuries. According to the scope of the phenomena of interest, models have different attributes and different applications which can classify them into global (climate), mesoscale and microscale (CFD).

Global climate models (GCM's) are models that reproduce macroscale phenomena. They are obtained by running models of Earth's climate from decades to centuries, covering all the world. Some examples are NASA-GISS-ER ($4^\circ \times 5^\circ$), BCC-CM1 ($1.9^\circ \times 1.9^\circ$) and GFDL-CM2 ($2^\circ \times 2.5^\circ$). Apart from this, another datasets that can be run globally are the reanalyses. The difference with GCM's is that they objectively combine observations with NWP models to generate consistent time series of atmospheric variables for recent decades. Some examples are ERA-Interim (≈ 80 km), MERRA-2 (≈ 50 km) and NCEP-R2 (≈ 210 km).

Mesoscale models are NWP models that produce high resolution weather information; they can be used either for forecasting or research. This kind of models include both the physics and details of terrain that make it possible to simulate the complex atmospheric phenomena at small scales. Some examples are COAMPS (4 km), WRF (up to 1 km) and HRRR (3 km).

In the microscale category we can find several kind of models that run from 10 to 100 meters approximately. Strictly speaking they aren't NWP models, though they follow the same principles. Firstly, there are linear wind flow models such as WASP, MS3DJH/MsMicro3, and MSFD, which follow the simplifications made by Jackson and Hunt [20]. Next in order there are CFD models, as Reynolds-Averaged Navier-Stokes (RANS) models and Large Eddy Simulation (LES) models, both offering a better approach to linear wind flow models. Whereas RANS models only consider the linearised equations of conservation of mass and momentum (without energy equation), the LES models solve the Navier-Stokes equations with full physics parametrisation schemes. Therefore, the last one is able to resolve important eddies of a flow whereas parametrising small ones. In order to be useful, these models must be coupled with mesoscale NWP models, resulting in combinations such as WRF-LES or ARPS-LES [21]. Following, we show the trending models in WRA, their configurations and applications.

Currently, the use of techniques based on the output of NWP models for WRA is being developed for different locations and under a range of models and configurations.

According to the literature review, in this work we will assess the reliability to reproduce wind power production of two types of NWP models that are popular for WRA: reanalyses and mesoscale models.

Reanalyses

As a consequence of the rapid development of technology to produce observational data and model simulations, reanalysis data have been started to be applied widely for wind energy resource assessments [22]. Some of the advantages that reanalyses have for wind power resource assessment are:

- Coverage of large geographic areas, usually global.
- Availability of data for extended time periods.
- Data is accessible for locations where measured data is non-existent.

As an example of the increasing popularity of reanalysis for wind power applications, S. Rose et al. [23] researched about the bias and uncertainties between the Climate Forecast System (CFSR) reanalysis and wind speed observations in the USA. I. Staffell et al. [24] reported hourly power outputs derived from the Modern-Era Retrospective analysis for Research and Applications (MERRA) and MERRA version 2 reanalyses; it stands out that they could estimate power production for potential wind farms anywhere in Europe. They assessed Europe's national aggregated wind output over twenty years as well. D. Cannon et al. [25] assessed wind speed for 33 years using MERRA reanalysis over Great Britain: they aggregated wind speed to calculate hourly capacity factors to quantify wind power variability across GB. Another example is the research of C. Mattar et al. [26]: using data from ERA-Interim reanalysis between 1979 and 2014, they assessed offshore wind energy potential and its technical and economic feasibility in Chile.

In this work we will use MERRA-2 reanalysis [27]. This reanalysis has been found to be reliable, with ease of access, with good spatial and temporal resolution, stable over long time-scale and with heights closer to those of wind turbines [28, 24]. In section 3.2 we will review MERRA-2 in detail.

Mesoscale Models

Mesoscale modelling, often called dynamical downscaling, takes reanalysis as input and generate high spatial resolution forecasts. This is possible because of the inclusion of the type of terrain, land use and synoptic weather conditions. The process of mesoscale

numerical modelling involves many conditions, which have influence on simulated wind speed and therefore on the wind power derived from it.

The first phase for mesoscale models is to choose the **initial and boundary conditions** used to run them. This can be accomplished by means of a global analysis or reanalysis, as they usually have enough dates available and are fairly accurate. Some datasets used frequently are ERA-Interim, ECMWF reanalysis, NCEP-FNL, NCEP-GFS and MERRA [29, 30, 31]. For instance, D. Carvalho et al. [32] used different analysis to provide initial and boundary conditions for WRF in Portugal and then compared the simulated wind speed and wind direction with data from thirteen measurement stations. They used ERA-Interim, NCEP-FNL and NCEP-GFS, and found that the choice of the initial and boundary data supplied to the model constitute a significant error source for simulations.

A remarkable characteristic of the operation of mesoscale models is the incapability to explicitly solve processes smaller than its horizontal or vertical grid scale, processes that are too complex or brief, or processes that are poorly understood. To deal with these limitations, the effects of these processes are included in subgrid equations called **parametrisations**. Parametrisations can include processes related to land surface, vegetation, planetary boundary layer and turbulence, convection, microphysics, radiation in clear skies, cloud cover and radiation in cloudy skies, and orographic drag [33]. The importance of parametrisation relies on the strong influence on outputs and on the indirectly interaction between models' variables. For wind speed for example, Draxl et al. [34] compared the effect of seven planetary boundary layer (PBL) parametrisations on wind energy forecasting. They concluded that the ability of the parametrisations to reproduce mean wind speed and its variability depends on atmospheric static stability and on the terrain. They suggest the best model setup for a forecasting system, in a particular region, depends on the typical distribution of atmospheric stability conditions at the site. Also, they discuss the limitations of PBL parametrisations on complex terrain. These two findings are also reported on [35] and [36].

Along with parametrisations, another essential parameter in mesoscale models is the **resolution of the spatial grid**, i.e. the horizontal distance between points in the numerical grid. It has been found that grid resolution affects the accuracy of 24-h NWP forecasts in complex terrain [36]. Furthermore, some researches conclude that high spatial resolution in NWP can lead to clearly represent complex variations in topography and processes, for example for wind speed [30, 37, 38]. Nevertheless, in some cases having a fine horizontal resolution does not improve the simulations accuracy or it could even weaken the accuracy. [39, 40].

Another parameter useful to analyse atmospheric phenomena is the **temporal resolution**, in model time, at which a new output is generated. The temporal resolution

facilitates the analysis of processes from local to national level and from hourly to inter-annual scales. Many authors have found hourly resolutions convenient for WRA [32, 41, 42, 43, 44, 45]. In some cases, for the WRF model, hourly modeled winds have been able to reproduce the most important characteristics of inter-annual variability, annual and diurnal evolution of a wind profile [29]. Actually, P. de Jong et al. [46] found that uncertainty of NWP models increases with higher temporal resolutions than half an hour.

To sum up, with the correct combination of numerical parameters for mesoscale models, their outputs can be useful to forecast power production, plan energy systems or identify potential sites for wind farms, among many others [46, 47, 48].

In this work we will use WRF model 4.0 [49], as it is open source and free available, with very flexible settings and broadly tested with applications in real time weather forecasting, air chemistry, hydrology, fire weather, cyclones, urban meteorology, solar and wind energy, LES modelling and polar environments [50]. We will review WRF in detail in section 3.3.

1.1 Objective

The wind potential in Mexico has been estimated at 40,000 MW, distributed in the peninsula of Baja California, the central region, the Gulf Coast, the Yucatan Peninsula and the Tehuantepec isthmus [14]. Among these areas, the Isthmus of Tehuantepec is the one that has the best wind resources, yet more research and analysis is needed [51].

The aim of this thesis is to analyse the main characteristics of wind speed and direction, as represented in global reanalyses and model simulations, and assess their convenience for wind power production for the region of Arriaga, Chiapas, which is near the Isthmus of Tehuantepec.

Specific Objectives

1. To compare wind speed and power output of a wind turbine installed at Arriaga, Chiapas, with the outputs of a mesoscale model (WRF) and a reanalysis dataset (MERRA-2) at different resolutions (equivalent to grid spacing of 50 km for MERRA-2 and 1, 3, 15, 75 km for WRF).
2. To determine the amount of meaningful information in terms of indices related to wind power production at different grid resolutions.

Addressing the above mentioned objectives, the rest of the work is organized as follows: the next chapter (2) gives a framework of the concepts related to wind energy and the techniques for its study, chapter 3 gives the characteristics and configurations about the observational data, MERRA-2 and WRF, chapter 4 presents the methodology for the analysis of data, after which results are shown in chapter 5, and finally, chapter 6 concludes this work.

Wind Energy Theoretical Framework

2.1 Wind turbines

Wind turbines are devices that convert the kinetic energy contained in the wind into mechanic energy and then into electricity via a generator. The orientation of the rotating axis can be either horizontal or vertical, giving name to the two types of wind turbines available: horizontal axis wind turbines (HAWT) and vertical axis wind turbines (VAWT).

Due to performance and mature of technology, HAWTs are more commonly found in the mega-watt market than VAWTs. Some drawbacks of the latter are lower power coefficient (C_p), cyclic aerodynamic loads and poor self-starting capability. Nevertheless, small VAWTs have widespread applications in urban or rural areas due to their omnidirectionality, insensitivity to turbulence, cheaper materials and maintenance and less noise [52].

For this work we use measurements from a HAWT, thus in figure 2.1 are shown the main parts of these type of wind turbines. They include the following according to J. Manwell et al. [53]:

- The rotor, the blades and the supporting hub.
- The drive train, which includes the rotating parts of the wind turbine. It may consist of shafts, gearbox, coupling, brakes, and the generator.
- The nacelle wind turbine housing, bedplate, and the yaw system.
- The tower and the foundation.
- The machine controls.
- The balance of the electrical system, including cables, switchgear, transformers, and possibly electronic power converters.

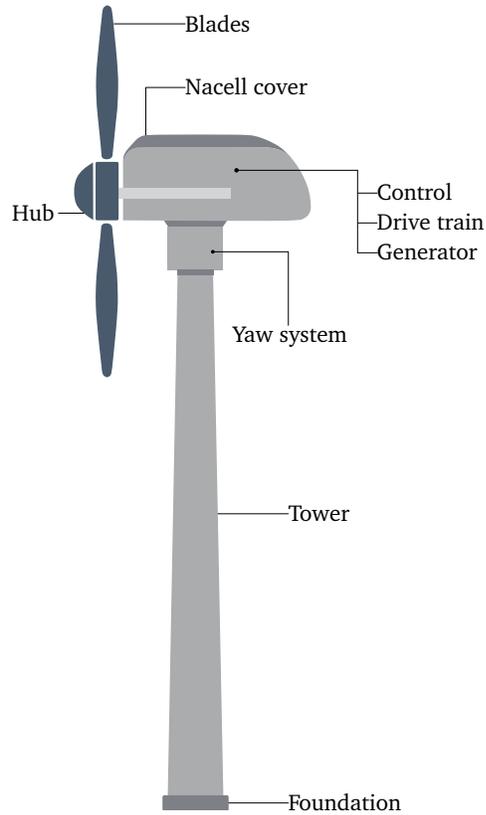


Figure 2.1: Main parts of a horizontal axis wind turbine

2.2 From gridded data to hub height values

Numerical models give the zonal (u) and meridional (v) wind component, so the magnitude of wind speed $|w_s|$ on each grid point is obtained using the following equation:

$$|w_s| = \sqrt{u^2 + v^2} \quad (2.1)$$

In order to compare data of WRF and MERRA-2 against observations, an interpolation is needed both horizontally and vertically. The procedure to convert gridded wind speeds into wind speeds at hub height is as follows[25]:

1. Carry out a bilinear horizontal interpolation to obtain wind speed values at the wind farm location.
2. Assume the wind profile scales logarithmically with height and calculate wind speed data at 90 meters. For WRF, we use a simple log-interpolation using values at 10 meters and at the first layer above 90 m. For MERRA-2, a log-spline (extrapolation) is used with values at 2, 10 and 50 meters height.

Bias Correction

Before calculating wind power production we must verify the modelled wind speeds for known systematic biases relative to real wind speed [54, 55]. This apply for analysis or reanalysis in first place and thus for simulations performed, since we take an analysis as initial conditions for simulations.

Some reasons behind the need for bias correction in reanalysis data are due to: (1) errors in the weather forecast model, (2) the spatial coarseness, which derives in the failure to resolve the detailed topography and thus causing the missing of speed-up and blockage effects and (3) the assimilated observations are not representative of wind farm sites: they are inferred from satellite data and ground observations from short met masts from airports, military bases and others, where there is not wind power production [24].

To avoid systematic biases, the steps we followed to calibrate MERRA-2 and WRF outputs are as follows:

1. Determine the percentile values of measurements and models' outputs.
2. Find the difference between measurements and models for each percentile.
3. Identify the proportional percentile for each value of models' time series.
4. Apply the correction to the models' time series by subtracting a proportional difference from each value.

2.3 Power Curve

A power curve is a graph that represents the relation between wind speed and power production of a specific wind turbine. Though each wind turbine manufacturer provides a power curve, in practice this theoretical power curve does not resembles the real operation of a wind turbine.

To be as accurate as possible, such relation can be approximated using real data and the error function $erf(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^x \exp(-t^2) dt$, where x represents the wind speed and $erf(x)$ represents the wind power production.

Based on measurements of wind speed and power production, the ideal fitting may be calculated through modifying the coefficients of the error function, a_0 , a_1 , a_2 and a_3 . These coefficients will depend on a penalty function, which optimizes the error function by minimizing the square of the difference between observed power and power extracted from the error function, thus:

$$P_W = a_0 \cdot erf\left(\frac{w_s - a_1}{a_2}\right) + a_3 \quad (2.2)$$

where w_s is the wind speed in m/s and P_w is the power in kW .

Beyond the fit of the power curve, to represent the operation of a wind turbine is an elemental factor to model accurately a power curve. We can talk specifically about including the cut-in and cut-out velocities. The first is the velocity at which the wind turbine starts to generate energy and the second is the velocity at which it stops generating due to safety reasons.

2.4 Annual Energy Production

To get the annual energy production (AEP) of a wind turbine, we can take two common approaches. One is through statistical methods and another with the direct use of time series of wind power.

In the first case, we assume a wind turbine that works with a power curve $P_w(w_s)$, and with wind speeds that are distributed with a known probability density function $p(w_s)$, commonly Weibull or Rayleigh [53]. Thus,

$$EAP = T \int_0^{\infty} P_w(w_s) p(w_s) dw_s, \quad (2.3)$$

where T is the operating time of the wind turbine, in this case, the total hours contained in a year.

In this work we will calculate the EAP using observational data directly. For this method we need to know the values of power P_w sampled at intervals Δt , thus,

$$AEP = \Delta t \sum_{i=1}^N P_w, \quad (2.4)$$

where i is the i^{th} value in the time series and N is the total samples of data.

2.5 Capacity Factor

The capacity factor (CF) represents the ratio of the AEP of a wind turbine to the energy that it could have been produced if the wind turbine would have worked at the rated power, P_R , compared in the same period, t (one year) [56]. Thus:

$$CF = \frac{AEP}{P_R \cdot t} \quad (2.5)$$

We will use the CF of a wind turbine as the main parameter to measure its performance. Moreover, the CF can be an indicator of the performance of a wind farm, instead

of a singular wind turbine. The concept is the same: the CF of a wind farm is the rate between the annual energy output of a wind farm to the theoretical energy produced at rated power during a year.

The capacity factor (CF) of a wind turbine is also related to the optimum design of a wind turbine. If a large rotor is coupled to a small generator, it will take any wind and produce a high CF, but with a low yearly energy output. Therefore, an optimal turbine design for a particular place, with a particular wind speed pattern, is needed to get the maximum AEP, even if the CF is not close to one.

2.6 Frequency Analysis

To examine the variability of wind speed or wind power production of a site, we can analyse them in the frequency domain. The spectral analysis helps us to understand the contribution of distinct time scale processes to the fluctuation present in wind speed or CF time series.

Power Spectral Density

Considering that a time series is a stationary signal, we can determine the contribution of every frequency f to the power (variance) present in that signal, namely, obtain its power spectral density (PSD). The basis of this method[53] is that a time series x_n can be expressed as a sum of sines and cosines (Fourier series), as shown in eq. 2.6:

$$\begin{aligned} x_n &= x(n\Delta t) \\ &= A_0 + \sum_{q=1}^{N/2} A_q \cos\left(\frac{2\pi qn}{N}\right) + \sum_{q=1}^{(N/2)-1} B_q \sin\left(\frac{2\pi qn}{N}\right), \end{aligned} \quad (2.6)$$

where Δt is the sample interval, $n = 0, 1, 2, \dots, N - 1$ is the sample number, N is the total samples in the time series, and q is the q^{th} component of the series of sines and cosines

(q^{th} frequency component). Additionally, A_0 , A_q and B_q are called Fourier coefficients:

$$\begin{aligned}
 A_0 &= \frac{1}{N} \sum_{n=1}^N x_n = \bar{x} \\
 A_q &= \frac{2}{N} \sum_{n=1}^N x_n \cos\left(\frac{2\pi qn}{N}\right) \quad q = 1, 2, \dots, \frac{N}{2} - 1 \\
 A_{N/2} &= \frac{1}{N} \sum_{n=1}^N x_n \cos(n\pi) \\
 B_q &= \frac{2}{N} \sum_{n=1}^N x_n \sin\left(\frac{2\pi qn}{N}\right) \quad q = 1, 2, \dots, \frac{N}{2} - 1
 \end{aligned} \tag{2.7}$$

For the analysis in this work, the mean value does not contribute any information about the variation of the signal and therefore is removed ($A_0 = 0$).

A more succinct form of equation 2.6 is

$$x_n = \frac{1}{N} \sum_{k=1}^N X_k \exp\left(j \frac{2\pi kn}{N}\right), \tag{2.8}$$

where the index $q = k$, $\frac{k}{N} \hat{=} f$ and X_k are the Fourier coefficients corresponding to equations 2.7. This form (eq. 2.8) is called the Inverse Discrete Fourier Transform (IDFT), because it transforms a signal from the frequency domain (X_k) to the time domain (x_n).

Conversely, the Discrete Fourier Transform (DFT) is used to obtain the frequency spectrum of a time series and can be written as

$$X_k = \sum_{n=0}^{N-1} x_n \exp\left(-j \frac{2\pi kn}{N}\right). \tag{2.9}$$

Once we know X_k , we are able to obtain the PSD of the signal x_n , $S_{xx}(f)$:

$$S_{xx}(f) = \frac{1}{N\Delta f} |X_k|^2 \tag{2.10}$$

The normalization of the DFT by the frequency bin width allows a correct comparison of signals that have different lengths. That is to say that we get rid of the dependency on bin width, so the variance per bin width is comparable regardless of the maximum frequency in the signals. Therefore, the integral of the PSD over a frequency band measures the average variance of a signal associated with that frequencies, and the integral of $S_{xx}(f)$ over the full spectrum is equal to the total variance of the signal x_n .

Red Noise

Red noise is a definition of a special noise that has enhanced low frequency fluctuations, derived from the interaction of white noise with slow response components of a system. The name is due to an analogy to visible light reduced in the short wavelengths, which is red. Red noise is commonly found as a background noise in climatic and hydrological time series, such as sea and air temperature and precipitation data [57].

As Gilman et al. describe in his research [58], a discrete red noise signal X can be expressed as a first order autoregressive process (AR1), thus,

$$X_n = \rho X_{n-1} + y_n, \quad (2.11)$$

where n is the n^{th} sample, y_n is a white noise component and ρ is the average lag-one autocorrelation coefficient, given as

$$\rho = \frac{\overline{X_n X_{n-1}}}{\overline{X_n^2}}. \quad (2.12)$$

Following the procedures in [58] the parametric estimation of the red noise spectrum L is obtained as:

$$L_h = \frac{1 - \rho^2}{1 - 2\rho \cos \frac{h\pi}{M} + \rho^2}, \quad (2.13)$$

where h is the frequency and M is the maximum lag. Finally, to get a precise level of red noise and following [57], we multiply the red noise PSD by the average of the PSD of the time series.

In the frequency analysis that we develop, we use the red noise power spectrum as null hypothesis to know how likely a peak is a periodic component of the time series or if it is just red noise. We assume that the red noise present in the time series can be treated as locally white, therefore its power spectrum is χ^2 distributed with $\nu = 2$ degrees of freedom. The ratio of power of a peak in the power spectrum to the red noise level can be assumed to be distributed as $\frac{\chi^2}{\nu}$, which can be compared with the tabulated probability distribution of χ^2 to determine peak significance [57, 59, 60].

For instance, if we choose a confidence level of 0.99, we have a 0.99 probability of a point of the PSD to be red noise. This means that we expect for 1% of the frequencies the null hypothesis to be false (1% of the PSD above the confidence level), that is to say, 1% of the spectrum is expected to be statistically significant and to be possibly cyclic components.

In this chapter we presented the concepts and theory behind this thesis. We addressed the procedures to calculate wind energy production and CF, the process of obtaining hub height values from models' data, how to calibrate and compare them, and the frequency analysis of time series and the concepts to evaluate it.

In the next chapter we describe the observational and models data. We will review the characteristics of the MERRA-2 reanalysis and the theory and configurations of WRF model simulations.

Wind resource from models and observations

In this chapter we will present the characteristics of the observational data to analyse. We will also present the general features of the MERRA-2 reanalysis and finally the theory, software components, parametrisations and configurations related to the WRF model simulations.

3.1 Observational Data

Measurements are available from a wind turbine located at the municipality of Arriaga, in the state of Chiapas in southern Mexico (figure 3.1). Arriaga is located between the Sierra Madre de Chiapas and the Pacific coastal plain. Its climate is warm sub-humid with rains in the summer and with a mean temperature of 27°C [61].

As to the wind turbine, the hub is 90 meters height with a rated power of 2 MW. Its wind speed and wind power data are for 2016 and taken every 10 minutes, then averaged hourly and 3-hourly for the frequency analysis (section 5.4.1).

3.2 MERRA-2 Reanalysis

The Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) is the latest reanalysis developed by NASA, and the next version of the original MERRA (January 1979 – February 2016) reanalysis produced by the global Modelling and Assimilation Office (GMAO). MERRA-2 covers from January 1980 to the present, with more than 40 data collections.

MERRA reanalysis version 2 intends to improve known deficiencies of its predecessor through the integration of recent developments of modelling and data assimilation, and

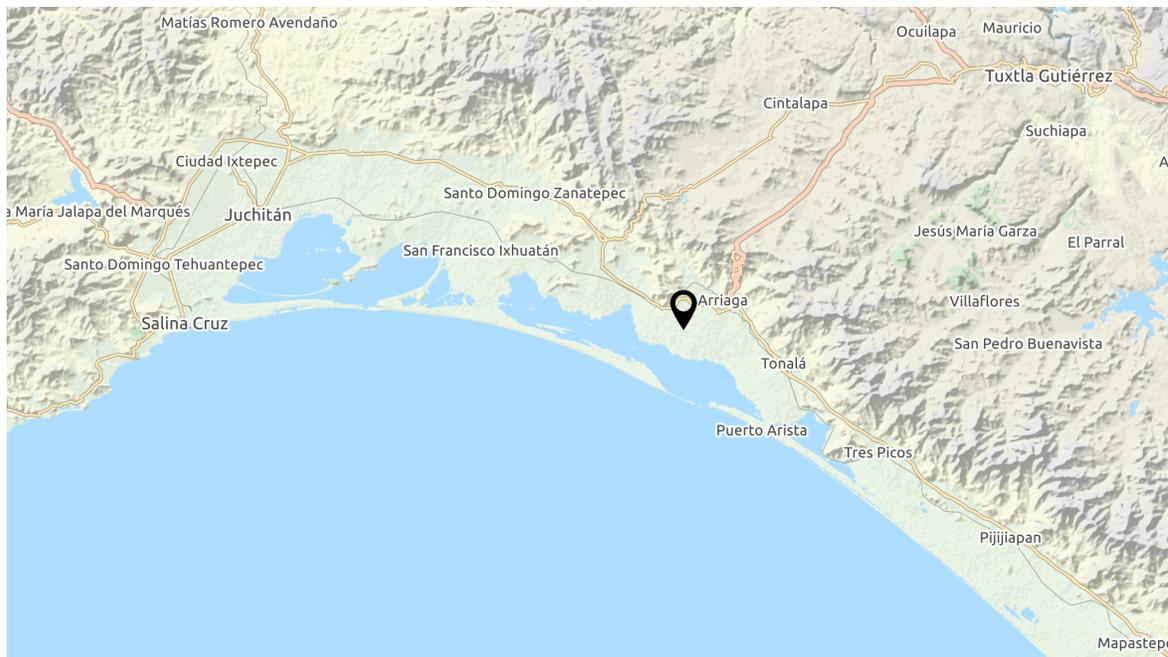


Figure 3.1: Surroundings of Arriaga, Chiapas. The black pin indicates the location of the wind turbine under study. (Maps © Thunderforest, Data © OpenStreetMap contributors)

at the same time to establish the basis of an Integrated Earth System Analysis (IESA) that will be able to couple the assimilation systems for the atmosphere, ocean, land and chemistry. The weaknesses that were solved include unphysical trends in precipitation, biases in some atmospheric and land surface processes, and a poor representation of the upper stratosphere. The updates made to MERRA include changes to the forecast model, the analysis algorithm, the observing sources, the radiance assimilation, the bias correction of aircraft observations, the mass conservation and water balance, the precipitation used to force the land surface and wet aerosol deposition, the boundary conditions for sea surface temperature and sea ice concentration, and reanalysis production.

Even though MERRA-2 addresses some limitations of MERRA, there are still some aspects that require attention due to the ongoing development of modelling and data assimilation techniques. Some progress are the assimilation of cloud and rain affected radiances (not only clear skies), direct assimilation of land surface observations, the implementation of a land model that includes dynamic phenology and photosynthesis, the developing of coupled ocean–atmosphere analysis system to improve the specification of ocean surface boundary conditions, and an accurate representation of aerosol effects on climate. All these efforts are made towards reducing the uncertainty and bias, and therefore to improve the quality of historical datasets for applications as significant as climate change.

MERRA-2 uses the Goddard Earth Observing System (GEOS) for atmospheric data

assimilation and the Gridpoint Statistical Interpolation (GSI) analysis scheme. The use of these systems result in a discretization of a resolution of $0.5^\circ \times 0.625^\circ$, equivalent to about 50 km in latitudinal direction, and 72 hybrid-eta levels from the surface to 0.01 hPa. The analysis produces 6-hourly outputs, but due to assimilation techniques MERRA-2 can produce higher frequency products. Also, the vertical grid can be the native 72 layers or a grid of 42 interpolated standard pressure levels [62].

The dataset that we use [27] has hourly outputs and includes values of zonal and meridional wind speed at 2, 10 and 50 meters over the terrain height, therefore there is no need to perform extra calculations to get wind speed at fixed heights.

3.3 WRF model

The Weather Research and Forecasting (WRF) model is an atmospheric modelling system developed by the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (NOAA) (represented by the National Centers for Environmental Prediction (NCEP) and the NOAA Earth System Research Laboratory (ESRL), the United States Air Force, the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration.

3.3.1 Governing Equations

WRF uses several equations to describe the motion of fluids in the atmosphere. Here, we present the basic governing equations, which are adapted to the model and presented later in this section along with additional equations.

First we present the **Navier-Stokes equations (conservation of momentum)** for an inviscid ($\mu = 0$), compressible ($\nabla \cdot \mathbf{v} \neq 0$) and non hydrostatic fluid ($\frac{\partial p}{\partial z} \neq -\rho g$):

$$\rho \frac{D\mathbf{v}}{Dt} = \rho \mathbf{g} - \nabla p + \mu \nabla^2 \mathbf{v}, \quad (3.1)$$

where ρ is the density of the fluid, $\frac{D}{Dt}$ is a material derivative, $\mathbf{v} = (u, v, w)$, g is the gravitational acceleration and p is the hydrostatic pressure.

WRF also uses the **conservation of energy** equation, for an adiabatic fluid ($\nabla^2 T = 0$):

$$\rho \frac{DT}{Dt} = k \nabla^2 T + \phi \quad (3.2)$$

where T is the temperature, k is the thermal diffusivity and ϕ is a source term.

Also there is the **conservation of mass** equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{v} = 0 \quad (3.3)$$

The conservation of moisture, the material derivative of geopotential and the ideal gas equation of state will be described in the following lines.

Additionally, the above equations are formulated for WRF using a terrain-following hydrostatic-pressure vertical coordinate denoted by η , defined as

$$\eta = \frac{p_d - p_t}{p_s - p_t}, \quad (3.4)$$

where p_d is the hydrostatic pressure of dry air, p_s is the value of p_d in the surface boundary and p_t is the value of p_d in the top boundary. Unlike traditional sigma coordinates, WRF hybrid sigma coordinate, η , allows the influence of the terrain on the horizontal surfaces to be removed more rapidly with increasing height, as illustrated in fig. 3.2.

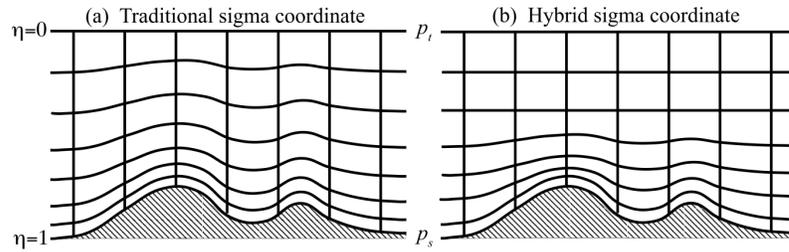


Figure 3.2: Eta Levels of WRF model [49]

In this scheme of hybrid sigma coordinates, the hydrostatic pressure is defined as

$$p_d = B(\eta)(p_s - p_t) + [\eta - B(\eta)](p_0 - p_t) + p_t, \quad (3.5)$$

where p_0 is a reference sea level pressure, $B(\eta)$ defines the transition between the terrain-following sigma coordinate and a pure pressure coordinate, such that η corresponds to the sigma coordinate near the surface, ($B(\eta) = \eta$), and to a pressure coordinate at upper levels ($B(\eta) = 0$). B is defined as a polynomial: $B(\eta) = c_1 + c_2\eta + c_3\eta^2 + c_4\eta^3$, such that

$$c_1 = \frac{2\eta_c^2}{(1 - \eta_c)^3}, \quad c_2 = \frac{-\eta_c(4 + \eta_c + \eta_c^2)}{(1 - \eta_c)^3}, \quad c_3 = \frac{2(1 + \eta_c + \eta_c^2)}{(1 - \eta_c)^3}, \quad c_4 = \frac{-(1 + \eta_c)}{(1 - \eta_c)^3}, \quad (3.6)$$

where the vertical coordinate metric is defined as

$$\mu_d = \frac{\partial p_d}{\partial \eta} = B_\eta(\eta)(p_s - p_t) + [1 - B_\eta(\eta)](p_0 - p_t), \quad (3.7)$$

and where the subscript η denotes differentiation and $\mu_d \Delta \eta = \Delta p_d = -g \rho_d \Delta z$ is proportional to the mass per unit area within a grid cell. Thus, the flux form for prognostic variables are defined as

$$\mathbf{V} = \mu_d \mathbf{v} = (U, V, W), \quad \Omega = \mu_d \omega, \quad \Theta_m = \mu_d \theta_m, \quad Q_m = \mu_d q_m, \quad (3.8)$$

where $\mathbf{v} = (u, v, w)$ are the covariant velocities and $\omega = \dot{\eta}$ is the contravariant vertical velocity. Additionally, θ is the moist potential temperature and Q_m represents the mixing ratios of moisture variables (water vapour ($m = v$), cloud water ($m = c$), rain water ($m = r$) and so on.

The above considerations result in the following equations:

Conservation of momentum:

$$\frac{\partial U}{\partial t} + (\nabla \cdot \mathbf{V}u) + \mu_d \alpha \frac{\partial p}{\partial x} + (\alpha/\alpha_d) \frac{\partial p}{\partial \eta} \frac{\partial \phi}{\partial x} = F_U \quad (3.9)$$

$$\frac{\partial V}{\partial t} + (\nabla \cdot \mathbf{V}v) + \mu_d \alpha \frac{\partial p}{\partial y} + (\alpha/\alpha_d) \frac{\partial p}{\partial \eta} \frac{\partial \phi}{\partial y} = F_V \quad (3.10)$$

$$\frac{\partial W}{\partial t} + (\nabla \cdot \mathbf{V}w) - g \left((\alpha/\alpha_d) \frac{\partial p}{\partial \eta} - \mu_d \right) = F_W \quad (3.11)$$

Conservation of energy:

$$\frac{\partial \Theta_m}{\partial t} + (\nabla \cdot \mathbf{V}\theta_m) = F_{\Theta_m} \quad (3.12)$$

Conservation of mass:

$$\frac{\partial \mu_d}{\partial t} + (\nabla \cdot \mathbf{V}) = 0 \quad (3.13)$$

Conservation of moisture:

$$\frac{\partial Q_m}{\partial t} + (\nabla \cdot \mathbf{V}q_m) = F_{Q_m} \quad (3.14)$$

Material derivative of geopotential:

$$\frac{\partial \phi}{\partial t} + \mu_d^{-1} [(\mathbf{V} \cdot \nabla \phi) - gW] = 0, \quad (3.15)$$

The terms at the right side of the equations 3.9-3.12, F_U , F_V , F_W and F_{Θ} represent forces that arise from model physics, turbulent mixing, spherical projections and earth's rotation.

Along with equations 3.9 - 3.15, there is the diagnostic equation for dry hydrostatic

pressure

$$\frac{\partial \phi}{\partial \eta} = -\alpha_d \mu_d, \quad (3.16)$$

and the relation for the full pressure (derived from the ideal gas equation of state), combining dry air and water vapour:

$$p = p_0 \left(\frac{R_d \theta_m}{p_0 \alpha_d} \right)^\gamma \quad (3.17)$$

In the above equations, $\alpha_d = 1/\rho_d$, is the inverse density of the dry air, $\alpha = \alpha_d(1 + q_v + q_c + q_r + q_i + \dots)^{-1}$ is the inverse density taking into account the full parcel density, $\phi = gz$ is the geopotential, $\gamma = C_p/C_v = 1.4$ is the ratio of the heat capacities for dry air, R_d is the gas constant for dry air, and p_0 is a reference pressure (typically 10^5 Pascals).

The operations $\nabla \cdot \mathbf{V}$ and $\mathbf{V} \cdot \nabla$ are defined as

$$\nabla \cdot \mathbf{V}a = \frac{\partial Ua}{\partial x} + \frac{\partial Va}{\partial y} + \frac{\partial \Omega a}{\partial \eta}, \quad (3.18)$$

and

$$\mathbf{V} \cdot \nabla a = U \frac{\partial a}{\partial x} + V \frac{\partial a}{\partial y} + \Omega \frac{\partial a}{\partial \eta}, \quad (3.19)$$

where a is a generic variable.

It is worth mentioning that for simplicity, the equations that contain map projections, Coriolis effect and curvature terms are not shown here, however they can be examined in [49].

3.3.2 Program components of WRF

In order to run a real data case, the WRF model must first obtain meteorological and static terrestrial data, together with initial and lateral boundary conditions. This task is achieved by using the WRF Preprocessing System (WPS) coupled with the Advanced Research (WRF) ARW dynamical core (fig. 3.3).

The WPS is a collection of programs that takes terrestrial and meteorological data and transforms them for input to the real ARW preprocessor program. The first component of WPS is *geogrid*, which defines the size and location of domains and interpolates static terrestrial fields to domain grids. The next is *ungrib*, which extract meteorological fields (time dependent) from GRIB files coming from an analysis or reanalysis data set. The last one is *metgrid*, which interpolates meteorological fields from *ungrib* to the grids defined by *geogrid*.

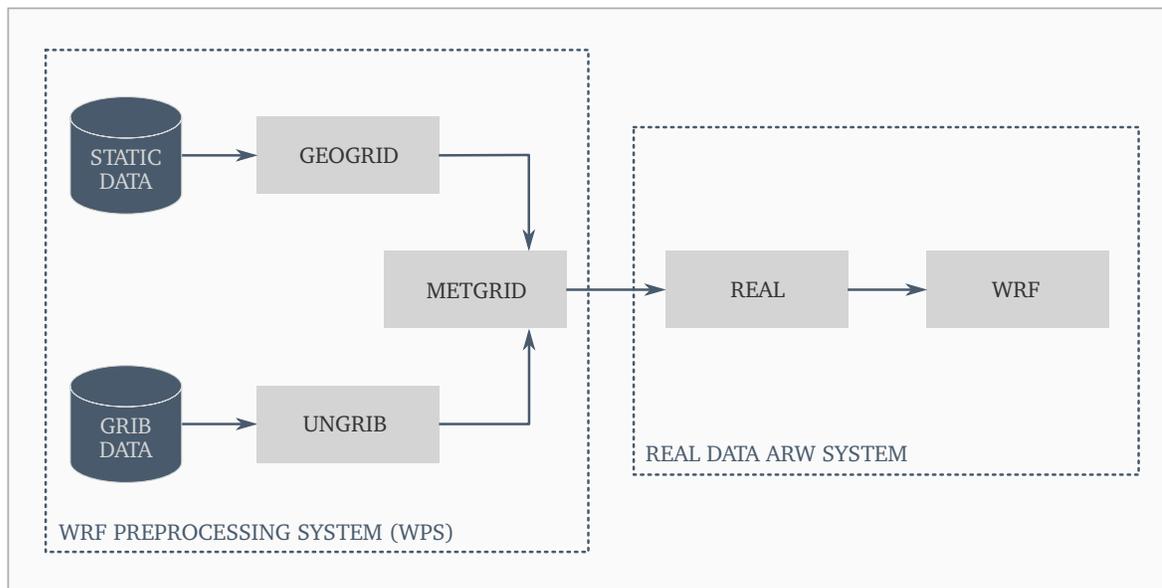


Figure 3.3: Data flow of WPS, and feeding of data to the ARW [49]

The output of the WPS is used by the ARW preprocessor of real data cases to build initial and boundary conditions. The initial conditions are built for the beginning of the simulation and lateral conditions are built for every period of the analysis used in WPS. Subsequently, the ARW solver can be ran to process the corresponding equations.

3.3.3 Configuration

WRF simulations are run every 24 hours, with 6 hours of spin-up. In other words, WRF is run for 30 hours, but the first six hours are ignored to let the model reach to an equilibrium state. Additionally, a daily restarting removes the errors from each integration period and allows the model not to diverge from the analysis over the integration time.

In this work we use four domains: 75 km, 15 km, 3 km and 1 km (fig. 3.4 and 3.5), with the NCEP-FNL analysis [63] as input for boundary and initial conditions. The output of WRF is every hour for all the domains, excepting the 75 km domain, which has 3-hourly output.

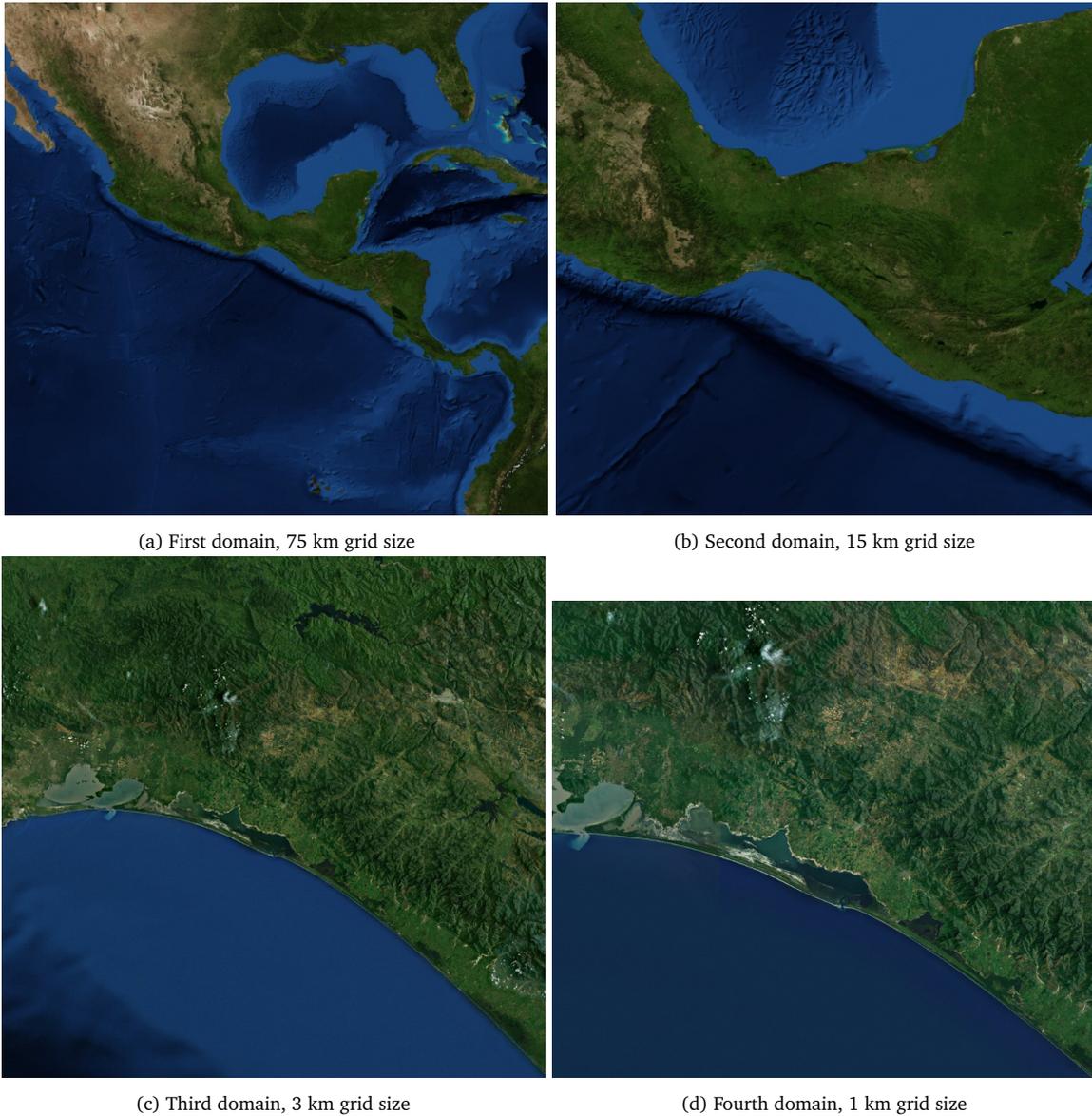


Figure 3.4: Domains of WRF model

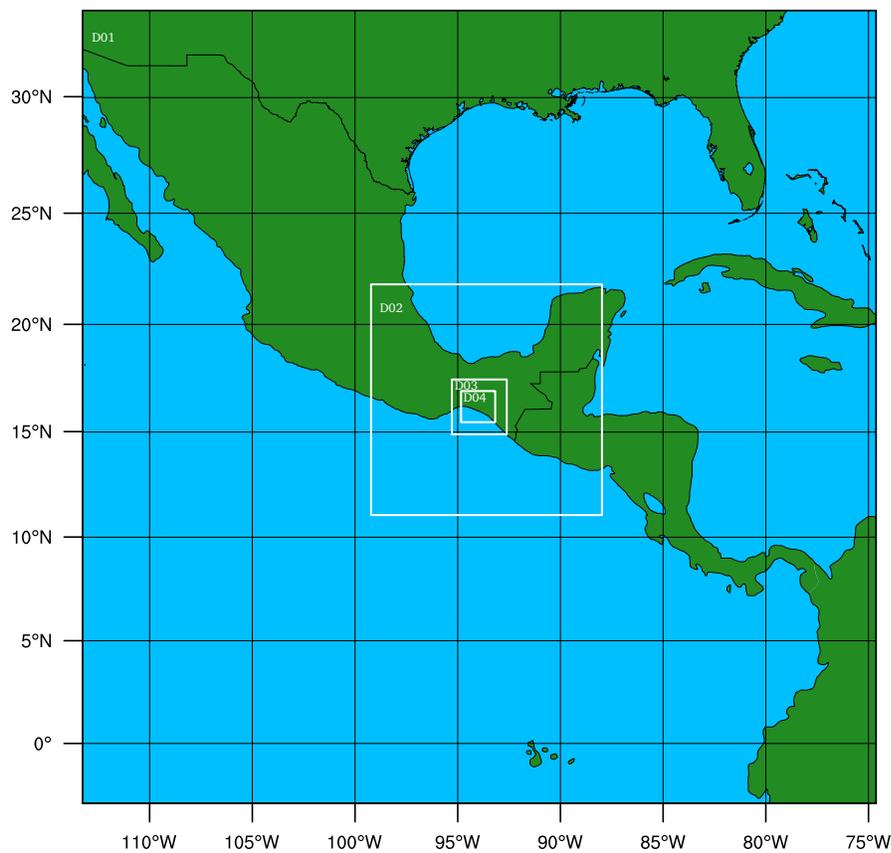


Figure 3.5: Nesting of WRF domains

Parametrisations

As mentioned in the introduction, the parametrisation of physical processes is vital to get accurate results for a NWP model. Table 3.1 shows the set of options used in our run and in the following lines we explain the basic concepts about them.

Table 3.1: WRF parametrisation and scheme setup used in this work

Parametrisation or scheme	Reference
Dudhia shortwave radiation	Dudhia (1989) [64]
Rapid Radiative Transfer Model for longwave radiation	Mlawer et al. (1997) [65]
WSM3 microphysical parametrisation	Hong et al. (2004) [66]
Kain-Fritsch convective parametrisation	Kain (2004) [67]
YSU planetary boundary layer scheme	Hong et al. (2006) [68]
Revised MM5 surface layer scheme	Jimenez et al. (2012) [69]
Unified Noah land surface model	Tewari et al. (2004) [70]

Radiation-shortwave: Dudhia The shortwave radiation originates from the radiation coming from the sun (shortwave) at wavelengths between 0.1 and 4 μm (visible spectrum

mainly). While this radiation moves through the atmosphere it can be diffused, absorbed or reflected. Additionally, at the moment of contact with the earth's surface, a significant amount of it can be reflected back to the atmosphere, whereas other can be absorbed, depending on the albedo of the surface [71].

Dudhia shortwave scheme only calculates the downward radiation flux, using a simple radiative transfer model, with no horizontal communication. It includes absorption and scattering in clouds and in clear sky. To avoid long calculations, it uses static look-up tables, previously generated by a much more expensive scheme, to compare flux and cooling rates.

Radiation-longwave: RRTM The longwave radiation originates from the energy that the atmosphere and the earth radiates after absorbing energy from the shortwave radiation. The earth and the atmosphere have a much lower temperature than the sun, thus, the Earth's emitted radiation has longer wavelengths (in the infrared spectrum mainly). Again it can be subject to absorption, scattering and reflection [71].

The Rapid Radiative Transfer Model parametrisation (RRTM) solves the radiative transfer equation for upward and downward fluxes with calculations performed only in columns with no horizontal communication. It takes into account multiple absorption bands, trace gases and microphysics. Also, to avoid long calculations, it uses static look-up tables, previously generated by a much more expensive scheme, to compare flux and cooling rates.

Microphysics: WRF Single-moment 3-class scheme The microphysics solves the aerosol-radiation-water interactions. This parametrisation is about how these particles are involved in the formation of clouds and their classification and evolution in time. Derived from the physical processes of particles like condensation, collision-coalescence and nucleation, these particles can be either liquid or solid and with different sizes and shapes. Thus, these physical processes define the growth of cloud particles, as well as any transformation of phase like precipitation [72].

The WRF Single-moment 3-class scheme is a simple and efficient scheme that includes ice and snow processes, suitable for mesoscale grid sizes.

Cumulus: Kain-Fritsch (2004, JAM) The cumulus physics parametrisation covers the aggregated effects of the convection related to subgridscale cumulus clouds. This parametrisation involves processes as convective initiation, deep convection, shallow convection, updrafts, downdrafts and precipitation.

The (new) Kain-Fritsch parametrisation is a scheme that includes deep and shallow

convection, cloud entrainment and detrainment. It uses a mass flux approach with up-drafts and downdrafts, with CAPE removal time scale closure. This parametrisation is only applicable to the 75-km and 15-km grids, because the cumulus convection is solved explicitly for smaller grids.

Boundary layer: YSU Hong et al. (2006, MWR) The planetary boundary layer (PBL) parametrisation tries to reproduce the development of the PBL through the parametrisation of the vertical turbulent diffusion of the subgrid-scale fluxes of heat, momentum, and vapour. As wind turbine hub heights are within the PBL during most part of the day, the PBL parametrisation is an essential part in the simulations of wind speed.

The YSU scheme is a first order closure, non local bulk scheme. The characteristic of this scheme is that at the top of the PBL, it explicitly determines the top layer at the level at which minimum turbulent flux exists (heat, momentum and water vapour). Also, to allow for non-local vertical fluxes, a non-local term is added.

Surface layer: revised MM5 The surface layer is the part of the PBL that is in direct contact with the earth's surface. This parametrisation act as a link between the land surface and the boundary layer: it provides surface heat, humidity and momentum fluxes, from the model surface to the boundary layer. This is done through the calculation of friction velocities (log wind profile) and flux exchange coefficients.

The revised MM5 scheme follows the Monin-Obukhov similarity theory. It comes from the standard relations used in the parametrisations of MM5 model, but has updated stability functions and it is modified to cover the full range of atmospheric stability.

Land surface: Noah Land Surface Model The land surface parametrisation is intended to provide boundary conditions to the atmosphere in terms of heat and humidity fluxes. It tries to represent the complex biophysical, hydrological, and bio-geochemical interactions between the land-surface and the atmosphere. It has impact in several processes as surface radiation, low level clouds, the wetness of soil, and thermal balance.

The unified NCEP/NCAR/AFWA scheme takes into account four layers for calculating soil temperature and moisture. It uses algorithms for fractional snow cover and for the computation of the physics in frozen soil.

In this chapter we presented the characteristics of the observational data, of MERRA-2 reanalysis and of the WRF model. A description of grids and models' configurations was also presented. This description include governing equations, grids location and resolution, parametrisations and work flow of the WRF model.

Once we have presented the datasets we are going to work with, in the next chapter we show the methodology we will follow to accomplish the objectives of this work.

Methodology

In this chapter, we will explain the methodology followed in this work to compare and analyse observational data of a wind turbine against data from MERRA-2 and WRF at different spatiotemporal scales (fig. 4.1).

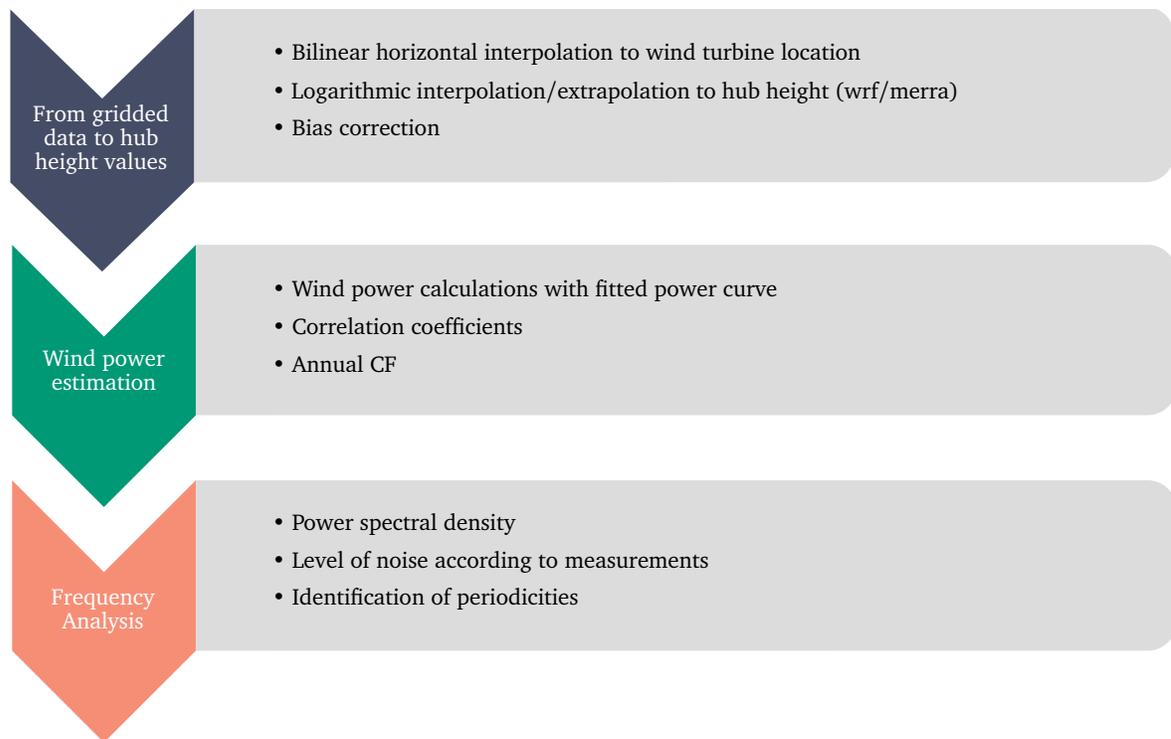


Figure 4.1: Methodology

The first step to be able to compare models' data against real measurements, is to obtain values of wind speed and wind power available in a uniform numerical grid and transform them into a time series of wind speed at hub height and a time series of wind power derived from it. Therefore, a horizontal bilinear interpolation and a log scale estimate is performed to get wind speed at hub height. Before calculating wind power

production, we remove systematic errors of wind speed data with a bias correction and we look at its effect on the distribution of data. We compare wind speeds and power production observations against models' outputs with the Pearson correlation coefficient and through scatter plots to visualize bias and distribution of data.

To obtain wind power production, we need to apply the bias corrected wind speed to a power curve. Thus, we use an error function to fit the wind power measurements as a function of wind speed and then calculate the correspondent wind power production. Subsequently, we compare the outputs of the models with real measurements to see how they correlate. Also, we obtain annual CF for each numerical grid, for wind power produced from wind speed measurements applied to the power curve and for wind power measurements.

Finally, we get the power spectral density for every time series of wind speed and wind power, and calculate the level of red noise present in measurements. In this analysis, we estimate the PSD of both models and observations (with hourly data). Then, based on the background level of red noise in observations, we are able to detect periodicities and evaluate the effect of spatial resolution of models on the distribution of power at different time scales. We apply this methodology for wind speed and CF. At last, we calculate the PSD for 3-hourly subsampled time series to assess the impact of subsampling on the distribution of variance, for both wind speed and CF. Again, we use the red noise level of measurements to compare the spectra of different grid sizes.

In this brief chapter we have described the methodology of this work. The first stage is to convert gridded data to be comparable with observations, so we can analyse the relation of wind speed observations and models values. Next in order, a power curve is proposed and wind power is computed, along with CF. Finally, we perform the frequency analysis of wind speed and CF for hourly and 3-hourly time series.

The next chapter shows the results obtained and their discussion.

Results and Discussion

5.1 Power Curve Fitting

As a start point we assessed the relation between observed wind speeds and observed power production for the wind turbine in 2016. Figure 5.1 shows the the empiric power curve (scattered points) and and its fitting (red line).

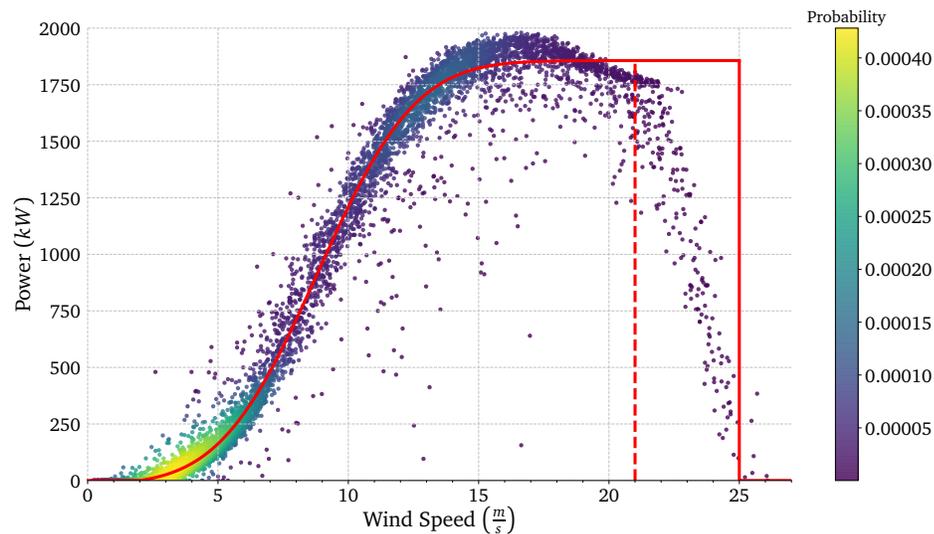


Figure 5.1: Empirical power curve (color represents the probability for each point) and in red the fit using the error function. Similarly as Cannon et al., the red dashed line represents the wind speed at which the wind turbine starts power production after the cut-out speed of 25 m/s [73].

The power curve reveals two concentrations of values that represent two main states of the operation of the wind turbine. The first for less than 500 kW and the second around 1750 kW. This plot also shows a non desired response in the wind turbine, as it is producing less than the rated power for wind speeds higher than 20 m/s. Hence, to fit the power curve we did not consider the values above 20 m/s.

Using the empiric power curve with the above consideration and based on equation 2.2, the best fit is as follows:

$$P = 937.7 \cdot \operatorname{erf}\left(\frac{w_s - 8.826}{4.159}\right) + 918.9 \quad (5.1)$$

where w_s is the wind speed in m/s and P is the power in kW .

5.2 Overall Performance of Simulations

In this section we will address one of the main purposes of this thesis, which is to compare wind speed and wind power production from models' and observations of a wind turbine in Arriaga, Chiapas for 2016.

5.2.1 Wind Speed

The 2016 wind speed observations have two seasonal trends: the first, with remarkable peaks during winter and spring, and the other with lower crests during summer and autumn (fig. 5.2). At first glance, this mean that in spring and winter seasons we can have higher energy production than in summer and autumn.

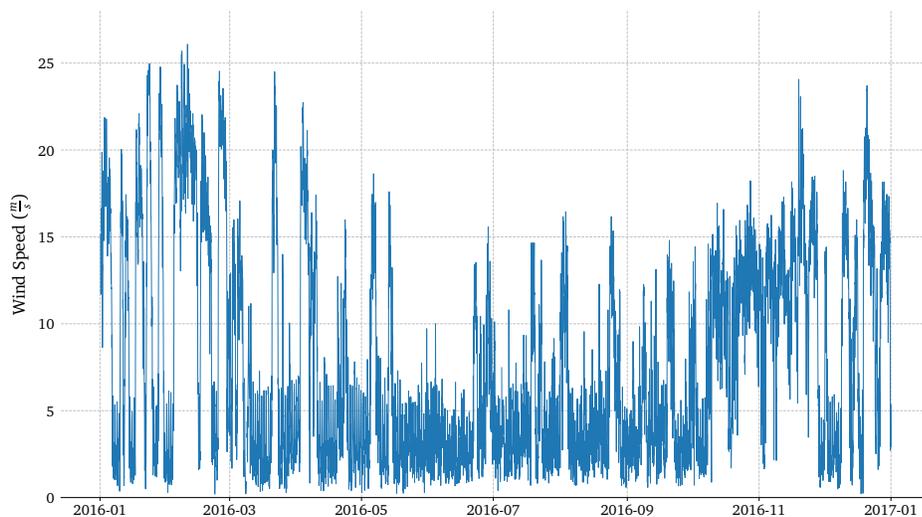


Figure 5.2: Wind speed measurements of a wind turbine for 2016 in Arriaga, Chiapas

Figure 5.3 shows the distribution of wind speed observations regarding its azimuthal direction, with the length representing the frequency percentage of wind in that direction. From this wind rose and the power curve (fig. 5.1), it can be seen that much of the wind used for power production (5-25 m/s) comes from the north-northeast direction.

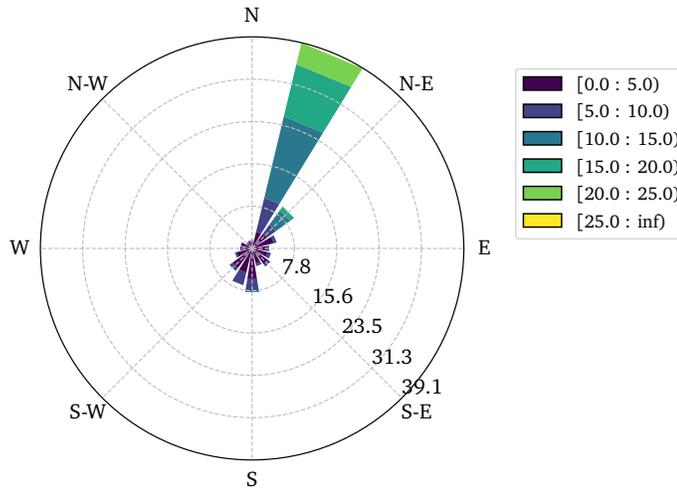


Figure 5.3: Wind rose for hourly observations of wind speed (m/s) in 2016

WRF model and MERRA-2 interpolated wind speeds were obtained at hub height with 3-hourly resolution for the 75 km grid and hourly for the 15, 3, 1 km grid and MERRA-2 data. We found that wind speeds derived from bias corrected models, follow the overall patterns and are correlated with observations to a certain extent (fig. 5.4).

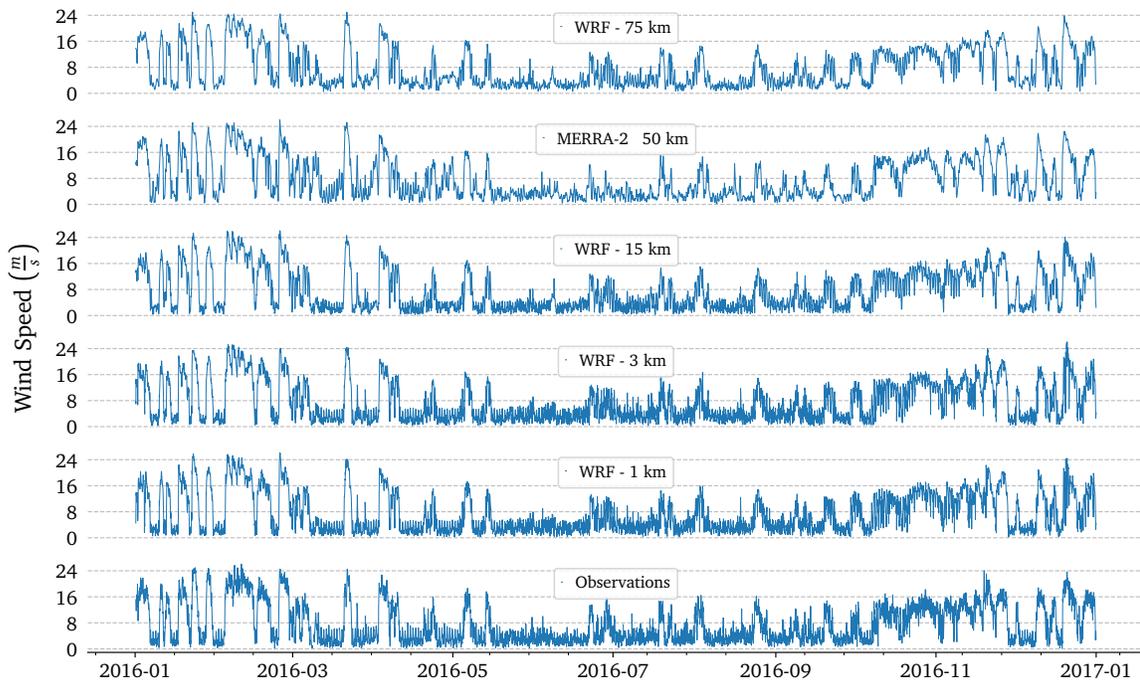


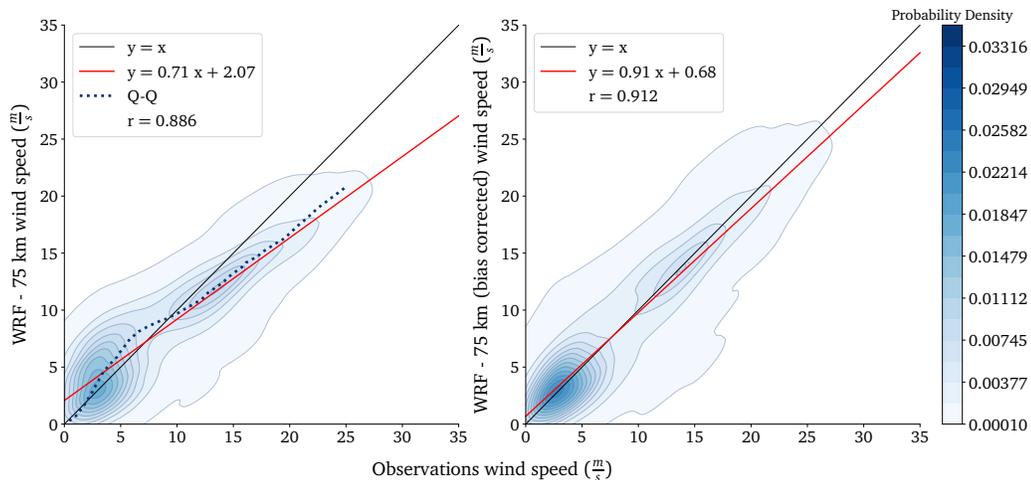
Figure 5.4: Wind speed measurements and simulations for 2016. Models time series reflect the close inter-connection with the observed wind speed. Nevertheless, this comparison is not enough to assess the grids in use.

To have an exact estimation for the correlation of simulations with real values we can

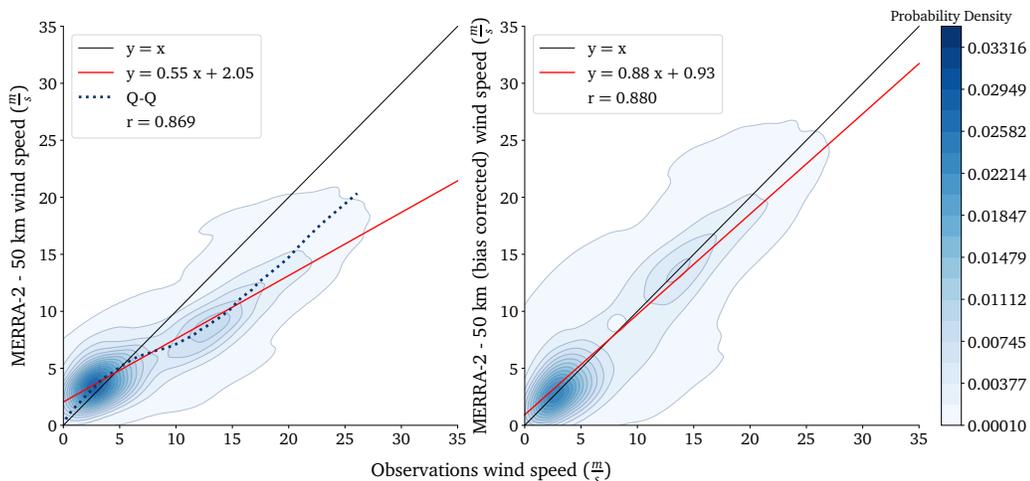
look to the scatter plots (fig. 5.5), which compare models and observations, for both non bias-corrected data and for bias-corrected data. The first impression from all figures in 5.5, is the reproduction of a bimodal distribution of data, with a major cluster of values located under 5 m/s, and a second one near 15 m/s.

Regarding the raw output of models, all grid sizes show a good correlation against observations, with correlation coefficients above 0.86. Whereas large grid sizes present underestimation of high wind speeds, small grids present overestimation. The 75 km WRF grid underestimate wind speeds higher than 10 m/s and MERRA-2 presents considerable deviations, underestimating values higher than 5 m/s. On the other hand, the 15, 3 and 1 km WRF grids overestimate wind speeds greater than 5 m/s. In this case, the best is the 3 km grid as it has the best least squares fit and $r=0.908$.

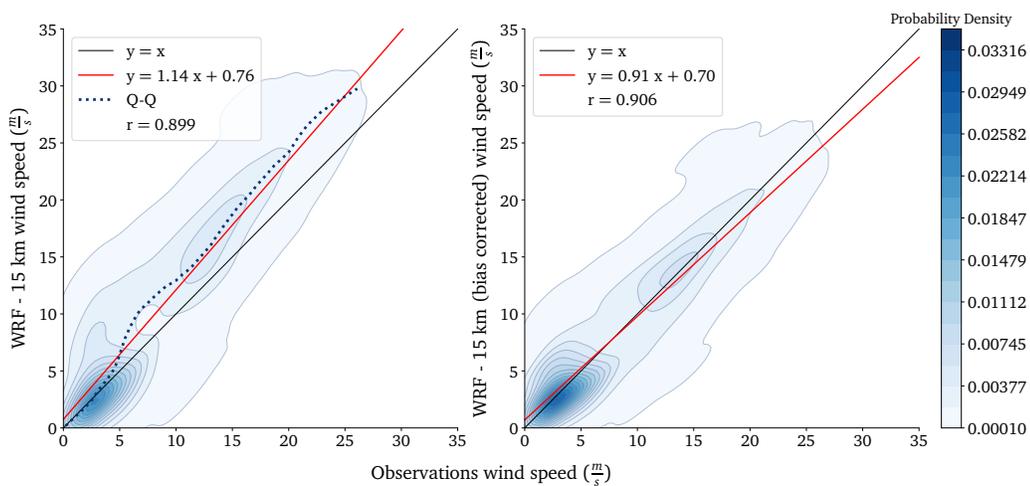
After the bias correction, the correlation coefficients are better, specially for grids coarser than 15 km. Nevertheless, the true effect of bias correction is shown in the linear least squares fit: for all cases the slope is close to 1 and the overall distribution of data (red line) is aligned towards the identity line (black). Remarkably, the 75 km WRF grid has the highest correlation coefficient ($r=0.912$); this performance is due to the fact that this grid is a 3-hourly dataset, so if we review the equation of the Pearson correlation coefficient: $r = \frac{cov_{xy}}{\sigma_x \sigma_y}$, the subsampling of a time series reduces the standard deviation in data and hence increases the correlation coefficient. For the hourly averaged time series, the 15, 3 and 1 km WRF grids have similar correlation coefficients and similar least squares fit as the 75 km grid. Among those, the best are the 1 and 3 km grids and the worst is MERRA-2 grid with correlation of 0.88.



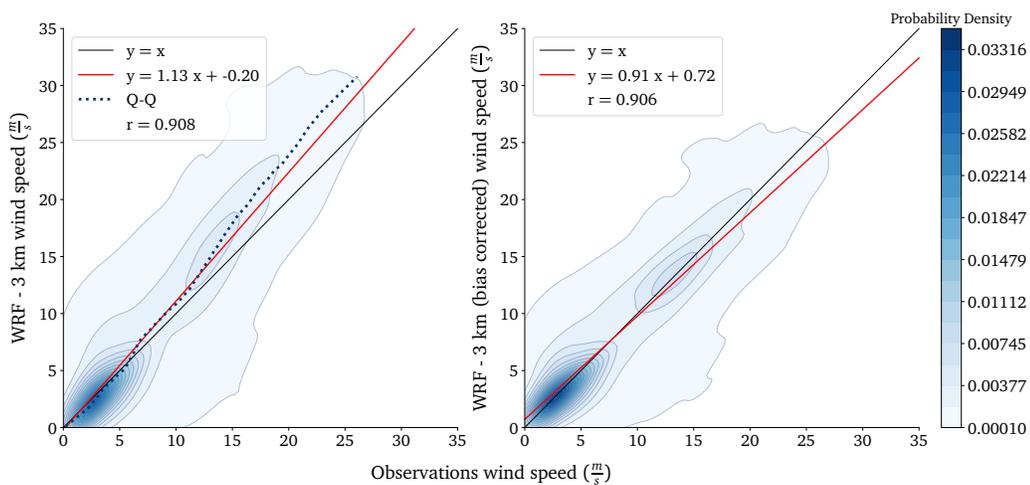
(a) WRF-75 km wind speed correlations



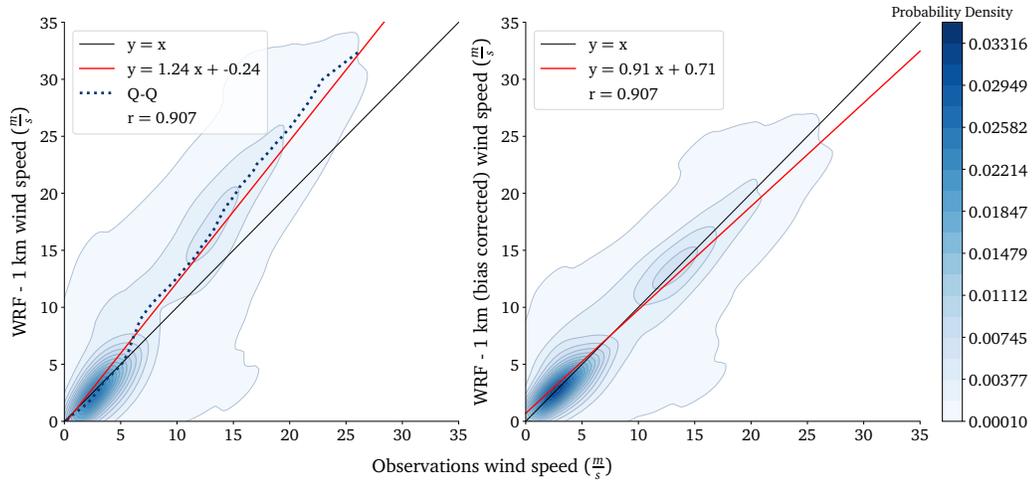
(b) MERRA-2 wind speed correlations



(c) WRF-15 km wind speed correlations



(d) WRF-3 km wind speed correlations



(e) WRF-1 km wind speed correlations

Figure 5.5: Comparisons of wind speed observations against non bias-corrected (left) and bias-corrected (right) wind speeds from models. (a) WRF model with 75 km grid, (b) MERRA-2 reanalysis with approx. 50 km grid, (c) WRF with 15 km grid, (d) WRF with 3 km grid and (e) WRF with 1 km grid. The black line indicates an identity relationship, the red line represents a linear least squares fit to the data and the dotted line shows the quantile-quantile distribution. "r" is the Pearson correlation coefficient.

5.2.2 Wind Power

In this section we will analyse the wind power production, calculated from bias-corrected wind speeds and using the power curve fit. Specifically, we look at the ability of models to reproduce the real power production. Therefore, the graphs in fig. 5.6 present a comparison between models and measurements.

Because of the good correlation of models in terms of wind speed, we could expect a strong correlation as well for wind power production. Instead of a diagonal shape in the plots, data is spread out throughout the whole graphs, and along with the spread, two major clusters of values appear: the first near null power production and another around the rated power. This arrangement of correlation is common for all the grid sizes.

This behaviour could mean a poor wind power modelling, but looking at the correlation coefficients, the lowest is 0.834, corresponding to MERRA-2, and the highest is 0.885, corresponding to the 75 km WRF grid. Similarly as for wind speed, the performance of the 75 km grid is due to the fact that it is 3-hourly subsampled. All the hourly averaged time series, the 15, 3 and 1 km WRF grids, have similar correlation coefficients ($r = 0.87$) and similar least squares fit, being the best the 1 and 3 km grids.

These correlations suggest that the data spread is not a significant source of error in the power production of models and that most of the data is located in the two aforementioned clusters.

Additionally, to test the effect of the power curve fit on estimated power produc-

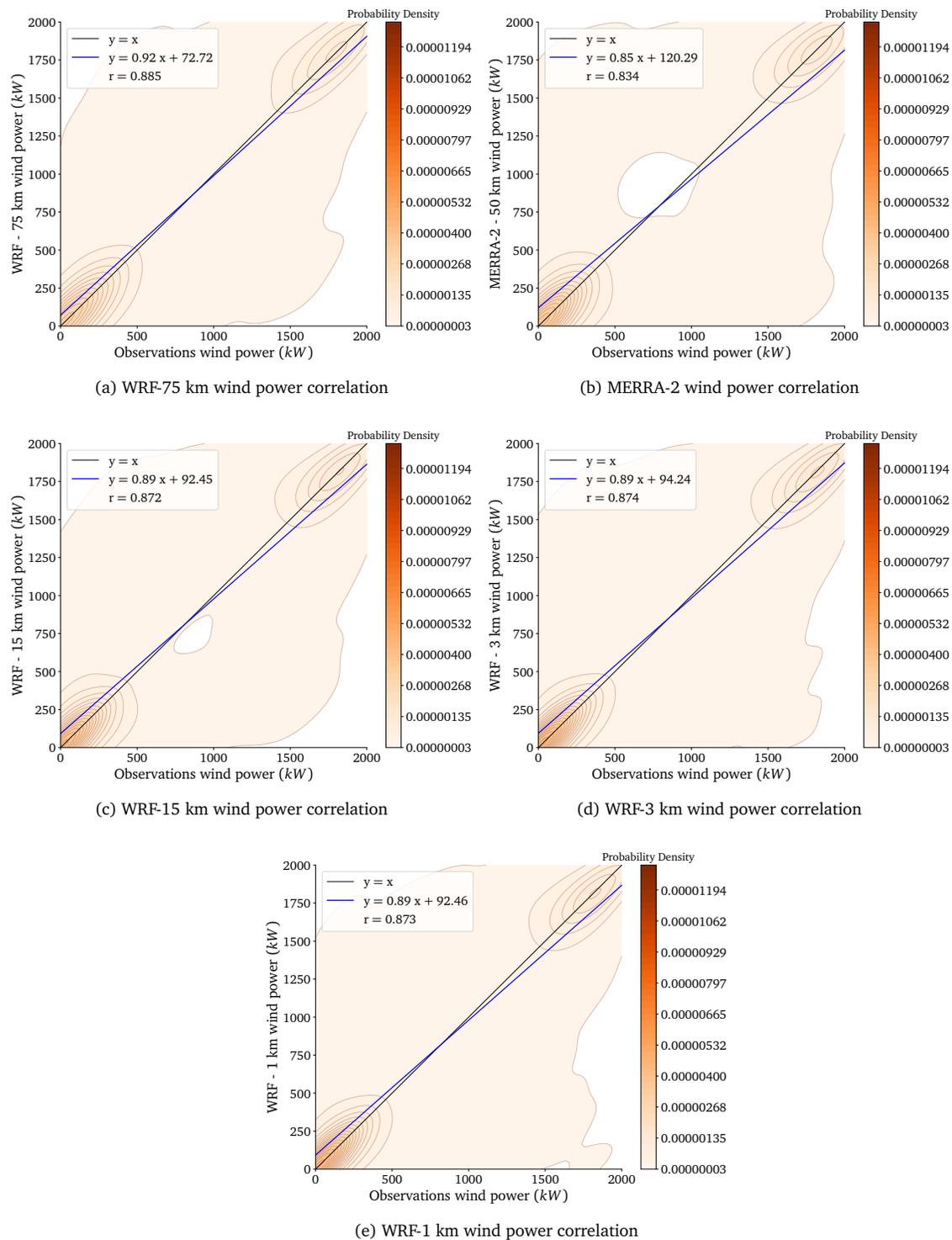


Figure 5.6: Comparisons of wind power observations against wind power derived from models. (a) WRF model with 75 km grid, (b) MERRA-2 reanalysis with approx. 50 km grid, (c) WRF with 15 km grid, (d) WRF with 3 km grid and (e) WRF with 1 km grid. The black line indicates an identity relationship, the blue line represents a linear least squares fit to the data and "r" is the Pearson correlation coefficient.

tion, we compared the on-site wind power production and wind power calculated using observed wind speed and the power curve fit (fig. 5.7). Both data sets show a high correlation, with $r=0.982$ and a slope of 1 for the least squares fit. The distribution of data follows the 1:1 diagonal and again presents two clusters in the same areas as in models' estimations. Also, there is a tendency towards the upper left, which is the result of low wind power production instead of rated power production at the end of the empirical power curve (values above 20 m/s in fig. 5.1).

The power curve fit is a good approximation to calculate wind power, due to the near diagonal arrangement in fig. 5.7 (black and blue lines respectively), which shows the imminent relation between estimated wind power with on-site measurements. Also, the idea of two states in power production is reaffirmed, due to the concentrations of power below 500 kW and around 1750 kW . In the end, a possible explanation for the spreading in the plots of figure 5.6 is the sensibility of the power production to changes in wind speed.

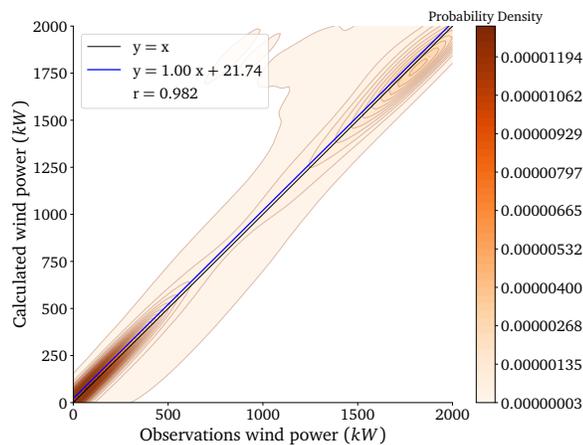


Figure 5.7: Comparison of observed wind power against calculated wind power derived from wind speed measurements. The black line indicates an identity relationship, the blue line represents a linear least squares fit to the data and "r" is the Pearson correlation coefficient.

5.3 Annual Capacity Factor

As mentioned in chapter 2, one of the main parameters to compare and measure the performance of wind turbines and wind farms is the Capacity Factor (CF). Using the equation 2.5, we calculated CF's for MERRA-2, WRF grids, power derived from observed wind speed and for on-site power production.

Figure 5.8 shows the contrast between models and observations: in all cases, CF is overestimated. The maximum difference is 1.8 percentage points, corresponding to MERRA-2, and the minimum is for the 15 km WRF grid with 1.5 percentage points. These values correspond to an error of 5.54% and 4.62%, respectively.

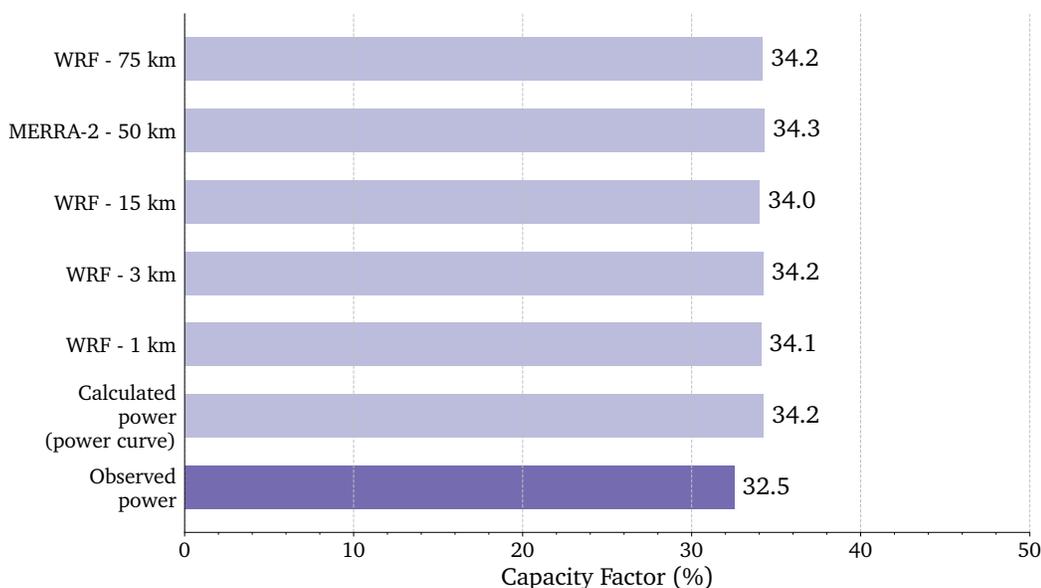


Figure 5.8: Annual capacity factor for observations, calculations and models

For future analysis, we could modify the power curve fit, to let the calculated wind power (using wind speed observations and the power curve) be equal to the wind power measurements. This could be to easily discard the inexactitude of wind power estimations from the power curve fit and have better approximations to on-site values.

5.4 Influence of spatiotemporal resolution on wind speed and wind power simulations

In this section we will focus on the second objective of this thesis, which is to determine the information available depending on the size of the grid of models. In the first part we will consider hourly datasets for wind speed and capacity factor and in the last part, we will deal with the same datasets, but subsampled at 3 hours, to identify the effects of subsampling on the information available (section 5.4.1).

Wind speed

The power spectral density of wind speed observations presents several peaks throughout all the frequencies, most of them above the mean red noise level (fig. 5.9). Following the methodology to discard red noise, for the 95% confidence level two zones stand out in the PSD, the first near the 7-day periodicity and the second for frequencies higher than 24 hours. Within the latter, three major peaks appear in the 24-h, 12-h and 8-h periodicities. Even though the 95% confidence level is used as a reference to discard red noise, we used a 99.9% confidence limit to focus our analysis on this three periodicities (green circles in fig. 5.9) and avoid possible false positives.

An interesting characteristic of the PSD of wind speed observations is the gap between the synoptic peak (near 7 days) and the diurnal peak (24 hours), resembling the gap found in the Van der Hoven's spectrum [74].

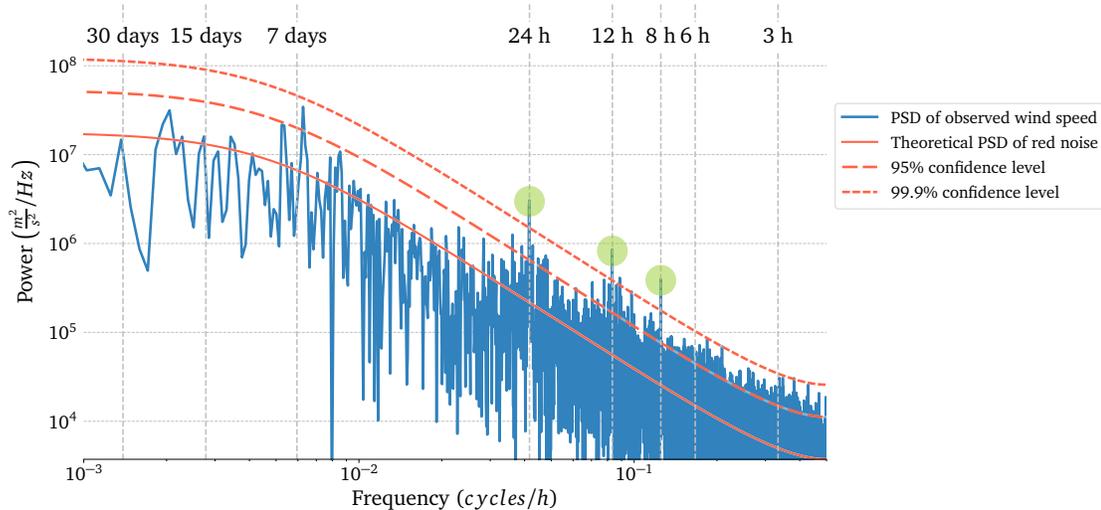


Figure 5.9: PSD of wind speed observations, hourly sampling

The models' PSDs reproduce the red noise background of observations, with the peculiarity of introducing a component at the 6-h periodicity. Also, all the PSDs, except for

MERRA-2, have the same gap as the PSD of observations, between the synoptic and the diurnal scale (fig. 5.10).

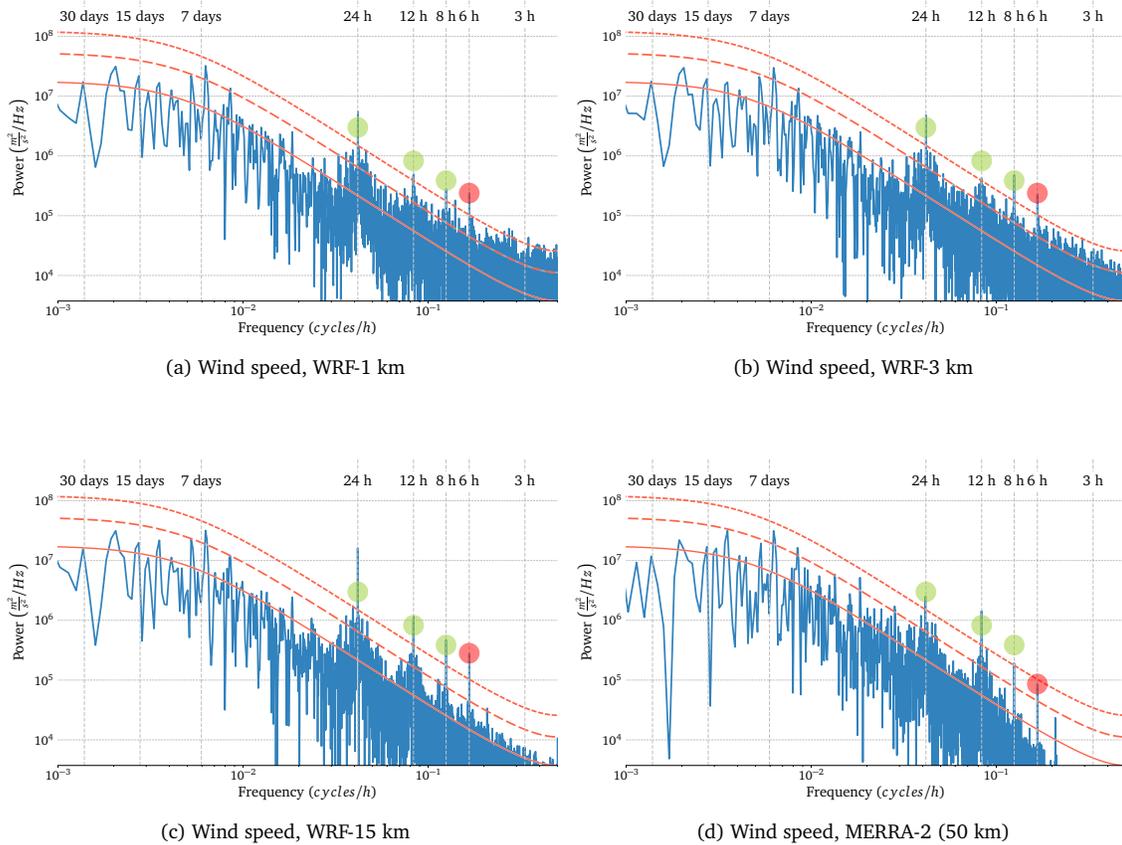


Figure 5.10: PSD of wind speed derived from models, hourly sampling. (a) 1-km WRF grid, (b) 3-km WRF grid, (c) 15-km WRF grid and (d) MERRA-2. The continuous orange line indicates the theoretical spectrum of red noise of the PSD of observations, the coarse dashed line indicates the 95% confidence level of red noise and the fine dashed line indicates a 99.9% confidence level of red noise. The green circles show the position of high frequency periodicities found in observations and the red circles show periodicities non existent in observations but that appear in models' PSD

Whereas the 1 km grid of WRF enhance the 24 hour component, it decreases the 12 and 8 hour component. Moreover, it is the only grid that presents false positives, for high frequencies, in the detection of red noise at the 99.9% confidence level. For the 3 km grid, although it has a close level of information for the entire PSD as in the one of observations, the absence of a clear peak at the 12 hour frequency is a remarkable characteristic for this grid size. Regarding the 15 km grid, the 24 hour periodicity is ten times higher than for observations; however, the 12-h and 8-h components are accurate. In comparison with the PSD of observations, the 15 km grid shows a general reduction of power for frequencies higher than 24 hours, the majority falling below the 95% confidence level. For MERRA-2 we found an inaccurate reproduction of the gap between the synoptic and diurnal scale, and a sharp decrease on the level of high frequencies, reaching in some

cases the mean red noise level. Nevertheless, the 24-h and 12-h frequencies remain above the 99.9% confidence level, and the 8-h frequency is the only periodicity under it.

The results of this section vary from grid to grid. It is clear that the variance of high frequencies changes when we look at the models: for high grid resolutions models have more spectral power in high frequencies than low grid resolutions. For low frequencies the variance in models does not differ greatly from observations.

Capacity factor

In terms of spectral density, the capacity factor derived from observed wind power has a similar red noise background as wind speed observations (fig. 5.11). For the analysis of CF, we assumed again a 95% and 99.9% confidence levels to detect periodicities based on the mean red noise level of the CF observations. Despite we are analysing CF, the PSD presents a similar gap that divide the synoptic and the diurnal scale. This is shown by the 95% confidence level, as there are a slight peak at 7 days and a group of frequencies at 24-h and higher. From this latter group, the frequencies that stand out from the 99.9% confidence level are the 24-h and 12-h peaks. Some small peaks also appear above this confidence level, but as they do not represent clear periodic phenomena, we did not take them into account for this analysis.

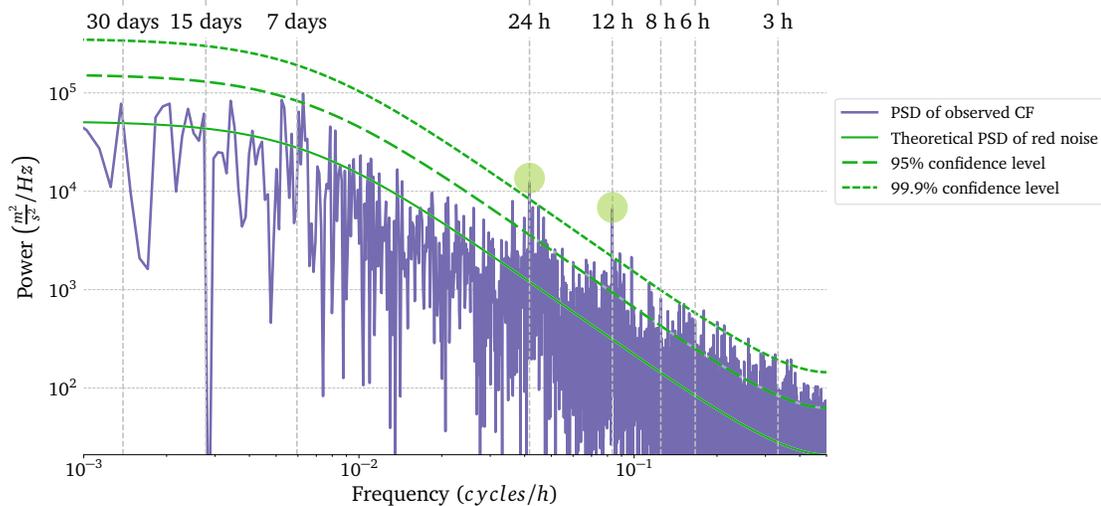


Figure 5.11: PSD of capacity factor observations, hourly sampling

In general, CF from models reproduce the red noise background of observations. The PSDs also have the same gap as of observations between the synoptic and the diurnal scale. In this case (only) some grids do introduce a component at the 6-h frequency (fig. 5.12).

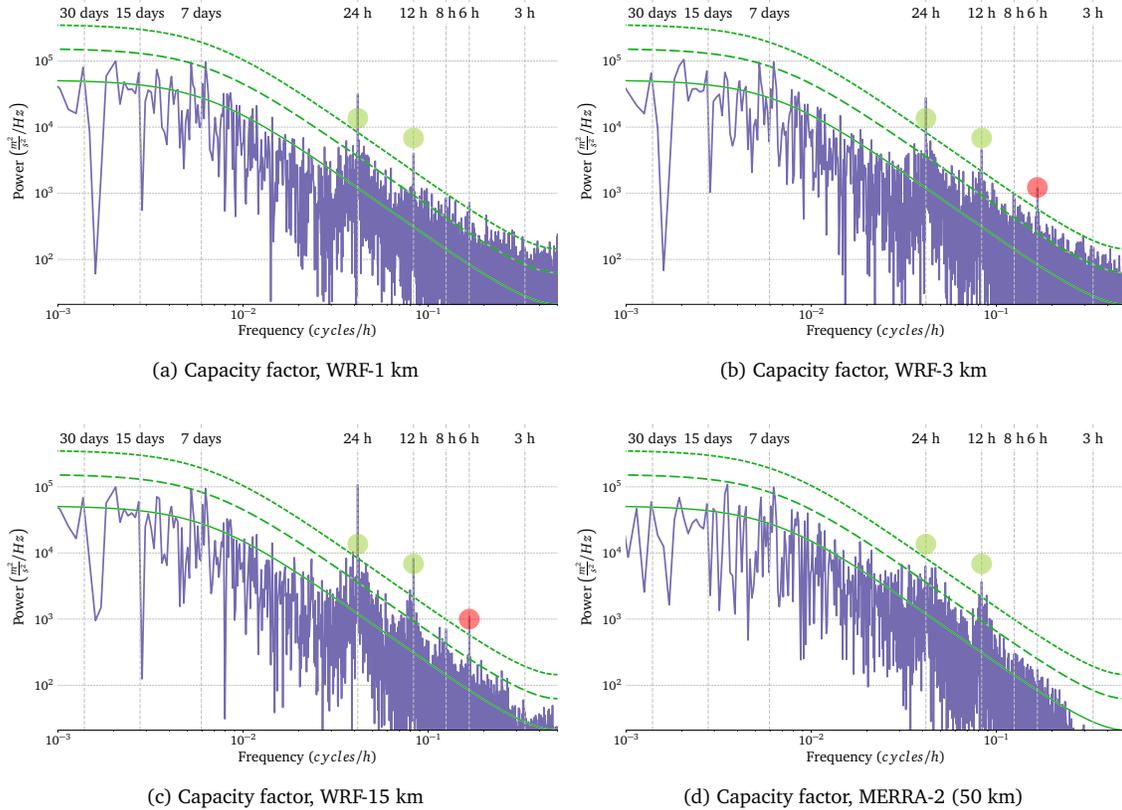


Figure 5.12: PSD of capacity factor derived from models, hourly sampling. (a) 1-km WRF grid, (b) 3-km WRF grid, (c) 15-km WRF grid and (d) MERRA-2. The continuous green line indicates the theoretical spectrum of red noise of the PSD of observations, the coarse dashed line indicates the 95% confidence level of red noise and the fine dashed line indicates a 99.9% confidence level of red noise. The green circles show the position of high frequency periodicities found in observations and the red circles show periodicities non existent in observations but that appear in models' PSD

For the finer grid of WRF, we obtained a similar behaviour as for the wind speed PSD: the 24 hour component is enhanced and the 8 hour component is decreased. The 1 km grid reproduces, for high frequencies at the 99.9% confidence level, the false positives present in the CF observations. The next grid in order (3 km grid), shows again the 24 hour component enhanced and the 8 hour component decreased. Additionally, this grid adds a component at 6 hours. As for the 15 km grid, the CF PSD has the same tendencies as its analogous for wind speed. The 24 hour periodicity is ten times higher than in observations, with an accurate reproduction of the 12-h component. Also a peak in the 6-h frequency appears. In comparison with the PSD of CF observations, the 15 km grid also shows a general reduction of power for frequencies higher than 24 hours, the majority falling below the 95% confidence level. Finally, MERRA-2 fails to reproduce the 24-h periodicity and gives a low amplitude to the 12-h frequency, Also, the gap between the synoptic and diurnal scale is not as clear as in observations. The decrease on the level of high frequencies is highly visible, reaching, for the highest, to the mean red noise

level.

Likewise the wind speed PSDs, the results of this section vary from grid to grid. However, the variance of high frequencies changes once again with the grid in use: the finer the grid resolution, the most spectral power in high frequencies. Furthermore, it seems that variance in low frequencies is not altered with the grid size.

5.4.1 Effect of subsampling on the frequency analysis

Wind speed

With 3-hourly subsampled wind speed observations we obtained a similar PSD as for hourly data, with almost identical levels of power. The 3-hourly observations also reproduce the gap found in the Van der Hoven's spectrum [74], and contain periodicities at 24-h, 12-h and 8-h. Similarly as with hourly observations, we took the 95% and 99.9% confidence level to detect periodicities.

One of the direct effects of subsampling in this section is the limit on the Nyquist frequency, which is the double than the frequency sampling. Figures 5.9 and 5.13 show this effect.

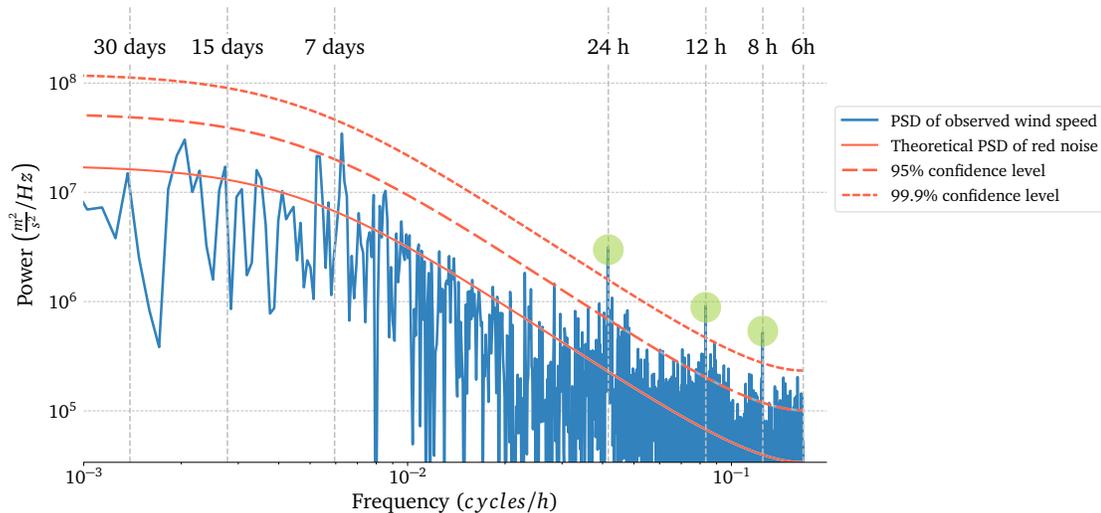


Figure 5.13: PSD of wind speed observations, 3-hourly sampling

In this case, the subsampled wind speeds from models have similar results than the output of models in section 5.4. However, there are some differences regarding 3-hourly observations and hourly models. They reproduce well the red noise background of observations, and also some grids introduce a component at the 6-h periodicity. Once again, all the PSDs, except for MERRA-2, have the same gap as the PSD of observations, between the synoptic and the diurnal scale. This time, we added a new grid to the analysis, the

75 km grid of WRF, which is the largest grid size available (fig. 5.14).

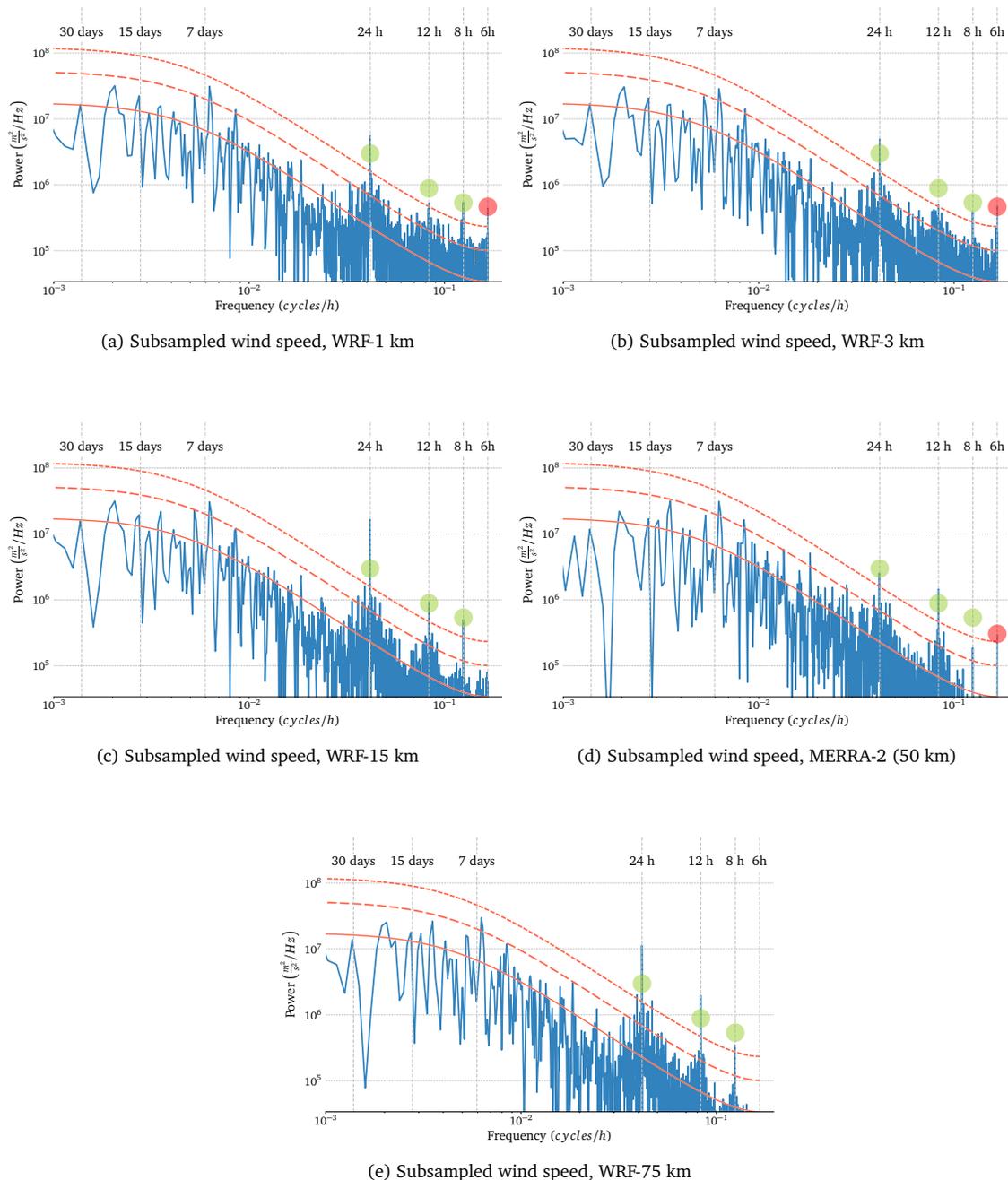


Figure 5.14: PSD of wind speed derived from models, 3-hourly sampling. (a) 1-km WRF grid, (b) 3-km WRF grid, (c) 15-km WRF grid, (d) MERRA-2 and (e) 75-km WRF grid. The continuous orange line indicates the theoretical spectrum of red noise of the PSD of observations, the coarse dashed line indicates the 95% confidence level of red noise and the fine dashed line indicates a 99.9% confidence level of red noise. The green circles show the position of high frequency periodicities found in observations and the red circles show periodicities non existent in observations but that appear in models' PSD

The 1 km grid of WRF enhance the 24 hour component, decreases the 12 hour component, is accurate for the 8 hour component and introduces a false periodicity at 6 hours.

Also, this grid does not give any false positives as the hourly data. For the 3 km grid, the 24 hour component is higher than for observations, and the 8 hour and 12 hour peaks diminish, the latter below the 99.9% confidence level. Again the 3 km grid introduces a 6 hour component. Regarding the 15 km grid, the 24 hour periodicity again is ten times higher than for observations, with the 12-h and 8-h components being accurate. As in previous wind speed and CF PSDs, the 15 km grid shows a general reduction of power for frequencies higher than 24 hours, falling close to the 95% confidence level. Nevertheless, this is one of the grid sizes that does not have an added 6-hour peak. For MERRA-2, we found a decrease on the level of highest frequencies, reaching in some cases below the mean red noise level. In spite of the decrease of power, only the 8-h frequency is under the 99.9% confidence level. Finally, the 75 km grid presents a growth in the 24-h and 12-h periodicities. The magnitude for high frequencies is marginally better than in the MERRA-2 PSD. Additionally the 8-h frequency is above the 99.9% confidence level and the extra frequency at 6-h does not appear.

The results of the models in this section vary from the results of hourly data sets. Even if there are some differences, we obtained similar PSDs and above all, we also observed how the variance of high frequencies diminish when the spatial resolution of the grid starts to decrease.

Capacity factor

The 3-hourly subsampled CF observations produce an almost identical PSD as the hourly CF data series. They also reproduce the gap between the synoptic and the diurnal scale found in the Van der Hoven's spectrum [74], and contain periodicities at 24-h, 12-h (fig. 5.15). The only difference is the lack of false positives, as in this case they would appear beyond the Nyquist frequency. Similarly as with hourly observations, we took the 95% and 99.9% confidence level to detect periodicities and compare it with models' PSDs.

The results of the PSDs of subsampled output of models vary from grid to grid. They follow the overall distribution of power as for hourly models: they reproduce well the red noise background of observations, and once again, all the PSDs, except for MERRA-2, have the same gap as the PSD of observations, between the synoptic and the diurnal scale. The differences between grids are the amplitude of the periodicities in high frequencies and the addition of extra components depending on the grid. In this analysis we added the 75 km grid of WRF as well (fig. 5.16).

Regarding the 1 km grid of WRF, we found the 24 hour component enhanced and an accurate reproduction of the 12 hour component. Unlike observations a peak appear at 6-h. For the 3 km grid, we found the periodicities of 24-h and 12-h equal as for the 1 km grid, we also detected an extra frequency at 6-h and a small decrement in the red noise

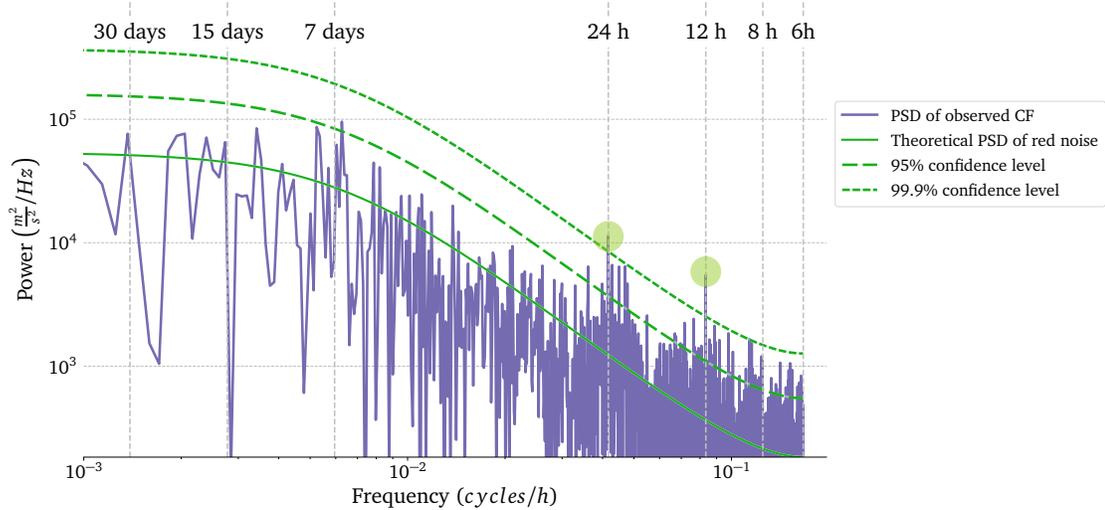


Figure 5.15: PSD of capacity factor observations, 3-hourly sampling

of high frequencies. In the 15 km grid, the 24 hour periodicity continue to be ten times higher than for observations; with the 12-h component being accurate. Congruently with the hourly dataset, this grid shows a general reduction of power for frequencies higher than 24 hours, falling close to the 95% confidence level. For MERRA-2 we found a decrease on the level of highest frequencies, reaching in some cases below the mean red noise level. This time, MERRA-2 do not reproduce the 24-h frequency, with the 12-h frequency hardly above the 99.9% confidence level. Finally, the 75 km grid presents a considerable growth in the 24-h peak, whereas the 12-h periodicity is well reproduced. The decrease in magnitude in the highest frequencies is again comparable to the MERRA-2 PSD, but with an extra periodicity at 8-h frequency above the 99.9% confidence level.

Likewise the hourly CF PSDs, the results of this section show how the variance of high frequencies changes with the grid in use: the variance of high frequencies diminish when the spatial resolution of the grid starts to decrease.

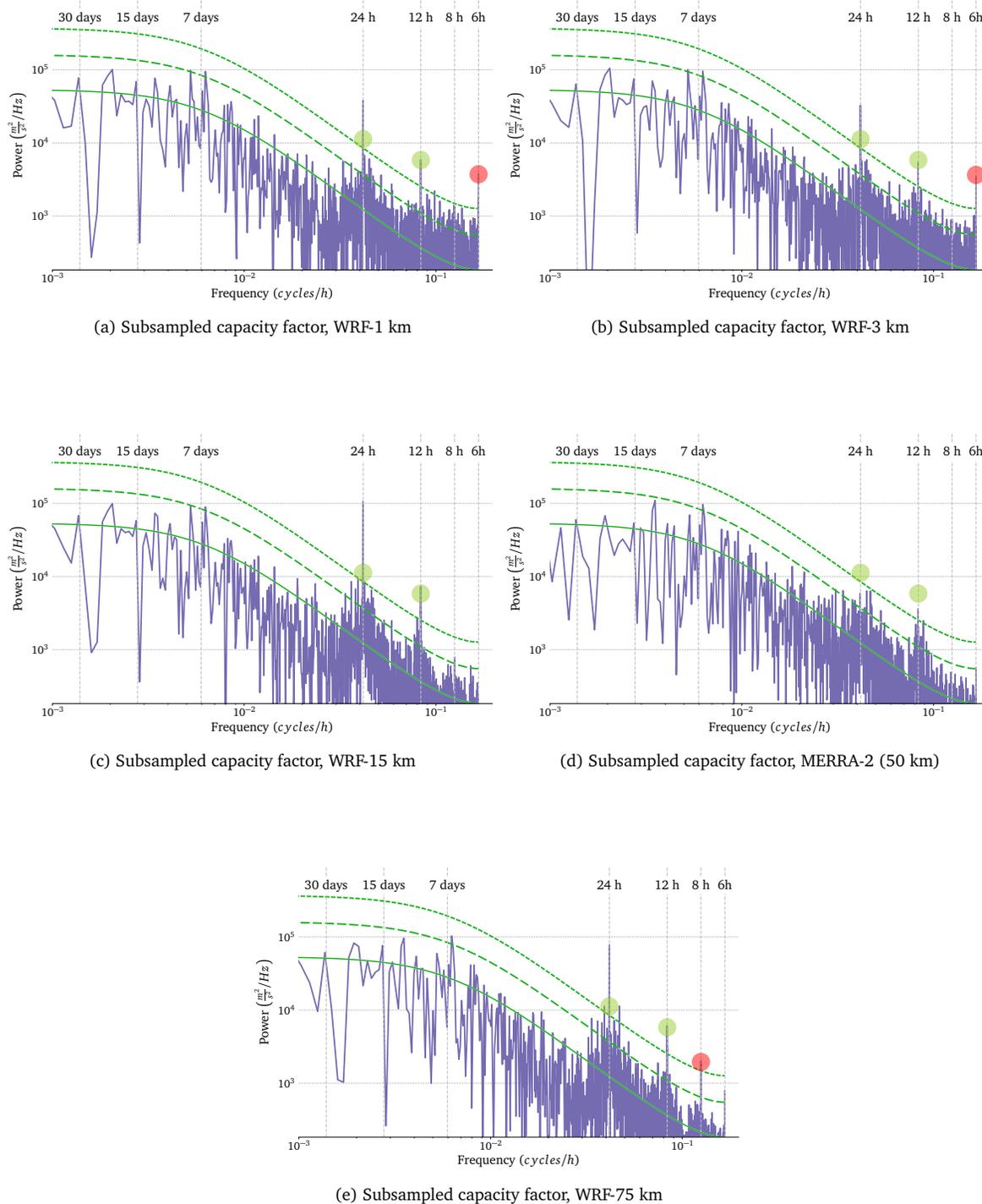


Figure 5.16: PSD of capacity factor derived from models, 3-hourly sampling. (a) 1-km WRF grid, (b) 3-km WRF grid, (c) 15-km WRF grid, (d) MERRA-2 and (e) 75-km WRF grid. The continuous green line indicates the theoretical spectrum of red noise of the PSD of observations, the coarse dashed line indicates the 95% confidence level of red noise and the fine dashed line indicates a 99.9% confidence level of red noise. The green circles show the position of high frequency periodicities found in observations and the red circles show periodicities non existent in observations but that appear in models' PSD

Conclusions

This thesis analyse the suitability of NWP models to represent the wind speed and derived wind power of a wind turbine in Chiapas, Mexico. We have used MERRA-2 reanalysis and WRF mesoscale model with 1, 3, 15, 75 km grids and hourly and 3-hourly data, to investigate the influence of the spatiotemporal resolution on the reproduction of wind speed and wind power production; a frequency spectrum analysis of wind speed and CF was carried out to determine the contributions of different grid sizes to the variance of time series.

In the process to estimate wind power production from wind speed data, the power curve is a key factor as it should represent correctly the performance of a wind turbine. To fit an error function as a power curve showed a good approximation for the wind turbine analysed. We were able to reproduce the annual CF with a maximum error of 5.54%. Also, based on fig. 5.7, the reproduction of wind power observations using the power curve fitting shows a high correlation (0.98), despite the inability to reproduce some of the data above 20 *m/s*.

The wind speed for 2016 in Arriaga, Mexico comes mainly from the north-northeast and has a bimodal distribution. Models' outputs are able to reproduce this bimodal pattern, with several differences depending on the grid size. For raw outputs of WRF model, the 1 and 3 km grids have the most accurate reproduction and as the grid size increases, the correlation decreases. After bias correction all the grids of WRF have very close correlations (around 0.91). Even though this comparison can be applied only to the hourly time series, we can conclude that the 3-hourly 75 km grid, has a good correlation and could be useful when high temporal resolution data sets are not needed. Additionally, the results show that in all cases, any grid of WRF performs better than MERRA-2.

Beyond the correspondence of models with measurements, the models are able to reproduce the two states of wind power production: the first near null power production and another around the rated power. Comparing all the grids, MERRA-2 again has the

poorest performance ($r = 0.83$) and all the grids of WRF perform almost equally ($r = 0.87$), excepting the 75 km grid, which is 3-hourly sampled and thus non comparable.

Furthermore, we found two relevant trends in the power spectral density of wind speed and CF. The first is the fact that all grids of WRF contain higher variance at the 24-h cycle than observed values (the overestimation may be due to the restarting of WRF simulations every 24 hours). The other relevant characteristic is that the majority of the spectra of WRF grids, for wind speed and CF, introduce an artificial periodicity at 6-h (this extra component may be due to the retrieving of the NCEP-FNL analysis every six hours). In both cases more research is needed to analyse in detail these findings and get solid conclusions.

To sum up, even if there are disparities between models and observations in the amplitude of the 24-h, 12-h and 8-h peaks, for wind speed and CF power spectra, we can reliably relate the decrease of information to the decrease of spatial resolution. This is, the smaller the grid spacing, the more energy in higher frequencies. Hence, for synoptic or longer cycles, low spatial resolution grids (75 km and MERRA-2) can reproduce the same amount of information as in observations, but for smaller cycles (diurnal or less), high resolution grids (1 and 3 km optimally) are the only ones that can reproduce the same amount of information as observations.

The results of this work suggest that NWP models are a tool that provide a reliable degree of certainty in the reproduction of wind dynamics. Even if they present deviations from reality, with a statistical correction they can increase their performance significantly and be a solid complement for traditional methods of WRA. Thus, we conclude that wind speeds derived from NWP models are useful reproducing the wind power production in Mexico.

To know the effect of the numerical grid on wind speed and CF is useful to understand the distribution of variance in different frequencies, from diurnal to monthly cycles, and to help in the decision of choosing the size of the numerical grid for future studies in southern Mexico. As is to be expected, the process to capture the physical models in the Mexican context, through the resolution of the grid, land use, parametrisations, etc. is still a work in progress. Part of this work is examined in the recommendations and future work.

6.1 Recommendations and future work

Based on the above conclusions, we can provide the following future work and recommendations:

- Analyse seasonal frequencies with time series longer than a year.
- Continue quantifying the effect of subsampling on spectral power (3, 6, 12, 24 h).
- Apply the same frequency analysis to technoeconomic indices relevant to wind power.
- Compare against new reanalyses and see how they compete with mesoscale models.
- Assess the suitability of models for seasonal and subseasonal CF.
- Continue to adapt the configurations of WRF (PBL parametrisations, initial conditions, ensembles) for the Mexican context.
- Know if there is a real influence of the 6 hourly data of boundary and initial conditions into the 6-h component present in some PSDs of the frequency analysis.
- Assess the impact of the daily running for WRF in the 24-h component of PSDs (possibly by changing the periodicity of WRF restarting).
- Improve the understanding of the effects of physical phenomena in southern Mexico and to explain the overestimation or underestimation of wind speed in models.
- Try different methods to represent accurately frequency spectra (filters, windows, multitapering).

References

- [1] Zachary A. Wendling. “Bridges beyond renewable energy: Decarbonizing the global electricity sector under uncertainty”. In: *Energy Research and Social Science* 48 (2019), pp. 235–245. ISSN: 22146296. DOI: 10.1016/j.erss.2018.09.020.
- [2] Kenneth Hansen, Brian Vad Mathiesen, and Iva Ridjan Skov. “Full energy system transition towards 100% renewable energy in Germany in 2050”. In: *Renewable and Sustainable Energy Reviews* 102 (2019), pp. 1–13. ISSN: 13640321. DOI: 10.1016/j.rser.2018.11.038.
- [3] H. Lund and B. V. Mathiesen. “Energy system analysis of 100% renewable energy systems-The case of Denmark in years 2030 and 2050”. In: *Energy* 34.5 (2009), pp. 524–531. ISSN: 03605442. DOI: 10.1016/j.energy.2008.04.003.
- [4] International Renewable Energy Agency. *Renewable Energy Prospects: Mexico, REmap 2030 analysis*. Tech. rep. Abu Dhabi: IRENA, 2014.
- [5] Ministry of Energy (Mexico). *Mapa de Ruta Tecnológica Energía Eólica en Tierra*. Tech. rep. Mexico City, 2017, p. 70.
- [6] Diario Oficial de la Federación de México. *LEY DE TRANSICIÓN ENERGÉTICA*. 2015.
- [7] IEA. *Renewables Information 2019: Overview*. Tech. rep. 2019, p. 12.
- [8] Ministry of Energy (Mexico). *Reporte de Avance de Energías Limpias Primer Semestre 2018*. Tech. rep. Mexico City, 2018, p. 21.
- [9] GWEC. *Global Wind Report 2017*. Tech. rep. 2018.
- [10] Mexican Wind Energy Association. *Mapas eólicos*. 2019. (Visited on 03/14/2019).
- [11] IRENA. *Renewable capacity statistics 2019*. Tech. rep. Abu Dhabi: International Renewable Energy Agency (IRENA), 2019.
- [12] BP. *BP Statistical Review of World Energy 2019*. Tech. rep. London: British Petroleum Co., 2019, p. 64.

- [13] Ezequiel Zárate-Toledo, Rodrigo Patiño, and Julia Fraga. “Justice, social exclusion and indigenous opposition: A case study of wind energy development on the Isthmus of Tehuantepec, Mexico”. In: *Energy Research and Social Science* (2019). ISSN: 22146296. DOI: 10.1016/j.erss.2019.03.004.
- [14] Eduardo Pérez-Denicia et al. *Renewable energy sources for electricity generation in Mexico: A review*. 2017. DOI: 10.1016/j.rser.2017.05.009.
- [15] IEA. *Wind Power Technology Roadmap*. Tech. rep. Paris: OECD/IEA, 2013, p. 63.
- [16] K. S.R. R Murthy and O. P. Rahi. “A comprehensive review of wind resource assessment”. In: *Renewable and Sustainable Energy Reviews* 72:July 2016 (2017), pp. 1320–1342. ISSN: 18790690. DOI: 10.1016/j.rser.2016.10.038.
- [17] Sultan Al-Yahyai, Yassine Charabi, and Adel Gastli. “Review of the use of Numerical Weather Prediction (NWP) Models for wind energy assessment”. In: *Renewable and Sustainable Energy Reviews* 14.9 (2010), pp. 3192–3198. ISSN: 1364-0321. DOI: 10.1016/J.RSER.2010.07.001.
- [18] Yuankun Wang et al. “A new method for wind speed forecasting based on copula theory”. In: *Environmental Research* 160 (2018), pp. 365–371. ISSN: 10960953. DOI: 10.1016/j.envres.2017.09.034.
- [19] R.D. Prasad, R.C. Bansal, and M. Sauturaga. “Some of the design and methodology considerations in wind resource assessment”. In: *IET Renewable Power Generation* 3.1 (2009), p. 53. ISSN: 17521416. DOI: 10.1049/iet-rpg:20080030.
- [20] P. S. Jackson and J. C.R. Hunt. “Turbulent wind flow over a low hill”. In: *Quarterly Journal of the Royal Meteorological Society* (1975). ISSN: 1477870X. DOI: 10.1002/qj.49710143015.
- [21] Philippe Beaucage, Michael C. Brower, and Jeremy Tensen. “Evaluation of four numerical wind flow models for wind resource mapping”. In: *Wind Energy* (2014). ISSN: 10954244. DOI: 10.1002/we.1568.
- [22] Chong Wei Zheng et al. “An overview of global ocean wind energy resource evaluations”. In: *Renewable and Sustainable Energy Reviews* 53:667 (2016), pp. 1240–1251. ISSN: 18790690. DOI: 10.1016/j.rser.2015.09.063.
- [23] Stephen Rose and Jay Apt. “Quantifying sources of uncertainty in reanalysis derived wind speed”. In: *Renewable Energy* 94 (2016), pp. 157–165. ISSN: 18790682. DOI: 10.1016/j.renene.2016.03.028.
- [24] Iain Staffell and Stefan Pfenninger. “Using bias-corrected reanalysis to simulate current and future wind power output”. In: *Energy* (2016). ISSN: 03605442. DOI: 10.1016/j.energy.2016.08.068.

- [25] D. J. Cannon et al. “Using reanalysis data to quantify extreme wind power generation statistics: A 33 year case study in Great Britain”. In: *Renewable Energy* 75 (2015), pp. 767–778. ISSN: 18790682. DOI: 10.1016/j.renene.2014.10.024.
- [26] Cristian Mattar and María Cristina Guzmán-Ibarra. “A techno-economic assessment of offshore wind energy in Chile”. In: *Energy* 133 (2017), pp. 191–205. ISSN: 03605442. DOI: 10.1016/j.energy.2017.05.099.
- [27] Global Modeling and Assimilation Office (GMAO). *MERRA-2 inst1_2d_asm_Nx: 2d,1-Hourly,Instantaneous,Single-Level,Assimilation,Single-Level Diagnostics*. Greenbelt, MD, USA, 2015. DOI: 10.5067/3Z173KIE2TPD.
- [28] Guorui Ren et al. “Characterization of wind resource in China from a new perspective”. In: *Energy* 167 (2019), pp. 994–1010. ISSN: 03605442. DOI: 10.1016/j.energy.2018.11.032.
- [29] Pedro A. Jiménez et al. “Role of synoptic- and meso-scales on the evolution of the boundary-layer wind profile over a coastal region: the near-coast diurnal acceleration”. In: *Meteorology and Atmospheric Physics* 128.1 (2016), pp. 39–56. ISSN: 14365065. DOI: 10.1007/s00703-015-0400-6.
- [30] Renalda El-Samra, Elie Bou-Zeid, and Mutasem El-Fadel. “What model resolution is required in climatological downscaling over complex terrain?” In: *Atmospheric Research* 203 (2018), pp. 68–82. ISSN: 01698095. DOI: 10.1016/j.atmosres.2017.11.030.
- [31] Lattawan Niyomtham et al. “On the Wind Resource Assessment along the Western Coast of Thailand”. In: *Energy Procedia* 138 (2017), pp. 1190–1195. ISSN: 18766102. DOI: 10.1016/j.egypro.2017.10.387.
- [32] D. Carvalho et al. “WRF wind simulation and wind energy production estimates forced by different reanalyses: Comparison with observed data for Portugal”. In: *Applied Energy* 117 (2014), pp. 116–126. ISSN: 03062619. DOI: 10.1016/j.apenergy.2013.12.001.
- [33] David J. Stensrud. *Parameterization Schemes*. Cambridge: Cambridge University Press, 2013. ISBN: 9780511812590. DOI: 10.1017/cbo9780511812590.
- [34] Caroline Draxl et al. “Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes”. In: *Wind Energy* 17.1 (2014), pp. 39–55. ISSN: 10954244. DOI: 10.1002/we.1555.

- [35] Sergio Fernández-González et al. “Sensitivity analysis of WRF model: wind-resource assessment for complex terrain”. In: *Journal of Applied Meteorology and Climatology* (2017), JAMC-D-17-0121.1. ISSN: 1558-8424. DOI: 10.1175/JAMC-D-17-0121.1.
- [36] David Siuta, Gregory West, and Roland Stull. “WRF hub-height wind forecast sensitivity to PBL scheme, grid length, and initial condition choice in complex terrain”. In: *Weather and Forecasting* 32.2 (2017), pp. 493–509. ISSN: 15200434. DOI: 10.1175/WAF-D-16-0120.1.
- [37] Dave MacLeod et al. “Transforming climate model output to forecasts of wind power production: how much resolution is enough?” In: *Meteorological Applications* (2018). ISSN: 14698080. DOI: 10.1002/met.1660.
- [38] David Carvalho et al. “A sensitivity study of the WRF model in wind simulation for an area of high wind energy”. In: *Environmental Modelling and Software* 33 (2012), pp. 23–34. ISSN: 13648152. DOI: 10.1016/j.envsoft.2012.01.019.
- [39] Elizabeth N. Smith et al. “WRF model study of the Great Plains low-level jet: Effects of grid spacing and boundary layer parameterization”. In: *Journal of Applied Meteorology and Climatology* 57.10 (2018), pp. 2375–2397. ISSN: 15588432. DOI: 10.1175/JAMC-D-17-0361.1.
- [40] Renyu Yuan et al. “Coupled wind farm parameterization with a mesoscale model for simulations of an onshore wind farm”. In: *Applied Energy* 206.April (2017), pp. 113–125. ISSN: 03062619. DOI: 10.1016/j.apenergy.2017.08.018.
- [41] Cristian Mattar and Dager Borvarán. “Offshore wind power simulation by using WRF in the central coast of Chile”. In: *Renewable Energy* 94 (2016), pp. 22–31. ISSN: 18790682. DOI: 10.1016/j.renene.2016.03.005.
- [42] Forest Cannon et al. “WRF simulation of downslope wind events in coastal Santa Barbara County”. In: *Atmospheric Research* 191 (2017), pp. 57–73. ISSN: 01698095. DOI: 10.1016/j.atmosres.2017.03.010.
- [43] Seán Collins et al. “Impacts of Inter-annual Wind and Solar Variations on the European Power System”. In: *Joule* (2018). ISSN: 25424351. DOI: 10.1016/j.joule.2018.06.020.
- [44] Pieter de Jong et al. “Integrating large scale wind power into the electricity grid in the Northeast of Brazil”. In: *Energy* 100 (2016), pp. 401–415. ISSN: 03605442. DOI: 10.1016/j.energy.2015.12.026.

- [45] Kelly Eurek et al. “An improved global wind resource estimate for integrated assessment models”. In: *Energy Economics* 64 (2017), pp. 552–567. ISSN: 01409883. DOI: 10.1016/j.eneco.2016.11.015.
- [46] Pieter de Jong et al. “Forecasting high proportions of wind energy supplying the Brazilian Northeast electricity grid”. In: *Applied Energy* 195 (2017), pp. 538–555. ISSN: 03062619. DOI: 10.1016/j.apenergy.2017.03.058.
- [47] George Caralis et al. “Development of Aeolian map of China using mesoscale atmospheric modelling”. In: *Renewable Energy* 74 (2015), pp. 60–69. ISSN: 09601481. DOI: 10.1016/j.renene.2014.07.055.
- [48] Theodore M. Giannaros, Dimitrios Melas, and Ioannis Ziomas. “Performance evaluation of the Weather Research and Forecasting (WRF) model for assessing wind resource in Greece”. In: *Renewable Energy* 102 (2017), pp. 190–198. ISSN: 18790682. DOI: 10.1016/j.renene.2016.10.033.
- [49] W.C. Skamarock et al. “A Description of the Advanced Research WRF Model Version 4 NCAR Technical Note”. In: *National Center for Atmospheric Research* (2019), p. 145. ISSN: 1477870X. DOI: 10.5065/1dfh-6p97.
- [50] Jordan G. Powers et al. “The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions”. In: *Bulletin of the American Meteorological Society* 98.8 (2017), pp. 1717–1737. ISSN: 0003-0007. DOI: 10.1175/BAMS-D-15-00308.1.
- [51] Quetzalcoatl Hernández-Escobedo, Alberto-Jesús Perea-Moreno, and Francisco Manzano-Agugliaro. “Wind energy research in Mexico”. In: *Renewable Energy* 123 (2018), pp. 719–729. ISSN: 0960-1481. DOI: 10.1016/J.RENENE.2018.02.101.
- [52] Jing Liu, Htet Lin, and Jun Zhang. “Review on the technical perspectives and commercial viability of vertical axis wind turbines”. In: *Ocean Engineering* 182.May (2019), pp. 608–626. ISSN: 00298018. DOI: 10.1016/j.oceaneng.2019.04.086.
- [53] J. F. Manwell, J. G. McGowan, and A. L. Rogers. *Wind Energy Explained: Theory, Design and Application*. 2010. ISBN: 9780470015001. DOI: 10.1002/9781119994367.
- [54] Mark Decker et al. “Evaluation of the Reanalysis Products from GSFC, NCEP, and ECMWF Using Flux Tower Observations”. In: *Journal of Climate* 25.6 (2012), pp. 1916–1944. ISSN: 0894-8755. DOI: 10.1175/JCLI-D-11-00004.1.
- [55] A Stickler and S Brönnimann. “Significant bias of the NCEP/NCAR and twentieth-century reanalyses relative to pilot balloon observations over the West African Monsoon region (1940-1957)”. In: *Quarterly Journal of the Royal Meteorological Society* 137.659 (2011), pp. 1400–1416. ISSN: 00359009. DOI: 10.1002/qj.854.

- [56] Sathyajith Mathew. *Wind energy: Fundamentals, resource analysis and economics*. 2007. ISBN: 3540309055. DOI: 10.1007/3-540-30906-3.
- [57] Michael E Mann and Jonathan M Lees. “Robust estimation of background noise and signal detection in climatic time series”. In: *Climatic Change* 33.3 (1996), pp. 409–445. ISSN: 1573-1480. DOI: 10.1007/BF00142586.
- [58] D. L. Gilman, F. J. Fuglister, and J. M. Mitchell. “On the Power Spectrum of “Red Noise””. In: *Journal of the Atmospheric Sciences* 20.2 (1963), pp. 182–184. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1963)020<0182:OTPSON>2.0.CO;2.
- [59] Torrence C. and Compo G. “A practical guide to wavelet analysis.” In: *Bulletin of the American Meteorological Society* 79 (1) (1998), pp. 61–78.
- [60] Michael Schulz and Manfred Mudelsee. “REDFIT: Estimating red-noise spectra directly from unevenly spaced paleoclimatic time series”. In: *Computers and Geosciences* 28.3 (2002), pp. 421–426. ISSN: 00983004. DOI: 10.1016/S0098-3004(01)00044-9.
- [61] Instituto para el Federalismo y el Desarrollo Municipal and Secretaría de Gobernación. *Enciclopedia de los municipios y delegaciones de México*. 2010. URL: <http://siglo.inafed.gob.mx/enciclopedia/EMM07chiapas/municipios/07009a.html>.
- [62] Ronald Gelaro et al. “The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)”. In: *Journal of Climate* 30.14 (2017), pp. 5419–5454. ISSN: 0894-8755. DOI: 10.1175/JCLI-D-16-0758.1.
- [63] *NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999*. 2000. DOI: 10.5065/D6M043C6. (Visited on 08/22/2018).
- [64] Jimmy Dudhia. “Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model”. In: *Journal of the Atmospheric Sciences* 46.20 (1989), pp. 3077–3107. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.
- [65] Eli J Mlawer et al. “Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave”. In: *Journal of Geophysical Research: Atmospheres* 102.D14 (1997), pp. 16663–16682. DOI: 10.1029/97JD00237.
- [66] Song-You Hong, Jimmy Dudhia, and Shu-Hua Chen. “A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation”. In: *Monthly Weather Review* 132.1 (2004), pp. 103–120. ISSN: 0027-0644. DOI: 10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

- [67] John S. Kain. “The Kain–Fritsch Convective Parameterization: An Update”. In: *Journal of Applied Meteorology* 43.1 (2004), pp. 170–181. ISSN: 0894-8763. DOI: 10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.
- [68] Song-You Hong, Yign Noh, and Jimy Dudhia. “A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes”. In: *Monthly Weather Review* 134.9 (2006), pp. 2318–2341. ISSN: 0027-0644. DOI: 10.1175/MWR3199.1.
- [69] Pedro A Jiménez et al. “A Revised Scheme for the WRF Surface Layer Formulation”. In: *Monthly Weather Review* 140.3 (2012), pp. 898–918. DOI: 10.1175/MWR-D-11-00056.1.
- [70] M. Tewari et al. “Implementation and verification of the unified noah land surface model in the WRF model”. In: *Bulletin of the American Meteorological Society*. 2004.
- [71] J.R. Garratt. *The Atmospheric Boundary Layer*. Cambridge: Cambridge University Press, 1994. ISBN: 0-521-46745-4.
- [72] D Lamb. “CLOUDS AND FOG | Cloud Microphysics”. In: *Encyclopedia of Atmospheric Sciences*. Ed. by Gerald R North, John Pyle, and Fuqing Zhang. Second Edi. Oxford: Elsevier, 2015, pp. 133–140. ISBN: 978-0-12-382225-3. DOI: 10.1016/B978-0-12-382225-3.00111-0.
- [73] Dirk Cannon et al. “Determining the bounds of skilful forecast range for probabilistic prediction of system-wide wind power generation”. In: *Meteorologische Zeitschrift* 26.3 (2017), pp. 239–252. ISSN: 16101227. DOI: 10.1127/metz/2016/0751.
- [74] Isaac Van der Hoven. “POWER SPECTRUM OF HORIZONTAL WIND SPEED IN THE FREQUENCY RANGE FROM 0.0007 TO 900 CYCLES PER HOUR”. In: *Journal of Meteorology* (1957). ISSN: 0095-9634. DOI: 10.1175/1520-0469(1957)014<0160:psohws>2.0.co;2.