Master Thesis

Assessing the impact of El Niño and La Niña on Brazilian wind-power generation with reanalysis based simulated time series

Submitted by

Dipl.-Ing. Katharina Gruber

1026526

In partial fulfilment of the requirements for the degree of

“Diplom-Ingenieur”

Within the Master’s Programme

“Stoffliche und energetische Nutzung nachwachsender Rohstoffe (NAWARO)”

University of Natural Resources and Life Sciences, Vienna

Department of Economics and Social Sciences

Institute for Sustainable Economic Development

Supervised by: Ass.Prof. Dipl.-Ing. Dr. Johannes Schmidt

Co-Advisor: Prof. Dr. Carlo L. Bottasso, Dipl.-Ing. Dr. Stefan Höltinger

Vienna, July 2018

Statutory declaration

I declare that I have developed and written the enclosed Master Thesis completely by myself, and have not used sources or means without declaration in the text. Any thoughts from others or literal quotations are clearly marked. The Master Thesis was not used in the same or in a similar version to achieve an academic grading or is being published elsewhere.

Vienna, 30th of July 2018 Signature Katharina Gruber

Katharina Gruber

Abstract

Brazil's hydropower generation, which makes up the majority of electricity generation, is subject to significant inter-annual variability. In order to reduce risks in the system and to increase power generation due to growing demand, it has become necessary to promote other generation technologies, in particular from renewable resources. Wind power has high potential, especially in the North-East region of Brazil. To assess if the El Niño and La Niña cycles have an influence on wind power generation and if wind power generation is correlated with water inflows into Brazilian hydropower plants, a simulation model of the long-term wind power generation (i.e. 1980-2016) in Brazil has been developed. The model is based on MERRA-2 reanalysis wind speed data and corrected with wind speeds measured by the National Meteorological Institute (INMET) and with wind power generation data from the homepage of the National Electrical System Operator of Brazil (ONS), and is spatially disaggregated to the level of states. Different methods of interpolation, data cleaning, and bias-correction were tested, to find the best procedure for the simulation of wind power generation. Finally, impacts of El Niño and La Niña on annual wind power generation and correlations with water inflows are assessed. Results show that applying bias-correction using wind power generation data, substantially improves the quality of the simulated time series. In contrast, bias-correction with wind speed data did not improve results in the South very much, but had a larger impact on wind power in the North-East. In most tested settings, the best results are obtained when Nearest Neighbour Interpolation is used with wind power bias correction. Results, however, differ significantly between single states. Results with El Niño and La Niña indices are ambiguous: For several cases impacts after one to three months are found, for other after five to eight. El Niño shows a higher impact than La Niña or both together – whether it is positive or negative may vary from region to region. Respecting these events can help in the prediction of future wind power generation. Water inflows show no significant correlation with wind power generation, when seasonality is removed from the timeseries.

Kurzfassung

Die Wasserkraft, die in Brasilien einen Großteil der Stromproduktion ausmacht, unterliegt beträchtlichen jährlichen Schwankungen. Um die dadurch entstehenden Risiken zu verringern und auch die steigende Stromnachfrage zu decken, ist es notwendig auch andere Technologien, insbesondere erneuerbare Energieproduktion, zu fördern. Windkraft bietet dafür besonders im Nordosten hohes Potenzial. Um die Auswirkungen von El Niño und La Niña auf die Windkraft sowie Zusammenhänge zwischen Wasserzuflüssen in Brasilianische Wasserkraftwerke mit Windkraft zu bewerten, wird ein Modell zur Simulation von Windkraft in Brasilien über mehrere Jahrzehnte (1980 – 2016) erstellt. Dieses Modell, welches auf MERRA-2 Reanalyse Windgeschwindigkeitsdaten basiert, wird einer Fehlerkorrektur mit Windmessungsdaten des Nationalen Meteorologischen Instituts (INMET) und Windkrafterzeugungsdaten des Nationalen Stromnetzbetreibers von Brasilien (ONS) auf der Ebene einzelner Staaten unterzogen. Unterschiedliche Methoden zur räumlichen Interpolation, Datenbereinigung und Fehlerkorrektur wurden getestet, um die beste Möglichkeit zur Simulation von Windkraft zu ermitteln. Anschließend werden die Auswirkungen von El Niño und La Niña auf die Windkraftproduktion sowie Zusammenhänge mit Wasserzuflüssen bewertet. Die Ergebnisse zeigen, dass die Windkraft-Fehlerkorrektur erheblich zu einer Qualitätsverbesserung der simulierten Zeitreihen beiträgt, die Windgeschwindigkeits-Fehlerkorrektur zumindest im Süden hingegen weniger, jedoch einen stärkeren Einfluss im Nordosten aufweist. Die Simulation, bei der die „Nächster Nachbar“ Methode in Kombination mit Windkraft-Fehlerkorrektor angewendet wird, erzielt im Allgemeinen die besten Ergebnisse – die Resultate unterscheiden sich jedoch in einzelnen Staaten. Die Analyse des Einflusses von El Niño und La Niña liefert keine eindeutigen Ergebnisse: In einigen Fällen werden Auswirkungen nach bis zu drei Monaten gefunden, in anderen erst nach fünf bis acht. El Niño scheint einen größeren Einfluss auf die Windkraftproduktion zu haben als La Niña oder beide zusammengenommen – die Art der Auswirkung (Steigerung oder Senkung der Windkraftproduktion) variiert je nach Region. Eine Berücksichtigung dieser Ereignisse kann dabei helfen, zukünftige Windkraftproduktion vorherzusagen. Die Korrelationen von Windkraftproduktion mit Wasserzuflüssen in brasilianische Wasserkraftwerke nach Entfernung saisonaler Effekte sind nicht signifikant.

Acknowledgement

First of all, I would like to thank my family, especially my parents, for supporting me not only during my thesis, but for the whole time of my studies and education. Thank you not only for your financial but also and even more for your moral support and for awakening my interest in science and learning since I was a little girl.

I am also very grateful for the outstanding support of my advisor Johannes Schmidt, who provided me with many resources and gave me lots of useful inputs for this thesis. Thank you for helping me with all my concerns.

Furthermore I want to thank my co-advisor Carlo Bottasso for his advice.

I also want to thank my other co-advisor Stefan Höltinger, who also provided important inputs for my work and took the time for all my questions.

My thanks also go to Johann Baumgartner, who provided help with the acquisition of data.

Moreover, many thanks to my friends and colleagues, who made my student days more pleasant.

And very special thanks to Peter Gsellmann for providing his knowledge and patience one more time.

Contents

List of Abbreviations

Interpolation methods

NN … Nearest Neighbour

BLI … Bilinear Interpolation

IDW … Inverse Distance Weighting

NNc … Nearest Neighbour with wind power correction

BLIc … Bilinear Interpolation with wind power correction

IDWc … Inverse Distance Weighting with wind power correction

Wind speed correction methods

x … monthly and hourly wind speed correction, without adaptations

r… monthly and hourly wind speed correction with long rows of 0 removed

m … monthly and hourly wind speed correction with adaptation of mean

rm … monthly and hourly wind speed correction with both of the above

nINc … without wind speed correction

List of Figures

List of Tables

1 Introduction

A growth in the demand of energy can, as in other emerging economies, also be observed in Brazil, a country which is characterised by a large share of renewable energy, mainly due to considerable amounts of installed hydropower capacity, covering around 70 % of electricity demand. In recent years, power generation from water resources was impeded as a consequence of droughts, especially in the years 2001-2002 and 2014-2015, leading not only to deficits of electricity and failures, but also to depletion of water reservoirs. Therefore, it has become essential to search for and foster other sources of energy, especially renewable ones.

Historically, during lacking hydropower generation, mostly fossil fuels, especially natural gas, were used to compensate deficiencies. In more recent years, also sugar cane has proven as an important resource for energy generation, but also wind power and photovoltaics have gained popularity. Wind power has been deployed in Brazil since 2006, especially in the North-East and South regions of the country. Due to the seasonally anticyclical generation patterns compared to hydropower, it offers a particularly interesting source of electricity for compensating lacks of hydropower generation.

These forms of energy generation have also been encouraged by political incentives, especially the PROFINA programme, which aimed at amplifying the share and diversity of renewables in the Brazilian energy matrix.

Due to the growing importance of wind power in the Brazilian electricity system, it is important to understand the dynamic characteristics of this resource. In this thesis I therefore simulate wind power generation data to assess its dynamic behaviour, in particular with respect to El Niño and La Nina events as well as water inflows into Brazilian hydropower plants. The following sections introduce to the topic by giving an overview of the energy landscape in Brazil and outlining the El Niño and La Niña phenomena, before presenting the research questions to be examined.

1.1 Introducing wind power into the Brazilian power system

Following the trend of global growth of demand for energy, also Brazil has witnessed a yearly increase in the need for electricity by about 4 % in the period of 2004 to 2013 and since then by even more, i.e. 4.2 % [1] [2]. The drivers for this growth and the subsequent expansion of generation capacities are economic as well as population growth, urbanisation, higher living standards, the increase of electricity access, which is currently at a level of 95 %, and also the need to prevent electricity crises as a consequence of lack of generation capacities due to extreme weather events [1] [3].

Brazil is a country, which has always been dominated by hydropower in electricity generation: The world's second largest hydropower plant, the Itaipu Dam, is located at the Brazilian-Paraguayan border [3] and a share between 69 % and 84 % of Brazil's electricity generation was provided by hydropower in the years 2004 to 2013 [1]. However, the role of hydropower has been declining in recent years: In 2000 as much as 95 % of electricity demand were covered by hydropower, but 15 years later, this percentage had already declined by more than 10 % to mere 83 % [3]. These reductions are an effect of slower capacity additions, which resulted from concerns in the hydropower industry due to droughts and the lacks in generation they caused [3].

Usually small seasonal or yearly deficiencies in hydropower generation can be compensated from the many reservoirs built in Brazil, which need to be refilled during times of higher water inflows in Brazilian summer and autumn months (January to June) [3]. However, this is not possible, if there are longer periods of droughts, as in the years 2001-2002 [4] or 2014-2015, when Brazil was struck by electricity crises [1] [2]. Consequences were deficits in electricity generation, causing black outs, loss of load, expensive dispatch of thermal power plants and an increase of electricity prices to more than double [1] [2]. Furthermore, with declining water levels in reservoirs, also other conflicts emerge: Trying to replenish them, the water resources lack in other domains, especially in agriculture [3].

In order to provide energy security, i.e. avoiding price volatility and supply disruptions, precautionary measures need to be taken in the form of expansion of generation capacities, diversification of Brazil's energy matrix and improvement of the electric grid for better interconnection between different regions [1] [2] [4] [5]. At the turn of the century, transmission links were already built between Brazil, Argentina and Uruguay, to take advantage of complementary hydrological cycles [3].

Also within Brazil it is possible to benefit from the complementary cycles of different electricity sources. In this context, attention has to be paid to climate change, which has various effects on electricity generation in Brazil: In the future, hydropower generation is expected to decline between 2 and 10 % due to effects connected to climate change [2]. As the country relies mainly on hydropower, changes in the usual hydrological cycle caused by extreme events affect rainfalls and therefore water inflows [4]. For example, the Brazilian state of Amazônia, which is normally a warm and humid region, has experienced low water levels or even dry rivers due to unusual weather conditions [4]. In years to come, a decrease in rainfall is expected in southern and central regions of this state, as well as in west central Brazil and large parts of the North-East, whereas in other parts of the country, including southern Brazil and the coastal region between Amapá and Ceará, there are predictions that there will be enhanced precipitation, with a trend of general reduction of water resources by 2100 [4].

One attempt to overcome this lack in hydropower generation due to climatic conditions and growth in demand is simply an expansion of the current capacity, for which the government is making plans in the North of Brazil [1]. However, an obstacle to these projects may be ecological and social problems: As many new plants are planned to be located in the amazon forest and building them uses large land patches, natural resources are impaired, habitats destroyed and populations need to be resettled [1]. Furthermore, creating large hydropower reservoirs in forest regions implies cutting down considerable areas of trees, which influences the hydrological cycles negatively and leads to declines in rainfalls due to reduced evapotranspiration [1].

These uncertainties and negative implications for environment and society lead to insecurities in electricity supply and concerns about future energy provision and therefore create the need to adapt Brazil's energy sector and foster other means of generating electricity, in order to compensate for lacks of hydropower production [2] [6] [4]. Renewable electricity generation from PV and wind power provide good alternatives, compared to natural gas, because they are environmentally friendly and abundantly available in Brazil: There are plenty of solar resources all over the country and wind power, which is predicted to rise by up to 10 % in the coming years, is deemed profitable especially in the North-East and in the South regions [1] [2] [7] [8]. The optimal mix of these two and hydropower has been examined in [1] and resulted in 37 %, 9 % and 50 % for solar, wind and water, respectively, when taking into account the capacity of existing hydropower reservoirs for balancing purposes. It has also been found that integrating several renewable resources not only reduces variability in electricity generation but also the need for reserves [6]. Apart from higher energy security, an energy diversification also delivers more stable electricity prices and if renewable resources are used, also aims of decarbonisation, health and environmental benefits are met [3].

Due to the anticyclical behaviour of wind and water resources, wind power seems especially convenient as an alternative to hydropower expansion, as it can help in stabilising supply [2] [7] [4]. Wind power has been used in Brazil since 2006 [1], and witnessed a rapid increase especially since 2010 (see Figure 3), reaching 1.2 % of demand by 2014 [1]. A decrease in costs of wind power of about 20 % since 2010 is one of the driving forces of this expansion [3]. As the government recognised this potential, political programmes were initiated, to promote wind power: The 10-year Energy Expansion Plan of 2014 intends to increase the share of wind power to 8 % by 2024 [3]. Auctions constitute a very popular instrument for fostering renewable energy, not only in Brazil [3]. Many were held in Brazil within the scope of PROFINA, the Alternative Energy Source Incentive Programme, which aimed at including a higher share of renewable energy in the Brazilian energy matrix and was supported financially by the Brazilian Development Bank [8].

However, a higher share of wind power in the electricity generation matrix is accompanied by a higher variability due to the intermittent nature of wind resources [9], which influences not only the availability of wind power but also the electricity market, bidding strategies and decisions of agents [10] [11]. In order to find out about the potential of integration of this intermittent source of electricity into the power system, it is important to understand seasonal, annual but also variabilities in the short and the very long-term [11]. Compared to other sources of renewable energy, wind power shows the biggest uncertainties [10], and therefore needs to be adequately modelled to estimate needs for flexibility [11] [12].

Spreading wind parks over different parts of the country can help in reducing variability of wind power by harnessing complementarity between different regions [12]. Therefore, if assessing incorporation of different electricity resources into a power system, it is necessary to examine aggregates of geographically dispersed power plants, to understand the smoothing effect [12] [13] [14]. Attempts of modelling wind power generation by upscaling wind power production generated from measurements of a single location do not provide accurate information about the actual possible needs for reserves, as regional differences and synergies are ignored [9].

Apart from helping in the management of variability of water resources, increasing the share of wind power also brings other benefits of socio-economic nature, such as creating jobs and support of local industrial development [3].

These advantages and circumstances show the important role of wind power in Brazil, especially in the future of Brazil's energy sector, which is facing challenges from more extreme climatic conditions as well as from growing demand and political sanctions in favour of environmental protection.

1.2 El Niño and La Niña and their meaning for the Brazilian energy sector

In the face of the impact that global climate change has on renewable energy in Brazil, especially on wind power, it is also important to identify other sources of extreme events, which pose challenges for transmission system operators and investors. One of these effects are the El Niño and La Niña Oscillations, which are known to affect hydrological cycles and thus hydropower generation and therefore the majority of Brazil's electricity sector, especially in the North-East of the country [7] [4]. The names "El Niño" and "La Niña", also known as "ENSO" (short for El Niño Southern Oscillation) stand for regularly recurring events of oceanic-atmospheric phenomena characterised by warm (El Niño) or cold (La Niña) currents of superficial water of the Pacific Ocean [15] [16] [17]. These streams of warm water affect regional as well as global climate by altering wind directions and therefore influencing where water masses are transferred [15] and in consequence causing floods, droughts, tropical storms and heat waves [17]. Although El Niño and La Niña phenomena do not occur annually, they show a considerable seasonality, when looking at the monthly averages (Figure 6). In eastern Latin America this is experienced as dry and hot periods during El Niño in the more northern regions but wet weather in the south, whereas La Niña brings cold and rainy conditions in the months between December and February [15]. In the Brazilian winter months (especially between June and August), different conditions are perceived: During La Niña events the area is affected mainly by heat, whereas during El Niño also rainfalls above average are usual [15].

Only two years ago, at the turn of the year from 2015 to 2016, one of the strongest El Niño events since records exist occurred and brought exceptionally hot weather but also increased rainfalls worldwide, which led to floods and landslides in north-east Brazil and may even be responsible for the outbreak of the Zika virus [16] [18] [19]. On the other hand, in the South the change in weather conditions led to extremely hot and dry weather, favouring forest fires and diminishing sugar, corn and soy bean production [18] [19], but also reducing water resources and emptying reservoirs, which caused political unrest [20]. São Paulo was one of the most strongly affected cities, as its water management system had difficulty coping with the masses of water coming down [19]. However, there are also critical voices, which doubt connections between these phenomena and changes in weather, such as the reduction in winds in the US in 2015 [21].

Nevertheless, it cannot be denied that there is an impact of El Niño and La Niña events, at least in certain regions, where extreme weather conditions lead to consequences for agriculture, power generation, health, ecosystems among others [5] [17]. Therefore, it is of utter importance to understand these phenomena, in order to take actions in advance to prevent the worst impacts. Since the consequences of ENSO on so many areas of the world are nothing new, already in the seventies of the past century, after a severe decline in fishery in Peru induced by El Niño in 1972-1973 affecting world economy, effort was put into studying these events [17]. Nowadays, with renewable energy becoming increasingly popular, and especially in Brazil which electricity is supplied mainly by volatile sources, it is also necessary to study the impacts of El Niño and La Niña on the power system, as growing demand in the future could challenge the grid and lead to instabilities, if it is jeopardised by such events, but if these ramifications are known beforehand, it is possible to make provisions for energy supply [3]. This may be particularly important, considering that ENSO often brings droughts, which impair hydropower and thus, as it is the dominating electricity resource of Brazil, the whole power grid's stability is affected [5]. Also for future projects, such as new wind parks, it is necessary to consider the impact El Niño and La Niña events can have, as they may influence the statistical characteristics of the site and its wind resources [22].

1.3 Aim and structure of the thesis

For this thesis, a model originally developed only for the North-East of Brazil is extended to simulate wind power generation for the whole country from freely available reanalysis data, using different bias correction methods for wind speed and wind power correction and validating against observed wind power generation data. The results are analysed and the most favourable method is chosen, and applied in an evaluation of impacts of El Niño and La Niña on wind power generation as well as correlations with water inflows into Brazilian hydropower plants. The resulting time series and insights gained from these analyses can help in creating energy system models, which are an essential tool for planning future energy provision and therefore support policymaking and political decisions.

The thesis aims at analysing the following research questions:

- Which interpolation and wind speed correction methods can improve approximation of a model based on MERRA-2 reanalysis data to observed wind power generation data?
- Which correlations of resulting time series with water inflows into Brazilian hydropower plants as well as El Niño and La Niña can be identified?
- Which differences between different regions can be observed in quality of the simulation as well as impacts of El Niño/La Niña and of water inflows?

2 Data and methods

In this chapter, the data used for creating and bias-correcting the model presented in this thesis as well as the data used for validation and subsequent analysis are described. The graph in Figure 1 gives an overview of the used data and methods, which will be outlined in the following. Some of the data as well as methods are already applied in an analysis of the wind power generation in the North-East of Brazil in [23], but some new methods are added and also new data are used, in order to make the simulation more exact and also with the aim to simplify the process by automation.

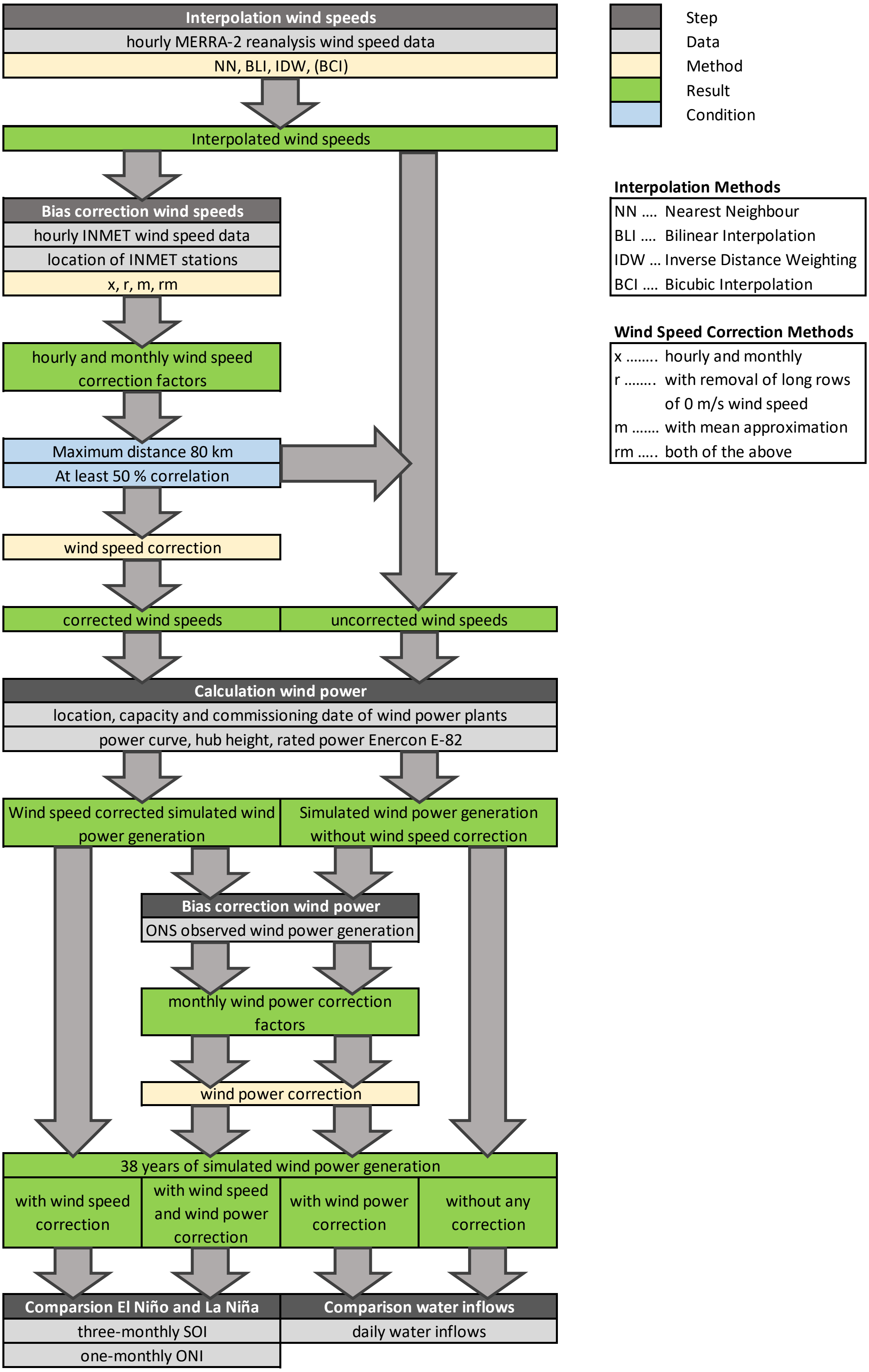

Figure 1: Overview of the approach used in this thesis (own representation)

2.1 Data: Access and preparation

There are several sources of data, which are used in this thesis for two purposes: the generation of a model, which simulates wind power generation output from reanalysis wind speed data and the examination of possible correlations of wind power generation with the El Niño and La Niña cycles as well as with hydrological cycles in Brazil. The used data are listed in Table 1, together with their temporal availability (at the moment of download):

Table 1: Summary of data used for modelling of wind power and analysis (own representation)

Data set name	Description	Temporal resolution	Coverage	Source
MERRA-2	reanalysis data, modelled wind speed data	hourly	1980-August 2017	NASA
BDMEP	wind speed measurement data	hourly	1999-2016	INMET
Wind farms	wind park data, geographical locations and installed capacities with commissioning dates	monthly	1998-2017	The Wind Power
Enercon E-82 wind turbine	power curve			Enercon
Histórico da operação	wind power generation data	monthly	2006 - July 2017	ONS
Histórico da operação	wind power generation data	daily	2006 – October 2017	ONS
Oceanic Niño Index	El Niño and La Niña index	monthly	1951 - November 2017	NOAA
Southern Oscillation Index	El Niño and La Niña index	three-monthly mean	1950 – October 2017	NOAA
COPA	inflows into Brazilian hydropower plants	daily	1979-2014	COPA Model

For the wind power generation data, it has to be mentioned that the start date depends on the beginning of wind power generation in the specific region and therefore is not 2006 but the installation year of the first wind turbine in the region. Furthermore, the data are downloaded per state as well as per subsystem and for all of Brazil (the latter two only for analysis and the first also for bias correction).

In the following, the data will be described, as well as the access to them and how they were prepared before use.

2.1.1 MERRA-2

The MERRA-2 data are a reanalysis dataset, which means they are modelled from an analysis of satellite data. The data are provided for free by the National Aeronautics and Space Administration (NASA). As they have already been described in [23], they will not be further explained here. Contrary to the first approach of the simulation presented in [23], which was only for the North-East of Brazil, in the thesis at hand the data are not downloaded via the command line with wget64, but with a script which can be found at [24]. This and other tools for dealing with MERRA data can be found at [24]. For using the script, the borders of the area that shall be downloaded have to be specified (latitudes between -36 and 5.5, longitudes between -74.1 and -33). Also the time span as well as the variable that shall be downloaded need to be specified. As data are available since 1980, all the data were downloaded until the date when the download was started (31st of August 2017). Five variables are downloaded, the wind speeds in u- and v-direction at 10 and 50 m height (U10M, U50M, V10M, V50M) as well as the disposition height.

The points downloaded amount to 5544 (in a grid of 66 in longitude and 84 in latitude). The files are downloaded in the .nc-format per day for the whole region, but are read per point during the simulation, which makes it necessary to transform them. They are first converted to the .feather format, which makes them faster readable. As it was impossible to load data for all years for such a big area as all of Brazil into the memory and the daily data shall be joined, this is done five-yearly. Later, when the data are rearranged and saved per point, the time series are merged to nearly 38 years (1980 until August 2017). The longitudes and latitudes as well as the dates are saved separately. Later the data can be read by a reading function ("getMerraPoint") by passing the longitude and latitude of the wanted location.

2.1.2 Wind speed measurements

Wind speed measurement data provided by the National Meteorological Institute of Brazil (Instituto Nacional de Meteorologia, INMET[1]) are used to bias-correct reanalysis wind speeds. The download of the data is carried out with a script, which is attached at the end of the thesis. There are 481 files downloaded, one for each wind speed measurement station; in a meta data file, the stations' names together with the locations are listed. With the help of the states' abbreviations, three of the stations can be filtered from the whole data, as they start with "AA" and "UY", which stands for Antarctica and Uruguay, respectively. It is confirmed, that no other station is outside of the downloaded area, by checking whether minima and maxima of longitudes and latitudes of the stations are inside of the

[1] http://www.inmet.gov.br/

limits for the download of MERRA-2 data. The distribution of the remaining 478 INMET wind speed measurement stations is depicted in Figure 2.

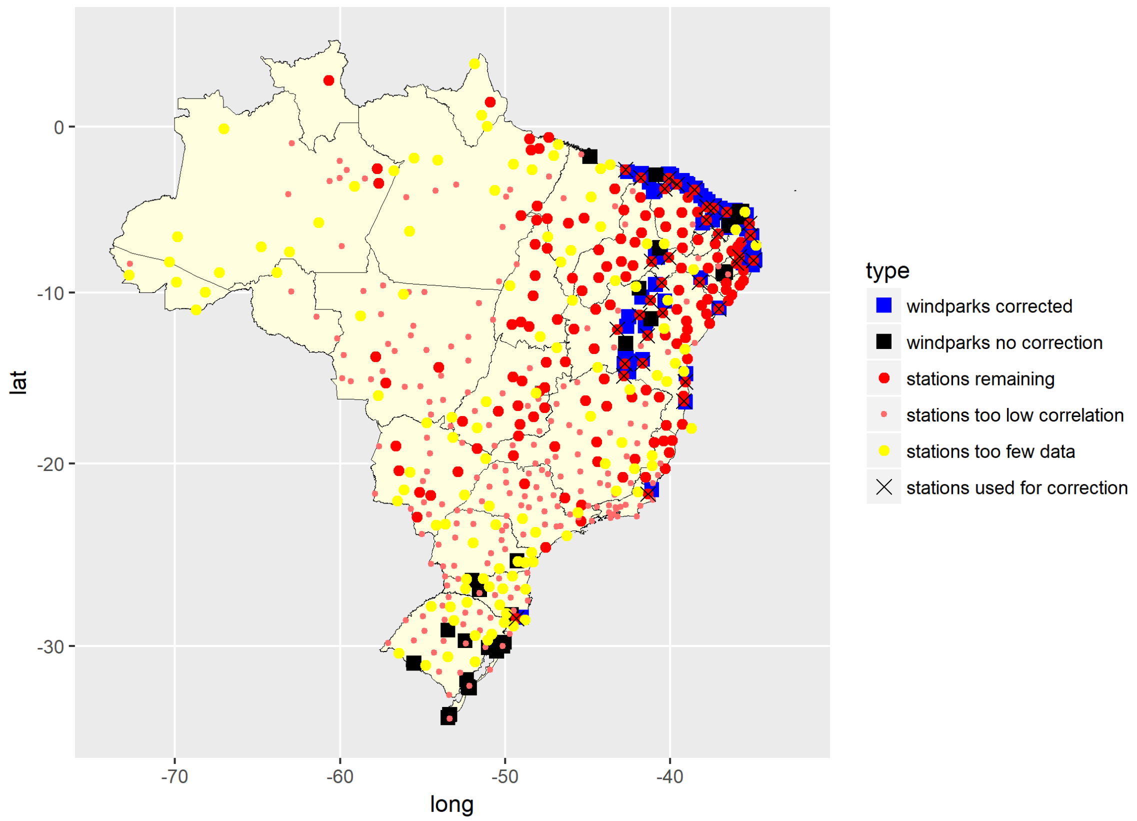

Figure 2: Location of INMET wind speed measurement stations and wind parks (own depiction with data from [25], [26] and [27])

2.1.3 Wind parks

As it was found in [23], that the wind park data used provided by ANEEL[2], did not provide very good accurateness in terms of geographical location, because only the municipalities of the particular wind parks were known, for this simulation different data are used. They are available on the homepage of The Wind Power[3]. The information provided comprises the name of the wind farm, the country and county (state) it is located in, the city at which it is located, the commissioning date, the number and type of installed wind turbines, the installed capacity, the geographical coordinates and some more. For further use, information on the name of the wind farm, the state it is located in, the installed capacity, the longitude and latitude of the location, the number and type of turbines installed, the type (onshore or offshore, but only onshore wind farms are selected) and the commissioning (year and

[2] http://www.aneel.gov.br/

[3] https://www.thewindpower.net/

month) are extracted. However, not always all of the information is specified; therefore, missing information is identified and has to be searched from other sources.

For parts of two wind parks, Taiba and Ventos de São Benedito, the installed capacities were missing when the data were downloaded, but added later, and therefore could be added from the website (5000 and 2800 kW, respectively). 48 wind parks were lacking their geographical coordinates. First, it was attempted to find the municipalities of the wind farms with missing coordinates in data from the National Agency of Electrical Energy of Brazil (ANEEL), which were also used in [23]. The data are downloaded with a script, which can be found in the appendix. 18 stations could be matched and the coordinates were fetched from the barycentres of the municipalities, which were calculated from a shapefile downloaded from [25] (for more detailed information on this step see [23]). Three municipalities in the ANEEL file (João Câmara, Caldeirão Grande do Piauí and Curral Novo do Piauí) could not be matched to those of the coordinates calculated from the shapefile due to special characters, so they were inserted manually. For the remaining 30 wind parks the approximate location of each wind farm is searched on the internet; most of the municipalities are given in the wind park data on the websites of The Wind Power. This way, the exact location can be identified on openstreetmap.org, where wind turbines are mapped. The links, where information about the (approximate) location of wind farms with missing coordinates in The Wind Power data was found, are listed in the appendix. Table 2 shows the location data added (sometimes there were several parts of one wind park, in those cases the location is listed only once). Some of the longitudes have the wrong sign, which is therefore changed.

Nine wind parks lacked commissioning dates, of which two (Corredor do Senandes I and Primavera) were found not to be finished yet, and therefore the dates were set to the future (2018/01). The commissioning dates found for the remaining seven wind parks are displayed in Table 3 (Santo Inácio consists of two parts but is listed only once). The sources which indicate commissioning dates are listed in the appendix. For the wind park Ventos de Santa Edwiges an obviously wrong commissioning date (21706) is listed, which is corrected to 2017/06.

In five cases (Pedra Cheirosa I – II, Ventos de Bahia II, Ventos de Santa Edwiges, Ventos de Santa Regina and Ventos de Santo Adriano) the states in which the wind parks are located were missing. The coordinates were searched online and this way the states (Ceará, Bahia and for the last three Piaui) could be determined and were added to the wind park data.

Four wind parks in the list (Eurus I, Eurus II, Fernando de Norhona, Terral) are not located in Brazil, but in Mexico (the first two), on an island about 400 km away from the mainland of Brazil (and therefore outside the region for which reanalysis data were downloaded – computational effort of downloading a larger area of reanalysis data is considered to exceed the merit of including this location in the analysis) and in Spain and are therefore eliminated from the data by checking whether the coordinates are inside the longitude and latitude borders of the realm selected for the download of reanalysis data.

Table 2: Missing locations of wind parks (sources see appendix)

Name	Latitude	Longitude
Antônio Pimentel de Sousa	-3.70292	-38.47351
CGE Delta 3 II	-2.7269	-42.5862
CGE Delta 3 IV	-2.7269	-42.5862
CGE Delta 3 V	-2.7269	-42.5862
CGE Delta 3 VI	-2.7269	-42.5862
Colonia	-3.6078	-38.9664
Geraldo Júnior Cavalcante Lopes	-3.70292	-38.47351
IMT	-25.42770	-49.27363
Instituto Federal de Educação-Ciência-Tecnologia Sul	-31.75757	-52.33329
Lagoa Seca	-2.8715	-40.0901
Malhadinha	-3.9221	-41.0289
MEL 02	-4.9447	-36.9656
Miassaba III	-5.1128	-36.3874
Pajeu do Vento	-14.1200	-42.5523
Pedra Cheirosa I - II	-2.9528	-39.8891
Pedro Pedron	-3.8744	-38.3833
Santo Inácio	-4.7710	-37.3082
Stela Maris Zambelli	-3.8744	-38.3833
Tarlene Guedes Bessa	-3.70292	-38.47351
Trari	-3.26969	-39.26847
Vento do Oeste	-2.8707	-40.0902
Ventos de Bahia II	-11.9965	-41.5109
Ventos do Araripe III	-7.3434	-40.5657
Vila Para I	-5.2115	-37.0242
Vila Para II	-5.2115	-37.0242
Vila Para III	-5.2115	-37.0242

Table 3: Missing commissioning dates (sources see appendix)

Name	Commissioning
Assurua	*2017/10*
Eolica Sao Cristovao	*2014/03*
Olinda	*2008/07*
Santo Inácio	*2017/07*
Taiba	*1998/12*
Ventos de São Benedito	*2016/05*
Ventos do Brejo A-6	*2011/06*

Figure 3 depicts where and in which year wind parks were commissioned and also the capacity of wind parks (since 2006, because prior there was no notable wind power capacity installed). This shows the main wind power generation regions: the North-East and South. Furthermore, it can be observed that until 2010 wind power was not as important as since then, because the majority of installations happened in the past seven years.

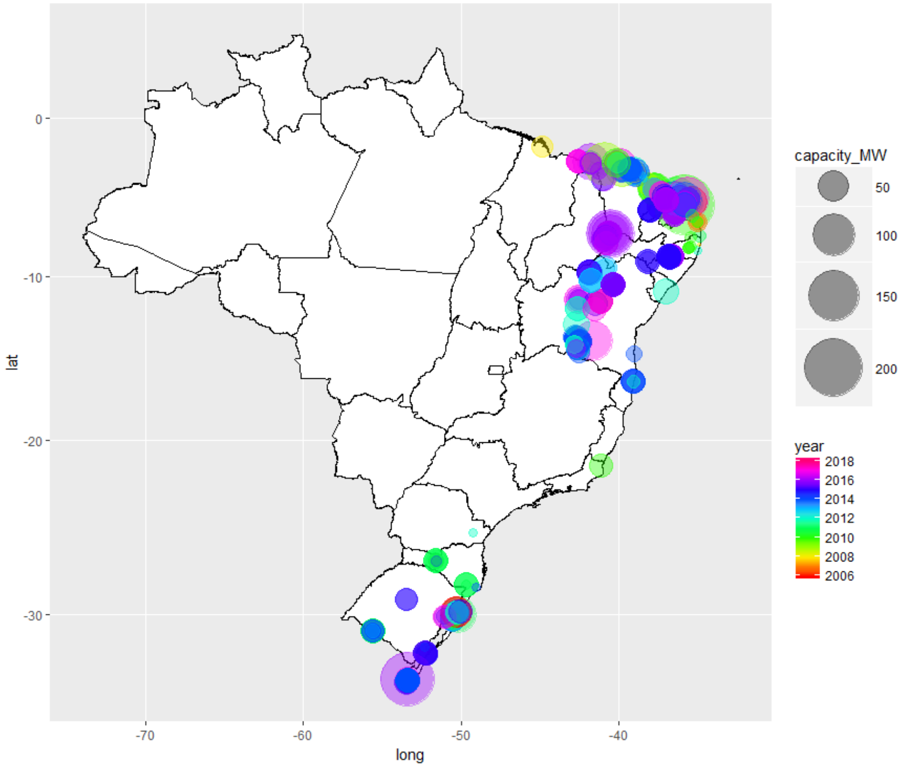

Figure 3: Location, capacity and commissioning year of wind parks (own depiction)

2.1.4 Power curve

For simulating wind power from wind speed, the power curve of a wind turbine and the height are needed. The data of the number and type of installed wind turbines were included in the list of wind parks, because it was planned to use this information for the simulation. However, it turned out that the information was not always complete as sometimes there was only given the number of turbines, for other locations only the manufacturer but not the model or no information at all. Furthermore, in Brazil's wind parks about 50 different types of wind turbines were installed and for each of them data about the power curve would have to be identified. As this was considered too much effort for this thesis, especially if good results can be achieved when using only one standard turbine (as in [23]), this method was chosen also for this simulation.

The selected wind turbine is Enercon E-82, although this turbine did not appear in the wind park data. Only other models from Enercon (Enercon E-48 with 0.8 MW rated power and Enercon E-40 with 0.6

MW rated power) were listed, but among the installed turbines, there were also several ones with a higher rated power (for example Gamesa G90 with 2 MW rated power, Suzlon S95 with 2.1 MW rated power or Acciona AW-3000/125 with 3 MW rated power), which is why a wind turbine in a higher range with 2 MW rated power, a rotor diameter of 82 m and a hub height of 108 m was selected for the simulation. The power curve can be found in the factsheet of the wind turbine [28] or in [23].

2.1.5 Wind power generation

For wind power generation correction and later also for comparison, observed wind power generation data are needed. They are downloaded from the homepage of the National Electrical System Operator (Operador Nacional do Sistema Elétrico, ONS) of Brazil [29] as .csv-files. This website has been updated since [23] was written and now data are available in different spatial and temporal resolutions: Per day, per week, per month as well as per year and for all of Brazil, per subsystem, per state as well as per power plant. For the simulation and validation monthly as well as daily data are downloaded per state, subsystem and for all of Brazil in GWh. The monthly data are downloaded starting in 2006 (if available, depending on when wind power generation started) until July 2017. For comparison with results from the simulation also daily and monthly data are downloaded per subsystem (South and North-East, because only in these regions notable amounts of wind power are generated) as well as for all of Brazil until August 2017 and daily data per state until October 2017. Table 4 gives and overview of the downloaded data and their use.

Table 4: Wind power generation data used for validation (own depiction)

Temporal resolution	Spatial resolution	Available parts	Data until	Use
daily	states	Bahia, Ceará, Maranhão, Pernambuco, Piaui, Rio Grande do Norte, Rio Grande do Sul, Santa Catarina	October 2017	comparison
daily	subsystem	South, North-East	August 2017	comparison
daily	Brazil		August 2017	comparison
monthly	states	Bahia, Ceará, *Maranhão*[4], Paraíba, Paraná, Pernambuco, Piaui, Rio de Janeiro, Rio Grande do Norte, Rio Grande do Sul, Santa Catarina, Sergipe	July 2017	correction and comparison
monthly	subsystem	South, North-East	August 2017	comparison
monthly	Brazil		August 2017	comparison

[4] Maranhão is only used for comparison but not for wind power correction, as wind power generation started only recently in May 2017 and therefore not enough data are available.

2.1.6 El Niño and La Niña

With the help of the determined correction factors, different simulations of nearly 38 years (37 years and eight months) of wind power in Brazil at full capacity can be calculated. These simulations are compared to indices of El Niño and La Niña events. For investigation of the impact of these events on wind power generation in Brazil, two different indices are consulted: The Oceanic Niño Index (ONI), an index based on a three-monthly mean, and the monthly Southern Oscillation Index (SOI). The indices are available at the homepage of the National Centers for Environmental Prediction (NCEP) [30] and the National Climatic Data Center (NCDC) [31] from the National Oceanic and Atmospheric Administration (NOAA). The first is copied as a table and saved as .csv-file, after removing the header lines every ten rows; the latter is provided in .csv-format at [32] and saved as such. The monthly or three-monthly data are available since 1950 until present.

2.1.7 Water inflows

Apart from influences of El Niño and La Niña events on wind power generation, also the correlations between water inflows into Brazilian hydropower plants and simulated wind power generation are examined. The aim is to find out whether, after removing seasonal fluctuations in wind power generation and water inflows, there is a correlation between the two, meaning that times of higher rainfall and therefore higher water inflows can be associated with an increase or decrease in wind power generation. The daily water inflow data are available between 1979 and 2014 per subsystem from an analysis of the optimal mix of renewable energies in Brazil [1] and can be downloaded at [33].

2.2 Methods

This chapter describes the methods used in this thesis. In many aspects they are similar to the methods used in [23], but in others there are differences or several variations of methods have been applied. The focus in the description of methods will be on the additions made to methods already applied in the simulation of wind power only in the North-East of Brazil.

2.2.1 Wind speed interpolation methods

When examining the wind power generation in a certain area, the wind speed at the specific location needs to be determined. For this, four different methods are used: the Nearest Neighbour Method, the Bilinear Interpolation, the Bicubic Interpolation, and Inverse Distance Weighting. Contrary to [23], all these methods are actually applied in the simulation. The Nearest Neighbour method is the simplest of these methods and consists in defining the grid point which has the smallest geographical distance to the desired point of interpolation and applying data from this point [34]. It is not only fast, but also suitable if many data points are available [35]. Bilinear interpolation uses four surrounding points and is a simple, yet accurate method, which can be applied if data are available in a grid [36]. Inverse Distance Weighting is a method where it is assumed that data from surrounding points around the point of interest influence data on this point inversely to their distance. It is fast and can be used for interpolation from points [34]. For keeping this method simple and fast and also because grid points are not very close with a distance of about 50 km, only four neighbouring points are considered for interpolation. Inverse Distance Weighting is a commonly used method in meteorology [35], which in several cases has proved [37] [38] to be the best method for interpolation among others, such as Kriging. It sometimes occurs that during Bilinear or Bicubic Interpolation one point of interest is on the longitudinal or latitudinal line between two MERRA-2 grid points (and not in a square surrounded by four); in this case, the interpolation is only performed between two points (for the Bilinear Interpolation) or only between six points (for the Bicubic Interpolation). Only the Bicubic Interpolation is discarded when calculating wind power generation, as in some cases (at wind speed measurement stations in Buriticupu, Caixas, Nova Maringá, Nova Ubiratã and Ituporanga) it delivered negative wind speeds, because in a few cases, when interpolation was done, the resulting function was negative on the point of interpolation, even though the wind speeds at the 16 surrounding points were positive. For an example (taken from interpolation at wind speed measurement station in Buriticupu for data of 31st August 2017 at 8:00:00) of one of these cases, see Figure 4: The wind speeds at the 16 MERRA grid points (within longitudes -47.5 and -45.625 and latitudes -5 and -3.5) are all positive (black dots), whereas the interpolated wind speed (white dot, at longitude -46.4495 and latitude -4.320597) is in the negative range at – 0.013 m/s (the plane where wind speed is zero is in red).

Table 5 provides a summary of these spatial interpolation methods. During interpolation, the effective wind speed needs to be calculated, which is done by the Euclidean norm. For inter- and extrapolation the logarithmic wind profile (see [23]) is applied, where the wind speed in a certain height depends on the ratio of the heights and an exponent (alpha friction coefficient), which is higher the lower the surface roughness. In the following, the subsequent steps will be explained.

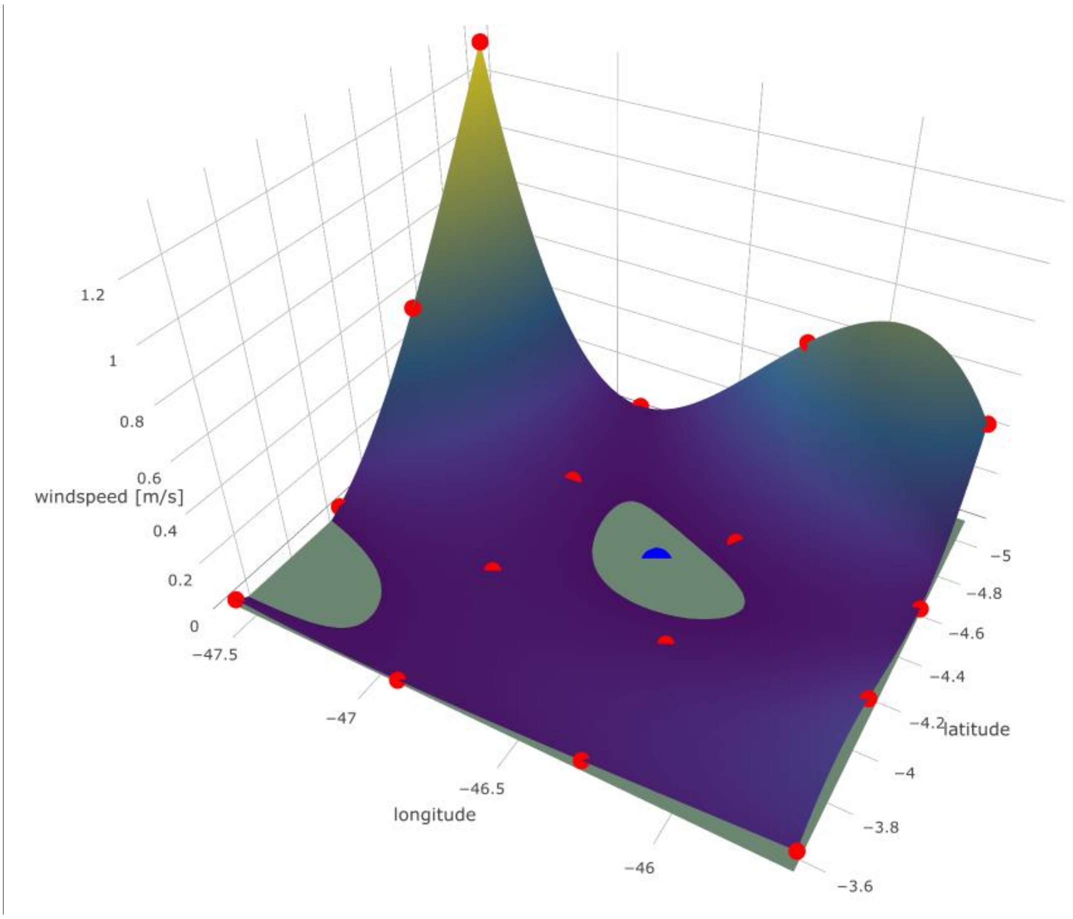

Figure 4: Example of Bicubic Interpolation, where 16 positive wind speeds result in a negative wind speed (own depiction)

Table 5: Summary of interpolation methods with short description. In the following the suffix "c" is added to the abbreviations if wind power correction is applied (own representation)

Method	Abbr.	Description
Nearest Neighbour Method	NN	Wind speeds of the closest MERRA-2 grid point are used
Bilinear Interpolation	BLI	The four surrounding points in a square around the point of interest are first linearly interpolated in one direction (resulting in two points), which are afterwards also interpolated in the other direction
Bicubic Interpolation (not applied)	BCI	16 closest points in square around point of interest are inserted in a cubic equation and with the resulting coefficients calculated by solving the equation system, the wind speed at the point of interest is determined
Inverse Distance Weighting	IDW	Wind speeds are calculated on the point of interest from the four closest grid points, where each of these point is weighted inversely to its distance from the point of interest

2.2.2 Wind speed correction and validation

The first part of the process is the validation and correction of wind speeds. For this, wind speed measurement data of the 478 INMET-stations are used. As the data are available at a height of 10 m above the ground, reanalysis wind speeds need to be extrapolated to the same height, as they are only available at 10 m above disposition height and 50 m above the ground. For extrapolation the power law is applied (see [23]). After vertical interpolation, reanalysis data can be compared to the measured wind speeds. As the hourly and monthly wind speed correction proved to be the best method in [23], only this method is applied for this simulation, but with different variants.

The first method is the same as in [23]. Hourly and monthly sums are compared between INMET and reanalysis data, wherever values are available in the INMET data. Subsequently, correction factors are calculated. Previous analysis [23] revealed some long sequences of the same wind speed in measured data, especially long sequences of 0 m/s wind speed. As this was considered to be unrealistic and therefore points to an error in the data, these long time series were removed in the second method. This was done using the function “rle”, which returns the length of rows of values. A limit of five days is set, so any row of same values with a length of at least 120 hours, is removed from the data. For the third method, monthly means are adapted, which means the mean of selected (without NAs) measured wind speeds are compared to the means of the reanalysis wind speeds for the same time spans and the latter are adapted by multiplying with a factor, so that the mean of reanalysis wind speeds becomes the same as the one of measured data. The fourth method combines the latter two methods and first removes long rows of same values and afterwards adapts the mean of reanalysis wind speeds to the mean of measured wind speeds. Thus, all in all, for each of the wind speed measurement stations that are not discarded due to bad data quality, 16 different sets of wind speed correction factors are calculated, for each of the four interpolation methods and for each of the four correction methods. Table 6 gives a summary of the applied wind speed correction methods.

It was determined, which stations provide wind speed data in sufficient quality by setting three limits: In order to use wind speed measurement data for comparison and wind speed bias correction, at least four “complete” months per month need to be available (that is four Aprils, four Junes et cetera). For a month to be “complete” a limit of 30 days was chosen. This of course excludes all Februaries, which is why it was set to allow one month to have less than 30 days. If a station was selected according to this rule, data from a particular month were only used, if the month had at least 10 days (240 hours).

Table 6: Summary of wind speed correction methods (own representation)

Method	Abbr.	Description
Hourly and monthly wind speed correction	x	Calculation of 288 correction factors, for each hour of the day and each month of the year by division of monthly and hourly sums for simulated and observed wind speeds
Hourly and monthly wind speed correction with removal of long series of 0 m/s wind speed	r	Additionally to the “x” method, before calculating hourly and monthly sums, long rows of 0 m/s wind speed in measured wind speed data are removed, as they are considered errors
Hourly and monthly wind speed correction with adaptation of means	m	The means of simulated and observed wind speeds are calculated and simulated wind speeds are multiplied by the proportion so that means are adjusted, afterwards bias correction is applied as in the “x” method
Hourly and monthly wind speed correction with removal of long series of 0 m/s wind speed and adaptation of means	rm	Both of the above methods are combined, first series of 0 m/s wind speed are removed, then means are adjusted and finally hourly and monthly wind speed correction factors are calculated
No wind speed correction	nINc	No wind speed bias correction is applied

A detailed analysis of the results of wind speed correction and correlations between reanalysis and observed wind speed data can be found in [23]. With the above mentioned limits, only 365 of the 478 wind speed measurement stations were selected for wind speed bias correction when long rows of same wind speed values in measured data were removed, and 384 when this method was not applied (method “x” and method “m”). But also two other criteria were used to determine, where wind speed bias correction shall be applied: The distance of the wind park to the wind speed measurement station may be a maximum of 80 km (which is about 10 km more than the diagonal distance between MERRA points) and the correlation between measured and reanalysis wind speeds after correction needs to be at least 0.5. Figure 2 shows the locations of INMET wind speed measurement stations (red and yellow) and wind parks (blue and black) using wind speed correction method “rm” (removing long rows of 0 m/s wind speed and adapting the mean value) with Nearest Neighbour interpolation. Only the larger red dots are stations, which data can be used for wind speed correction, the smaller red dots are stations where an insufficient amount of data is available and the yellow stations show a too low correlation between measured and reanalysis wind speed data of the remaining 365 stations. 164 stations show a too low correlation to be used for wind speed correction (for the Nearest Neighbour method). The wind speed measurement stations that are actually used for correction are marked with an X. Only the wind parks in blue can be corrected, the ones in black are either too far away (more than 80 km) from the nearest wind speed measurement station or the correlation of wind speeds is too low (less than 50 % after wind speed correction).

2.2.3 Calculation of wind power output and validation

After wind speed correction, wind power generation can be calculated with the help of the power curve. Wind speeds that lie between two values given from the power curve, are linearly interpolated. Wind power generation is calculated starting in 2006, as prior to that no noteworthy capacities were installed (eight wind parks have commissioning dates before 2006 with a capacity of 28.1 MW in total) and there are no data for comparison neither for that period.

After calculating wind power generation, the time series are aggregated per state, compared to recorded time series of wind power generation and used to calculate wind power correction factors. As in [23], monthly bias correction factors are identified, not per subsystem but for each of the eleven states listed in Table 4. The correction factors are calculated for each of the three interpolation methods and each of the four wind speed correction methods as well as without wind speed correction.

2.2.4 Simulating wind power generation for 38 years

When wind speed as well as wind power bias correction factors have been determined, it is possible to simulate wind power generation with different methods: three interpolation methods, four wind speed correction methods and without either wind speed or wind power correction or without any correction. This results in 30 different simulations, which are all compared to observed wind power generation data provided by the National Electric System Operator of Brazil. The comparison is done on daily and monthly scale as well as per state, per subsystem and for all of Brazil. Different statistical parameters (correlation, mean, RMSE, differences) are calculated and compared between different methods, as well as states or subsystems and before and after correction. The plots are generated with "ggplot" [39].

Using the correction factors determined from wind speed and wind power validation, a simulation of nearly 38 years (37 years and 10 months) can be performed. Again, the simulation is conducted for different interpolation and wind speed as well as wind power correction methods, resulting in 30 different time series. For the simulation, currently (per October 2017) installed wind power capacities are assumed. The simulated wind power generation time series are aggregated per state, as saving the power generation of single wind farms would use a considerable amount of storage.

2.2.5 Impact of El Niño and La Niña

The different simulations of 38 years of wind power generation are subsequently compared to indices for El Niño and La Niña. In order to avoid identifying possible seasonal correlations between the indices and wind power generation, wind power generation data as well as the indices for El Niño and La Niña are normalised by dividing by the mean of each month (for example: the wind power generation of January 2005 is divided by the mean of wind power generation in Januaries between 1980 and 2017). For comparison of wind power generation with the three-monthly Oceanic Niño Index, also the three monthly mean of wind power generation is calculated. The impact of El Niño and La Niña are examined separately, too. Furthermore, it is investigated with the help of the "lag" function, whether El Niño and La Niña currents have an impact on wind power generation a few months later than they emerge. Correlations are tested for significance with the R function "cor.test" [40] from the stats package.

Furthermore, it is analysed, whether it is possible to predict future wind power generation with the help of past wind power generation, putting an emphasis on the assessment of the effect of taking into account El Niño and La Niña indices. As an attempt with neural networks did not show good results, the analysis is only conducted with linear regression models. The general linear model is formulated as follows:

$$wp_{lag0} \sim \beta_0 + \sum_{i=1}^{nlag} (\beta_i \cdot SOI_{lagi} + \beta_{nlag+i} \cdot wp_{lag(nlag+i)}) + \sum_{j=1}^{12} \beta_{2 \cdot nlag+j} \cdot month_j$$

The wind power generation in a month (wp_{lag0}) is estimated using the wind power generation of past months ($wp_{lag(nlag+i)}$, where nlag is the number of lags respected), dummy variables of the months (sum at the end) and optionally the lagged Southern Oscillation Index (SOI_{lagi}) – if it is not included in the analysis, this part is removed from the linear model:

$$wp_{lag0} \sim \beta_0 + \sum_{i=1}^{nlag} \beta_i \cdot wp_{lagi} + \sum_{j=1}^{12} \beta_{nlag+j} \cdot month_j$$

From the first part of the analysis it is determined with which lags correlations are highest and these lagged time series of wind power and El Niño and La Niña events are used as an input for the linear model. In the results part it will be examined in detail which lags have the highest impact. For the linear model only part of the data (75 %) are used to train the model, the other 25 % (the last years) are used to test its prediction quality with the help of calculation of adjusted R-squared and RMSEs.

2.2.6 Correlations with water inflows

Similar to the comparison of El Niño and La Niña indices, also a comparison to water inflows into hydropower plants in Brazil is conducted. Data are aggregated monthly as well as for the whole of Brazil for both datasets. The correlation of wind power generation and water inflows is examined, as well as correlations with lagged time series. As for both, wind power generation and water inflows into Brazilian hydropower plants, seasonal behaviour is observed, they are deseasonalised, as mentioned before for El Niño and La Niña indices.

3 Results

This chapter presents some of the results from the simulation of twelve years of wind power (since 2006) and from validation with observed wind power generation. As there are plenty of figures presenting the outcomes, they are attached in the

Appendix and the results section is kept short by displaying results in tables for monthly and daily outcomes on the level of subsystems as well as for all of Brazil. For the abbreviations used in the tables and figures, see the List of Abbreviations. Moreover, the simulated wind power generation time series are compared to El Niño and La Niña indices as well as water inflows into Brazilian hydropower plants.

3.1 Monthly wind power generation per subsystem

For the comparison of synthetic monthly wind power generation with the observed one, several statistical parameters are calculated. First of all, correlations between simulated and observed monthly wind power time series are compared for the North-East and South. All correlations are above 0.95 and differences between various simulation methods or between North-East and South are not significant. They range between 0.983 (for Inverse Distance Weighting without any bias correction) and 0.994 (for three different simulations) in the North-East and from 0.957 (with Inverse Distance Weighting and removal of long rows of 0 m/s wind speed during wind speed correction) to 0.990 (for seven simulations with Bilinear Interpolation or Inverse Distance Weighting and wind power correction) in the South.

In the middle part of the table relative RMSEs are displayed, in order to facilitate comparison between the subsystems. Relative RMSEs in the North-East are between 0.194 and 0.714 and between 0.180 and 0.554 in the South. Wind power correction lowers relative RMSEs in the South of Brazil from between 0.415 and 0.554 to less than 0.253 and thus improves the approximation of simulated to observed wind power generation. In the North-East results are different as no clear improvement of the simulations with wind power correction can be discerned and generally, relative RMSEs are higher than in the South, except for some of the relative RMSEs of not wind power corrected monthly wind power generation. Surprisingly, also the relative RMSE of the simulation with wind power corrected Inverse Distance Weighting combined with wind speed correction with removal of long rows of 0 m/s wind speed, which shows the lowest relative RMSE in the North-East, is lower in the North-East than in the South.

Finally, the means of different simulations are compared to the mean of observed wind power generation in the right part of Table 7. For easier comparison between subsystems, the table features the relative deviations from the observed means. It can be observed, that before wind power correction, the means of the simulations in the South of Brazil in general deviate more from the means of observed monthly wind power generation than in the North-East, except if no wind speed correction is applied. From this, the assumption emerges that in the South wind speed correction has hardly any effect, but affects simulated wind power generation in a more positive way in the North-East, which is due to the fact that only few stations are available for bias correction in the South. Wind power correction also brings improvements in the aspect of adapting the mean monthly wind power generation of the simulations to the mean observed monthly wind power generation. The lowest relative deviations from means of observed wind power generation are at 0.024 in the North-East, when applying Inverse Distance Weighting with wind power and wind speed correction with adaptation of means, and at 0.047 in the South, when the Nearest Neighbour method together with wind power and wind speed correction with no adaptations or at least removal of long rows of 0 m/s wind speed is applied.

Table 7: Correlations, relative RMSEs and relative absolute differences in means of monthly wind power generation time series for the North-East and South of Brazil (own representation)

		correlations					**relative RMSEs**					**rel. abs. difference in means**				
		x	**r**	**m**	**rm**	**nINc**	**x**	**r**	**m**	**rm**	**nINc**	**x**	**r**	**m**	**rm**	**nINc**
North-East	**NN**	0.992	0.992	0.989	0.992	0.984	0.272	0.469	0.334	0.469	0.714	0.109	0.136	0.242	0.136	0.552
	BLI	0.991	0.992	0.988	0.992	0.984	0.280	0.459	0.433	0.328	0.691	0.093	0.123	0.343	0.038	0.530
	IDW	0.991	0.993	0.988	0.992	0.983	0.291	0.248	0.318	0.478	0.692	0.076	0.039	0.212	0.135	0.529
	NNc	0.993	0.994	0.989	0.994	0.991	0.291	0.248	0.333	0.248	0.305	0.028	0.030	0.027	0.030	0.032
	BLIc	0.993	0.993	0.990	0.993	0.991	0.284	0.252	0.317	0.255	0.303	0.032	0.034	0.031	0.027	0.030
	IDWc	0.992	0.994	0.989	0.993	0.991	0.291	0.194	0.336	0.263	0.306	0.032	0.036	0.024	0.034	0.030
South	**NN**	0.988	0.988	0.988	0.988	0.987	0.484	0.484	0.484	0.484	0.516	0.305	0.305	0.246	0.305	0.319
	BLI	0.987	0.987	0.983	0.987	0.987	0.464	0.464	0.415	0.467	0.491	0.290	0.290	0.243	0.291	0.302
	IDW	0.985	0.957	0.982	0.985	0.985	0.531	0.469	0.466	0.513	0.554	0.342	0.211	0.286	0.342	0.353
	NNc	0.989	0.989	0.987	0.989	0.989	0.180	0.180	0.219	0.180	0.190	0.047	0.047	0.051	0.047	0.048
	BLIc	0.990	0.990	0.988	0.990	0.990	0.181	0.181	0.205	0.182	0.191	0.048	0.048	0.049	0.048	0.048
	IDWc	0.990	0.981	0.988	0.990	0.990	0.185	0.253	0.208	0.185	0.194	0.048	0.052	0.050	0.048	0.049

From the results in this section it can be summarised that wind power bias correction usually improves the approximation of simulated to observed wind power generation. In the North-East, it is more difficult to draw precise conclusions, as none of the methods, neither interpolation, wind speed, nor wind power correction, always lead to better results. Some additional results (RMSEs, means and differences) can be found in the

Appendix.

3.2 Monthly wind power generation in Brazil

The results of different simulations are also aggregated for all of Brazil in order to compare them to wind power generation records for the whole country. This chapter aims at the comparison of simulated and observed monthly wind power generation in Brazil and starts with the correlation between the time series (see Table 8). Correlations are in a range of 0.985 (for Inverse Distance Weighting without any bias correction) and 0.994 (for four simulations with at least removal of long rows of 0 m/s wind speed during wind speed correction) and thus, neither between different simulations for all of Brazil nor compared to results from the North-East and South significant differences are identified.

Relative RMSEs between observed and simulated wind power generation for all of Brazil are also compared in the middle part of Table 8. They lie in a range between 0.244 and 0.714 before and between 0.176 and 0.273 after wind power correction, implying that wind power correction lowers RMSEs and thus improves the adaptation of the simulation to observed wind power generation. The lowest relative RMSEs can be observed with wind power correction, when wind speed correction with removal of long rows of 0 m/s wind speed is applied.

Table 8: Correlations, relative RMSEs and relative absolute differences in means of monthly wind power generation time series for Brazil (own representation)

	correlations					relative RMSEs					rel. abs. difference in means				
	x	**r**	**m**	**rm**	**nlNc**	**x**	**r**	**m**	**rm**	**nlNc**	**x**	**r**	**m**	**rm**	**nlNc**
NN	0.992	0.992	0.990	0.992	0.986	0.245	0.332	0.320	0.332	0.667	0.147	0.052	0.337	0.052	0.510
BLI	0.991	0.992	0.989	0.992	0.986	0.246	0.330	0.414	0.244	0.644	0.132	0.045	0.324	0.026	0.489
IDW	0.991	0.991	0.989	0.991	0.985	0.252	0.268	0.312	0.338	0.656	0.127	0.005	0.226	0.045	0.498
NNc	0.993	0.994	0.991	0.994	0.991	0.245	0.213	0.269	0.213	0.255	0.034	0.036	0.034	0.036	0.037
BLIc	0.993	0.994	0.991	0.993	0.992	0.236	0.213	0.262	0.215	0.250	0.038	0.039	0.037	0.033	0.036
IDWc	0.993	0.994	0.991	0.993	0.992	0.241	0.176	0.273	0.220	0.252	0.037	0.041	0.031	0.039	0.036

The right part of Table 8 compares the relative deviations of means of observed and simulated wind power generation in Brazil. The lowest value occurs with not wind power corrected data when Inverse Distance Weighting is applied with wind speed correction when long rows of 0 m/s wind speed are removed. It has to be pointed out that not for all simulations wind power correction improves results: Inverse Distance Weighting with removal of long rows of 0 m/s wind speed and also Bilinear Interpolation with wind speed correction combined with mean approximation as well as removal of long rows of 0 m/s wind speed during wind speed correction show lower relative deviations in means before wind power correction than after. When no wind power correction is applied, wind speed correction usually leads to lower relative differences in means, but this is not always the case if wind power correction is applied. Compared to results from the analysis in the subsystems, the relative deviations in means mostly are lower than in the South, but usually higher than in the North-East.

From the analysis of simulated monthly wind power generation for all of Brazil it can be concluded, that in most cases wind power correction improves results, but not always. The most significant improvement (at least in the comparison of relative RMSEs and relative absolute differences in means) can be seen when not applying wind speed correction beforehand, which shows, that probably at least one bias correction method should be applied. Wind speed correction with removal of long rows of 0 m/s wind speed seems to work best for most of the measures regarded, and for the interpolation method no preferential method can be detected.

3.3 Daily wind power generation per subsystem

This section presents results from the analysis of wind power generation in the North-East and in the South of Brazil, but on a daily basis. The left part of Table 9 shows correlations between simulated and observed daily wind power generation in the North-East and South of Brazil. As in the monthly comparison, correlations are above 0.9, with only one exception (0.896 with Inverse Distance Weighting and only wind speed correction combined with removal of long rows of 0 m/s wind speed in the South). In the North-East they range from 0.977 (with Inverse Distance Weighting and no bias correction) to 0.990 (Inverse Distance Weighting with wind power and wind speed correction with removal of long rows of 0 m/s wind speed) and are thus slightly but not significantly higher than in the South, where correlations are between 0.896 and 0.954 (for six different simulations).

The relative RMSEs of daily wind power generation per subsystem are compared in the middle part of Table 9. Relative RMSEs are in a range of 0.245 to 0.907 in the North-East and are in the range of 0.438 to 0.773 in the South. On a daily basis, the RMSEs are mostly smaller in the North-East than in the South, contrary to the results from the monthly analysis. In the North-East as well as in the South a clear improvement by wind power bias correction can be perceived, as relative RMSEs are smaller after correction: In the North-East relative RMSEs are reduced from a range between 0.265 and 0.907 to between 0.245 and 0.315, in the South before wind power correction relative RMSEs lie between 0.653 and 0.793 and after correction they are in a range of 0.438 to 0.591. In the North-East, the highest reductions in relative RMSEs by wind power correction can be observed when no wind speed correction is applied – before wind power correction relative RMSEs are significantly higher with the simulation without wind speed correction. In the South results are similar for different wind speed correction methods and without wind speed correction, which supports the assumption made before that wind speed correction does not have a big impact on results in the South but is more important in the North-East.

Table 9: Correlations, relative RMSEs and relative absolute differences in means of daily wind power generation time series for the North-East and South of Brazil (own representation)

		correlations					relative RMSEs					rel. abs. difference in means				
		x	r	m	rm	nINc	x	r	m	rm	nINc	x	r	m	rm	nINc
North-East	NN	0.985	0.986	0.983	0.986	0.979	0.321	0.409	0.452	0.409	0.907	0.183	0.078	0.325	0.078	0.655
	BLI	0.985	0.985	0.983	0.985	0.978	0.318	0.401	0.585	0.301	0.884	0.166	0.064	0.433	0.027	0.632
	IDW	0.984	0.988	0.982	0.985	0.977	0.316	0.265	0.420	0.417	0.886	0.148	0.108	0.292	0.077	0.632
	NNc	0.988	0.988	0.985	0.988	0.985	0.282	0.264	0.312	0.264	0.298	0.097	0.099	0.096	0.099	0.101
	BLIc	0.987	0.987	0.986	0.987	0.985	0.281	0.270	0.302	0.270	0.298	0.101	0.104	0.100	0.095	0.099
	IDWc	0.987	0.990	0.985	0.987	0.985	0.288	0.245	0.315	0.278	0.302	0.101	0.106	0.092	0.103	0.098
South	NN	0.952	0.952	0.941	0.952	0.951	0.721	0.721	0.665	0.721	0.748	0.343	0.343	0.283	0.343	0.358
	BLI	0.949	0.949	0.931	0.949	0.949	0.707	0.707	0.653	0.710	0.729	0.328	0.328	0.280	0.329	0.340
	IDW	0.946	0.896	0.930	0.946	0.946	0.773	0.667	0.697	0.773	0.793	0.382	0.187	0.324	0.382	0.393
	NNc	0.954	0.954	0.942	0.954	0.953	0.438	0.438	0.487	0.438	0.443	0.078	0.078	0.082	0.078	0.079
	BLIc	0.954	0.954	0.936	0.954	0.953	0.446	0.446	0.496	0.446	0.450	0.079	0.079	0.080	0.079	0.079
	IDWc	0.952	0.917	0.936	0.952	0.952	0.454	0.591	0.500	0.454	0.458	0.079	0.084	0.081	0.079	0.080

Relative absolute differences in means of observed and simulated daily wind power generation are displayed in the right part of Table 9. In the North-East, relative deviations in means of wind power generation are usually smaller than in the South before wind power correction: In the North-East they range from 0.027 to 0.325 (with four exceptions for simulations without any bias correction and Bilinear Interpolation with approximation of means), whereas in the South they lie between 0.187 and 0.393 (with the exception of Bilinear Interpolation with approximation of means during wind speed correction). In the North-East wind power correction does not always lead to lower relative deviations in means of simulated and observed wind power generation, but is significant when no wind speed correction is applied: Then, relative mean differences in means are reduced from around 0.6 to around 0.1. For some simulations, however, when wind speed correction with at least removal of long rows of 0 m/s wind speed is applied, relative deviations in means are increased. In the South, wind speed correction leads to only slightly lower relative deviations in means, but contrary to the North-East, wind power correction reduces relative differences in means of simulated and observed daily wind power generation significantly: Before wind power correction most relative differences lie between 0.3 and 0.4 and after wind power correction they are all below 0.1.

To sum up, in the North-East correlations and relative RMSEs are better compared to the South, but relative absolute differences in means are higher after wind power correction than in the South. Wind speed correction leads to improvement in terms of correlations, relative RMSEs as well as relative absolute differences in means in the North-East, but has little to no effect in the South. However, due to only few stations available for wind speed correction, wind power correction is more effective in the South, as in the North-East results are significantly improved only when no wind speed correction is applied. In the

Appendix some additional results can be found.

3.4 Daily wind power generation in Brazil

In this chapter, the results from the analysis of daily wind power generation in Brazil are presented. The left part of Table 10 shows the correlations of observed and simulated daily wind power generation in Brazil. Correlations range between 0.980 (Inverse Distance Weighting without any bias correction) and 0.989 (Inverse Distance Weighting with both corrections combined with removal of long rows of 0 m/s wind speed), showing no significant difference to monthly correlations or to those in the South or North-East or between different methods in Brazil.

The comparison of relative RMSEs (see middle part of Table 10) shows that wind power correction always improves the simulation (decreases relative RMSEs): Before wind power correction relative RMSEs range from 0.260 to 0.841 and after correction they are reduced to between 0.242 and 0.271. The largest reductions in relative RMSEs are observed if no wind speed correction is applied (from more than 0.8 to less than 0.3), leading to the conclusion that at least one type of bias correction method should be applied. After wind power correction, the Nearest Neighbour method results in the lowest relative RMSEs when at least long rows of 0 m/s wind speed are removed. However, no significant differences in relative RMSEs can be observed after wind speed correction. Relative RMSEs are lower than in the comparison of daily wind power generation in the South, but about in the same range as in the North-East. Compared to the relative RMSEs of monthly wind power generation, they are in a similar but slightly lower range.

Another parameter which is used to compare daily wind power generation in Brazil are the relative deviations in means of observed and simulated time series, which are represented in the right part of Table 10. Before wind power correction, they lie between 0.006 and 0.590 and after they range from 0.094 to 0.106, so some differences in means are reduced, but others are increased. For simulations without wind speed correction, relative deviations in means of simulated and observed wind power are always significantly reduced, from around 0.6 to around 0.1. However, if at least removal of long rows of 0 m/s wind speed is applied during wind speed correction, relative differences in means are increased by wind power correction. Therefore, wind power correction is not always recommendable when comparing the resulting wind power time series by this measure. Compared to relative absolute differences in the daily analysis of the South and North-East and also in the monthly analysis of Brazil, no clear trend can be identified, as some values are lower while others are higher.

Table 10: Correlations, relative RMSEs and relative absolute differences in means of daily wind power generation time series in Brazil (own representation)

	correlations					relative RMSEs					rel. abs. difference in means				
	x	r	m	rm	nlNc	x	r	m	rm	nlNc	x	r	m	rm	nlNc
NN	0.986	0.985	0.985	0.985	0.982	0.331	0.301	0.443	0.301	0.841	0.218	0.006	0.319	0.006	0.603
BLI	0.985	0.984	0.984	0.985	0.981	0.322	0.301	0.557	0.267	0.818	0.201	0.014	0.406	0.089	0.581
IDW	0.984	0.987	0.983	0.984	0.980	0.323	0.260	0.423	0.309	0.831	0.197	0.056	0.301	0.014	0.590
NNc	0.988	0.988	0.987	0.988	0.987	0.251	0.242	0.268	0.242	0.263	0.098	0.100	0.097	0.100	0.101
BLIc	0.988	0.988	0.987	0.988	0.986	0.250	0.246	0.266	0.246	0.263	0.101	0.103	0.101	0.096	0.099
IDWc	0.988	0.989	0.986	0.988	0.986	0.255	0.243	0.271	0.251	0.265	0.101	0.106	0.094	0.103	0.099

From the results in this section it can be concluded that wind power correction sometimes leads to better results but not in all cases, also depending on the parameter regarded. Correlations are always increased and relative RMSEs always reduced, but from the analysis of relative absolute differences in means no clear trend is observed. Wind speed correction, however, always leads to better results in terms of correlations, relative RMSEs as well as relative absolute differences in means, but for correlations the differences with and without wind speed correction are only marginal.

3.5 Evaluation and selection of the best method for simulation

As the analysis of thirty different simulations on different spatial and temporal resolutions delivers many results and these are hard to compare to each other, the (normalised) calculated statistical parameters are ranked for easier comparison and summed up for each method (sum of ranks of correlations, RMSEs and relative absolute difference in means per interpolation and wind speed correction method). The results of these rank sums are summarised in

Table 11. It can be seen clearly, that the Nearest Neighbour method with wind power correction in most cases leads to the best results regarding rank sums for correlations, relative deviations in means as well as relative RMSEs, except for the monthly comparison on the level of states as well as in Brazil on a daily basis, where Bilinear Interpolation yields better results. In two other cases also another interpolation method yields low rank sums: in the monthly comparison of wind power generation in the South, with Bilinear Interpolation with wind power correction with optional removal of long rows of 0 m/s wind speed. Regarding wind speed correction methods, also a clear trend is visible: For all spatial and temporal levels of analysis, methods where at least long rows of 0 m/s wind speed are removed during wind speed correction, and optionally also the mean is adapted, the lowest rank sums are obtained. In two cases (in the daily and monthly comparison in the South) also wind speed

correction without adaptation performs as good as with removal of long rows of 0 m/s wind speed. Nevertheless, as with the “r” and “rm” method, always the best results are achieved in terms of rank sums, finally the method with removal of long rows of 0 m/s wind speed is selected, as it is less complex and thus needs less computational time, but performs as well as the method where also the mean is adapted.

Figure 5 shows the time series of monthly wind power generation for the selected method, before and after wind power correction. It can be observed, that before 2014 not wind power corrected time series are closer to observed wind power correction, but in more recent years the bias corrected simulation fits observed wind power generation better.

Table 11: Rank-sums of all simulations for evaluation of the best method. The lowest sum for each spatial and temporal level is marked in green (own representation)

		Brazil					North-East					South					States				
		x	r	m	rm	nINc	x	r	m	rm	nINc	x	r	m	rm	nINc	x	r	m	rm	nINc
daily	NN	62	37	70	37	88	58	37	77	37	88	56	56	57	56	67	382	380	396	380	450
	BLI	63	46	77	39	85	56	37	79	28	86	54	54	58	58	63	395	400	359	362	437
	IDW	68	25	77	48	88	64	27	77	41	87	74	62	64	74	78	465	471	413	463	468
	NNc	19	16	35	16	38	20	16	38	16	40	3	3	47	3	18	195	176	267	176	205
	BLIc	26	27	39	14	37	31	30	40	19	37	11	11	47	10	23	193	176	208	162	173
	IDWc	28	25	38	29	38	32	22	37	31	38	26	57	50	26	31	203	245	239	196	188
monthly	NN	45	54	74	54	88	38	58	70	58	88	58	59	53	59	72	697	694	713	694	756
	BLI	52	52	79	21	84	45	55	76	45	85	55	55	60	60	67	675	673	702	621	747
	IDW	53	36	72	61	88	47	25	69	59	87	80	67	66	80	84	755	783	787	719	784
	NNc	21	12	41	12	45	19	10	46	10	42	12	10	46	10	24	166	151	260	151	199
	BLIc	26	21	45	14	31	26	23	48	12	37	10	10	35	11	21	130	124	190	144	146
	IDWc	26	19	40	26	32	34	17	46	24	38	16	59	38	16	24	163	215	256	158	165

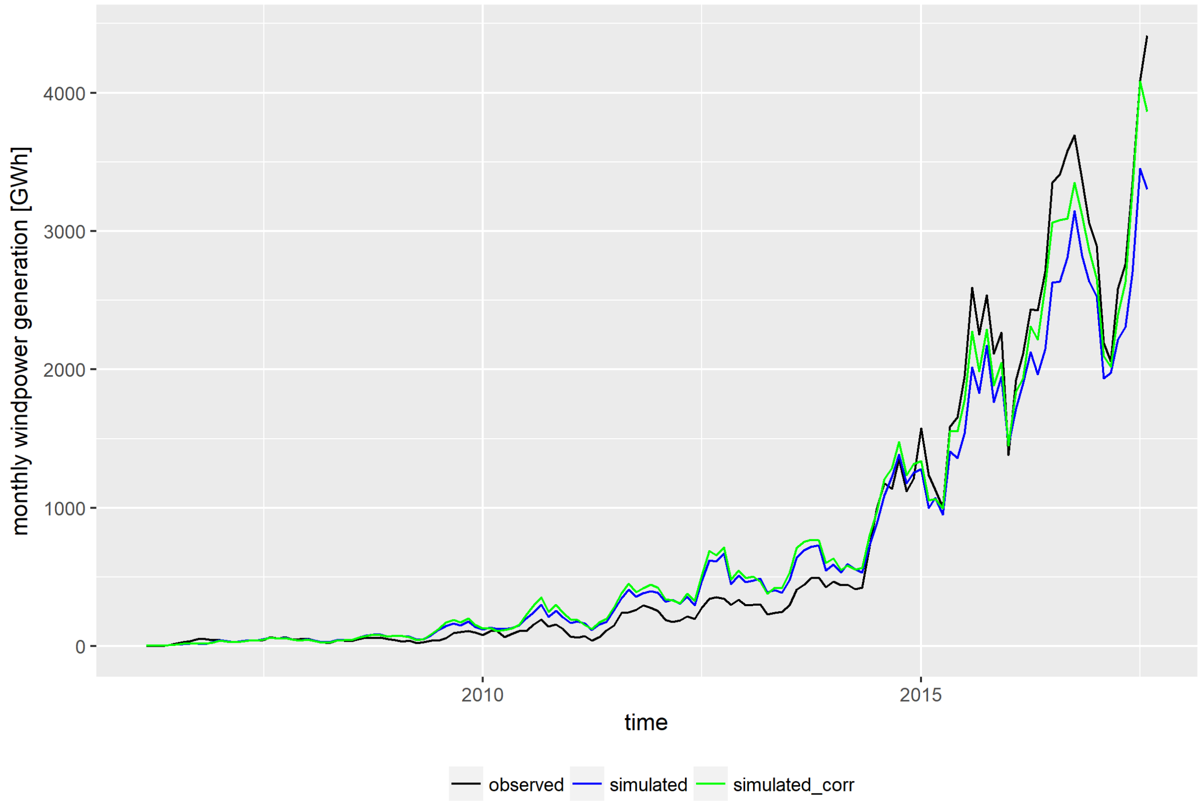

Figure 5: Comparison of observed and simulated monthly wind power generation in Brazil for the method which is evaluated the best before and after wind power correction (own depiction)

3.6 Impact of El Niño and La Niña

This chapter presents results from the examination of correlations between fluctuations of wind power generation and El Niño and La Niña currents. As the wind power corrected Nearest Neighbour method combined with removal of long rows of 0 m/s wind speed was identified as the best in previous sections, only this method is used for analysis here. For the comparison, theoretical wind power over a time span of nearly 38 years was calculated with the wind power capacity installed at the moment. As it is known that wind power generation varies with the season (see Figure 6: lowest wind power generation is observed from February to April and highest wind power generation between August and October), it was considered necessary to deseasonalise wind power generation, in order to eliminate correlation with seasonal fluctuations. However, seasonalities were not only found in wind power generation, but also in the indices for El Niño and La Niña (see Figure 6), despite the events not occurring seasonally but throughout several years (see Figure 7). Therefore, also those indices were deseasonalised before comparing them to simulated wind power generation.

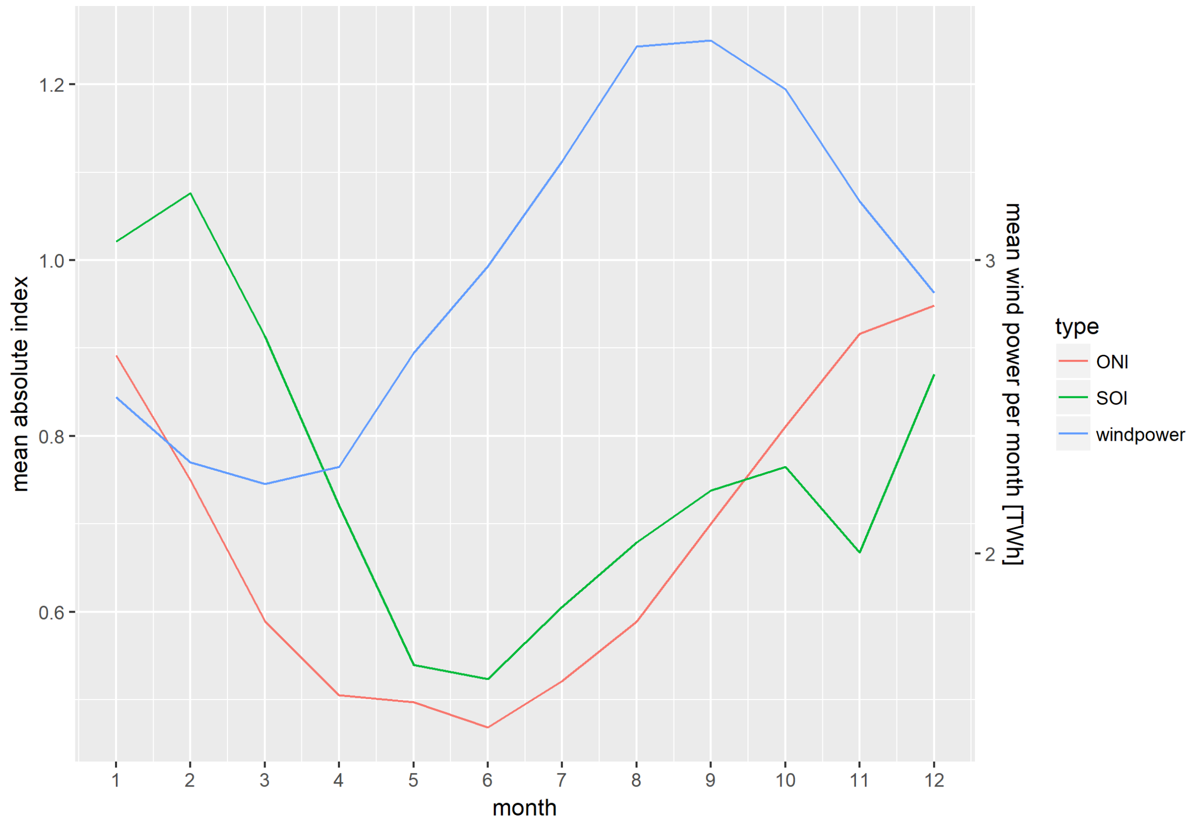

Figure 6: Mean of monthly wind power generation in Brazil and mean absolute monthly El Niño and La Niña indices (own depiction)

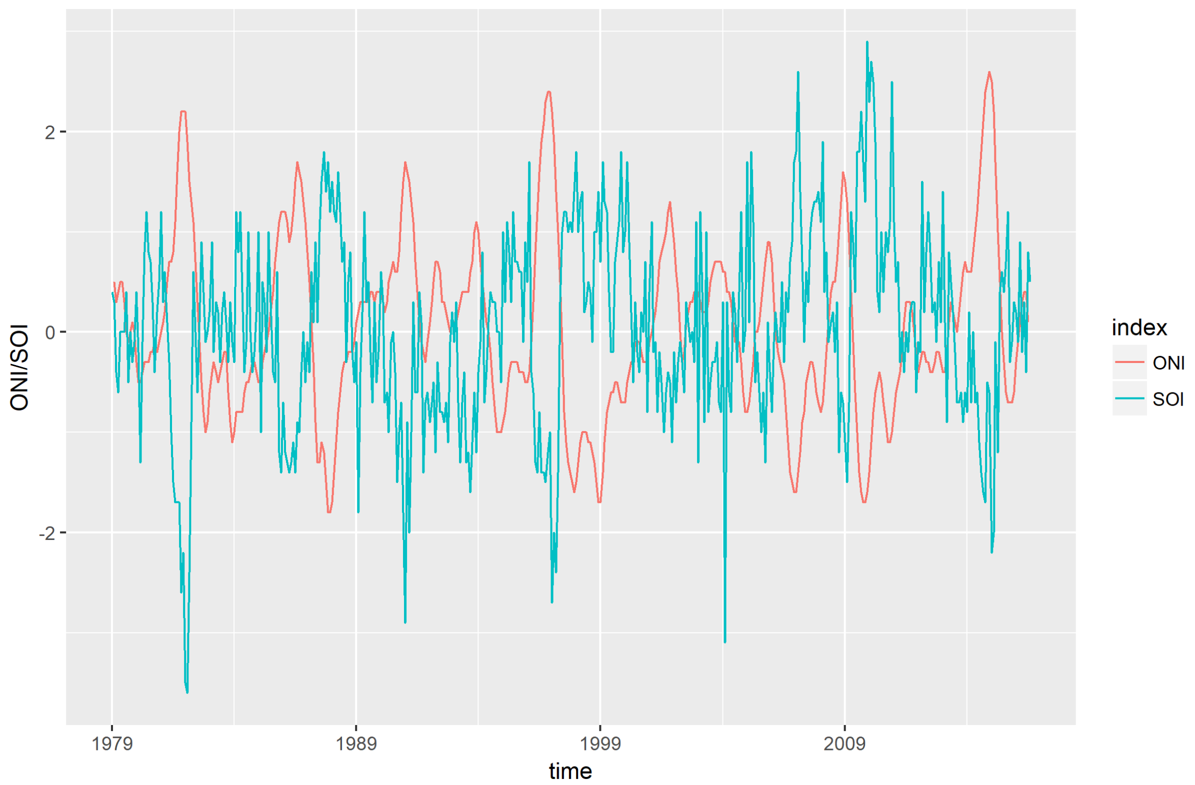

Figure 7: Oceanic Niño Index and Southern Oscillation Index since 1980 (own depiction with data from [30] and [32])

Monthly and yearly correlations of wind power generation and El Niño and La Niña time series were calculated for the one monthly (SOI) and three monthly indices (ONI) in two subsystems (North-East and South) and for all of Brazil (Table 12). In the comparison with the Oceanic Niño Index absolute correlations lie between 0.1 % and 7 % with lags of seven and three months, respectively. With the Southern Oscillation Index absolute correlations are in a similar range of 0.003 to 0.059. However, none of these correlations is significant. In the South the maximum absolute correlations are in a similar range, at 0.053 with the SOI and at 0.069 with the ONI with lags of four and five months, respectively. In the North-East slightly higher absolute correlations of 0.86 are obtained for the ONI with a lag of three months. However, neither these correlations are significant, when tested with "cor.test". The analysis of yearly time series shows even higher but neither significant correlations for Brazil (9.8 %) and the South (22.2 %).

Table 12: Yearly and monthly correlations of El Niño and La Niña events, represented by the Oceanic Niño index (ONI) and the Southern Oscillation index (SOI) with wind power generation in Brazil (B) and its North-East (NE) and South (S) with lags between 0 and 12 months. The highest absolute correlations are highlighted in green. None of the correlations are significant (own representation)

	both						yearly		
	ONI			SOI			SOI		
lag	B	NE	S	B	NE	S	B	NE	S
0	-0.027	-0.037	0.037	0.012	0.004	0.033	0.019	-0.013	0.201
1	-0.035	-0.047	0.042	-0.028	-0.026	-0.014	0.061	0.029	0.211
2	-0.056	-0.069	0.042	0.052	0.059	-0.015	0.059	0.036	0.153
3	-0.070	-0.086	0.050	0.059	0.053	0.036	0.063	0.041	0.153
4	-0.066	-0.085	0.065	-0.027	-0.014	-0.053	0.052	0.024	0.178
5	-0.049	-0.068	0.069	0.019	0.021	-0.002	0.043	0.015	0.179
6	-0.022	-0.036	0.052	-0.029	-0.033	0.009	0.037	0.010	0.172
7	-0.001	-0.005	0.019	-0.053	-0.066	0.034	0.025	-0.001	0.164
8	0.014	0.019	-0.019	-0.034	-0.046	0.037	0.033	0.007	0.159
9	0.015	0.027	-0.048	0.031	0.045	-0.045	0.078	0.050	0.186
10	0.011	0.024	-0.054	0.003	0.006	-0.014	0.098	0.068	0.205
11	0.011	0.020	-0.032	0.045	0.053	-0.021	0.067	0.037	0.197
12	0.011	0.010	0.006	0.015	0.027	-0.040	0.052	0.017	0.222

As no significant correlations could be identified when examining El Niño and La Niña events together, they are also investigated separately in Table 13. This results in higher and also significant correlations with one exception (the correlation with El Niño in the SOI events in the South is not significant). The highest correlations with La Niña Events with the ONI are at 16.1 %, 26.4 % and 29.8 % in the South, in Brazil and in the North-East with lags of six, three and four months respectively. When the SOI is examined, highest absolute correlations are lower between 13.3 % (South) and 18.8 % (Brazil) and occur with lags of two to three months. Thus, it can be expected that after three to four months after a La Niña event wind power generation will decrease in the North-East and in Brazil, whereas in the South La Niña's impact is not conclusive.

Equally to the analysis of La Niña, also the impacts of El Niño events on wind power generation in Brazil are examined in the right part of Table 13. With El Niño higher correlations than for La Niña are obtained in Brazil and in the South, for the ONI as well as the SOI. All of the correlations are positive, meaning an increase of wind power generation can be expected after an El Niño event. However, results are not definite regarding the lag with which El Niño might have an effect on wind power generation, as highest effects are observed after two, six or seven months.

Table 13: Monthly correlations of separate El Niño and La Niña events, represented by the Oceanic Niño index (ONI) and the Southern Oscillation index (SOI) with wind power generation in Brazil (B) and its North-East (NE) and South (S) with lags between 0 and 12 months. The highest absolute correlations are highlighted in green. Significant correlations (level of significance p < 0.05) are printed in bold (own representation)

	La Niña						El Niño					
	ONI			SOI			ONI			SOI		
lag	B	NE	S	B	NE	S	B	NE	S	B	NE	S
0	-0.060	-0.079	0.068	0.017	0.015	0.013	-0.009	-0.013	0.014	0.024	0.012	0.053
1	**-0.150**	**-0.165**	0.030	**-0.179**	**-0.161**	-0.116	0.055	0.045	0.052	**0.218**	**0.214**	0.053
2	**-0.226**	**-0.244**	0.026	**-0.138**	-0.111	**-0.133**	0.094	0.081	0.079	**0.226**	**0.227**	0.033
3	**-0.264**	**-0.292**	0.062	**-0.188**	**-0.176**	-0.100	0.128	0.110	0.107	0.132	0.107	0.116
4	**-0.260**	**-0.298**	0.106	**-0.162**	**-0.156**	-0.070	**0.170**	**0.143**	**0.149**	0.054	0.062	-0.023
5	**-0.227**	**-0.272**	**0.143**	-0.042	-0.056	0.038	**0.233**	**0.200**	**0.187**	-0.015	-0.045	0.118
6	**-0.164**	**-0.209**	**0.161**	-0.045	-0.049	0.001	**0.285**	**0.252**	**0.199**	0.002	-0.001	0.011
7	-0.106	**-0.148**	**0.159**	-0.013	-0.034	0.069	**0.311**	**0.281**	**0.193**	0.018	-0.010	0.111
8	-0.075	-0.112	**0.143**	-0.034	-0.054	0.060	**0.305**	**0.279**	**0.174**	-0.030	-0.037	0.020
9	-0.063	-0.097	0.133	0.010	0.027	-0.056	**0.266**	**0.246**	**0.145**	-0.027	-0.020	-0.031
10	-0.066	-0.099	0.124	0.003	-0.024	0.093	**0.199**	**0.185**	0.101	-0.088	-0.098	0.022
11	-0.048	-0.079	0.119	0.017	0.008	0.035	0.122	0.114	0.061	-0.063	-0.063	-0.010
12	-0.020	-0.049	0.116	0.031	0.032	0.004	0.061	0.052	0.052	-0.040	-0.060	0.073

For the analysis of correlations with thirteen different states in Brazil, not all results are presented, but only the highest correlations and the lags when these occur are summarised in Table 14. When analysing the correlations with the Southern Oscillation Index for El Niño and La Niña events together, absolute correlations between 3.4 % and 23.2 % are obtained, however only those above 10 % are significant. When considering only El Niño events, absolute correlations are higher for most states, up to 36.5 %, which is also the case with La Niña events only, where a maximum correlation of 39.9 % is observed. For the comparison with the ONI, correlations are lower in most cases than with the SOI, but are also usually higher if either El Niño or La Niña events are considered instead of both together. In the North-East mainly negative correlations with both or only El Niño events can be observed, which indicates that after an El Niño event lower wind power generation can be expected (except in Sergipe). On the other side, after a La Niña event wind power generation is more likely to increase, indicated by positive correlations, especially in the North-East. In the South and South-East results are not as conclusive, as fewer states are available for comparison and some correlations are negative, whereas others are positive. Considering lags, it is difficult to draw conclusions, as highest absolute correlations occur with no up to one year lags.

Table 14: Highest absolute monthly correlations and their lags of El Niño and La Niña events, represented by the Oceanic Niño index (ONI) and the Southern Oscillation index (SOI) with wind power generation in thirteen states in Brazil (BA…Bahia, CE…Ceará, MA…Maranhão, MG…Minas Gerais, PB…Paraíba, PR…Paraná, PE…Pernambuco, PI…Piaui, RJ…Rio de Janeiro, RN…Rio Grande do Norte, RS…Rio Grande do Sul, SC…Santa Catarina, SE…Sergipe) and three states (NE…North-East, S…South, SE…South-East). Significant correlations are printed in bold (own representation)

			NE	NE	NE	SE	NE	S	NE	NE	SE	NE	S	S	NE
			BA	CE	MA	MG	PB	PR	PE	PI	RJ	RN	RS	SC	SE
both	SOI	cor	**-0.117**	-0.034	**-0.124**	**-0.221**	**±0.105**	**0.232**	**0.107**	**-0.113**	0.054	-0.077	0.060	**-0.177**	0.090
		lag	**3**	5	**12**	**9**	**12, 3**	**0**	**9**	**12**	8	4	6	**10**	9
	ONI	cor	-0.021	**-0.017**	**0.127**	**-0.142**	0.057	**0.122**	0.067	**-0.101**	**0.097**	0.083	-0.052	0.087	0.063
		lag	9	**7**	**5**	**4**	11	**3**	2	**8**	**9**	11	4	3	2
El Niño	SOI	cor	**-0.258**	**-0.333**	**-0.350**	-0.073	**-0.185**	-0.106	**-0.157**	**-0.365**	0.110	**-0.245**	**0.161**	0.125	**0.196**
		lag	**3**	**5**	**5**	9	**4**	10	**2**	**3**	12	**4**	**6**	7	**0**
	ONI	cor	**-0.186**	**-0.220**	**-0.173**	**-0.138**	-0.111	-0.082	-0.056	**-0.167**	0.128	**-0.140**	**-0.137**	0.086	0.024
		lag	**1**	**3**	**3**	**3**	3	9	2	**3**	11	**3**	**2**	11	4
La Niña	SOI	cor	**0.269**	**0.233**	**0.311**	**0.249**	**0.325**	**0.399**	**0.216**	**0.249**	-0.109	**0.277**	**0.210**	**0.164**	-0.116
		lag	**9**	**7**	**2**	**9**	**8**	**0**	**8**	**7**	0	**7**	**6**	**0**	3
	ONI	cor	**0.170**	**0.241**	**0.315**	-0.105	0.105	**0.313**	**0.176**	**0.208**	0.089	**0.219**	0.120	**0.175**	**-0.141**
		lag	**2**	**2**	**2**	9	1	**3**	**2**	**2**	4	**1**	5, 7	**3**	**5**

The results from this analysis are used to construct a linear regression model for wind power forecast. As highest correlations were found after two to seven months, these are tested in the model. Table 15 shows results of this analysis: Adjusted R-squared range from 0.389 (Rio Grande do Sul) to 0.881 (Ceará), excluding Minas Gerais which has a very low adjusted R-squared of only 0.086. For most parts of Brazil, adjusted R-squared are higher and MSEs lower, if the SOI is included in the prediction. Only in the South, in Rio de Janeiro and in Rio Grande do Sul better results (lower MSEs and higher adjusted R-squared) are obtained without considering El Niño and La Niña. It is notable, that including El Niño in the analysis improves the results of prediction especially in the northern states. In the right part of the table, green coloured fields indicate that El Niño and La Niña events with this lag (two, four or six months) are significant. In eleven of sixteen cases the impact of the ENSO is significant with a lag of two months. Once more, in the South, less significant influence of the SOI is perceived. Also with lags of four months significance is observed, but only for three cases, and with a lag of six months only in Paraná impacts can be seen.

Table 15: Adjusted R-squared (adj R^2) and Mean Square Errors (MSE) from the analysis of linear models with lags of two, four and six months, with and without considering El Niño and La Niña events for Brazil, the North-East and the South and thirteen states. The higher correlations and lower MSEs, as well as the significant parameters are highlighted in green. (own representation)

	with Niño		without Niño		significance Niño		
	adj R^2	MSE [GWh]	adj R^2	MSE [GWh]	lag2	lag4	lag6
Brazil	0.792	94708.690	0.773	101324.240			
North-East	0.774	79352.185	0.753	88014.513			
South	0.408	6868.649	0.417	6865.726			
Bahia	0.680	4888.440	0.673	5250.066			
Ceará	0.881	3058.849	0.851	3603.748			
Maranhão	0.869	5.793	0.810	6.906			
Minas Gerais	0.086	0.002	0.065	0.002			
Paraíba	0.757	3.320	0.746	3.595			
Paraná	0.729	0.002	0.714	0.001			
Pernambuco	0.545	652.306	0.504	685.171			
Piaui	0.878	3643.566	0.867	3680.166			
Rio de Janeiro	0.552	0.396	0.554	0.387			
Rio Grande do Norte	0.727	10525.797	0.699	11539.912			
Rio Grande do Sul	0.389	6131.054	0.398	6114.658			
Santa Catarina	0.669	53.461	0.663	51.087			
Sergipe	0.531	1.036	0.519	1.032			

As the lag of two months resulted in most significant impacts, also a linear model with only this lag is tested. The results can be seen in Table 16: When only including ENSO time series with a lag of two months, in even more cases better results are obtained with than without considering El Niño and La Niña events. Only in three cases the adjusted R-squared, which ranges between 0.395 and 0.878 (except Minas Gerais which is at 0.088), are lower and also only in three cases MSEs are higher when the SOI is used in the analysis. However, including only the lag of two months in the analysis does not affect which impacts of the ENSO are significant. Compared to the results of the linear model which also includes lags of four and six months, adjusted R-squared are in a similar range and MSEs are lower (except in Bahia, Minas Gerais, Paraíba, Paraná and Pernambuco).

Table 16: Adjusted R-squared (adj R^2) and Mean Square Errors (MSE) from the analysis of linear models with lags of two months, with and without considering El Niño and La Niña events for Brazil, the North-East and the South and thirteen states. The higher correlations and lower MSEs, as well as the significant parameters are highlighted in green. (own representation)

	with Niño		without Niño		significance Niño
	adj R^2	MSE [GWh]	adj R^2	MSE [GWh]	lag2
Brazil	0.792	92885.666	0.773	101324.240	
North-East	0.775	78343.954	0.753	88014.513	
South	0.414	6830.902	0.417	6865.726	
Bahia	0.677	4892.366	0.673	5250.066	
Ceará	0.880	2941.693	0.851	3603.748	
Maranhão	0.864	5.637	0.810	6.906	
Minas Gerais	0.088	0.002	0.065	0.002	
Paraíba	0.758	3.336	0.746	3.595	
Paraná	0.727	0.002	0.714	0.001	
Pernambuco	0.521	665.128	0.504	685.171	
Piaui	0.878	3580.204	0.867	3680.166	
Rio de Janeiro	0.555	0.390	0.554	0.387	
Rio Grande do Norte	0.727	10335.868	0.699	11539.912	
Rio Grande do Sul	0.395	6092.036	0.398	6114.658	
Santa Catarina	0.670	52.930	0.663	51.087	
Sergipe	0.525	1.025	0.519	1.032	

It is also tested, if other lags of one, three, five and seven months show a different impact on the linear model. The adjusted R-squared, MSEs and significant parameters are presented in Table 17. For this case even more adjusted R-squared (fifteen of sixteen) are higher, if ENSO is included in the linear model. Of the MSEs eleven are lower if including El Niño and La Niña events in the prediction, which is more than with lags of two, four and six months (Table 15), but less than with a lag of only two months (Table 14). Although adjusted R-squared are in a similar range and MSEs in most cases are slightly higher than with lags of two, four or six months, lags of one, three, five or seven months are not very likely to improve results, as hardly any of them are significant: Only a lag of one month is significant in the South and in Rio Grande do Sul, and a lag of three months is significant in Maranhão, Paraná and Santa Catarina. Thus, it can be expected, that only few months (one to three months, most likely after two months) after an El Niño or La Niña event, this may have an effect on wind power generation. This is consistent with some of the results from the separate analysis of El Niño and La Niña events (Table 13 and Table 14), where for several cases the highest significant correlations were achieved with lags of one to three months.

Table 17: Adjusted R-squared (adj R^2) and Mean Square Errors (MSE) from the analysis of linear models with lags of one, three, five and seven months, with and without considering El Niño and La Niña events for Brazil, the North-East and the South and thirteen states. The higher correlations and lower MSEs, as well as the significant parameters are highlighted in green. (own representation)

	with Niño		without Niño		significance Niño			
	adj R^2	MSE [GWh]	adj R^2	MSE [GWh]	lag1	lag3	lag5	lag7
Brazil	0.791	89189.331	0.773	101324.240				
North-East	0.769	75779.930	0.753	88014.513				
South	0.431	7077.554	0.417	6865.726				
Bahia	0.667	5090.837	0.673	5250.066				
Ceará	0.879	2831.097	0.851	3603.748				
Maranhão	0.877	5.071	0.810	6.906				
Minas Gerais	0.087	0.002	0.065	0.002				
Paraíba	0.765	3.194	0.746	3.595				
Paraná	0.729	0.001	0.714	0.001				
Pernambuco	0.517	599.209	0.504	685.171				
Piaui	0.874	3343.621	0.867	3680.166				
Rio de Janeiro	0.555	0.365	0.554	0.387				
Rio Grande do Norte	0.740	9497.328	0.699	11539.912				
Rio Grande do Sul	0.411	6289.008	0.398	6114.658				
Santa Catarina	0.671	51.986	0.663	51.087				
Sergipe	0.540	1.096	0.519	1.032				

Finally, also only El Niño events and their suitability for helping in the prediction of wind power generation are studied. The results are shown in Table 18: Still in most cases, including El Niño events into the analysis lead to higher adjusted R-squared and lower MSEs. Fewer cases of significant parameters occur: The lag which shows most significance still is the one with two months, however, when only El Niño events are considered, only for five regions (North-East, Paraíba, Paraná, Pernambuco, Rio Grande do Norte) significant levels of impact on the wind power generation are determined. Lags of four or six months are less important.

Table 18: Adjusted R-squared (adj R^2) and Mean Square Errors (MSE) from the analysis of linear models with lags of two, four and six months, with and without considering El Niño events for Brazil, the North-East and the South and thirteen states. The higher correlations and lower MSEs, as well as the significant parameters are highlighted in green. (own representation)

	with Niño		without Niño		significance Niño		
	adj R²	MSE [GWh]	adj R²	MSE [GWh]	lag2	lag4	lag6
Brazil	0.783	98143.351	0.773	101324.240			
North-East	0.765	82860.401	0.753	88014.513			
South	0.407	6870.048	0.417	6865.726			
Bahia	0.673	5144.484	0.673	5250.066			
Ceará	0.876	3410.687	0.851	3603.748			
Maranhão	0.861	6.431	0.810	6.906			
Minas Gerais	0.060	0.002	0.065	0.002			
Paraíba	0.758	3.319	0.746	3.595			
Paraná	0.721	0.002	0.714	0.001			
Pernambuco	0.535	652.497	0.504	685.171			
Piaui	0.871	3634.479	0.867	3680.166			
Rio de Janeiro	0.553	0.393	0.554	0.387			
Rio Grande do Norte	0.721	10864.157	0.699	11539.912			
Rio Grande do Sul	0.388	6133.223	0.398	6114.658			
Santa Catarina	0.661	52.908	0.663	51.087			
Sergipe	0.527	1.056	0.519	1.032			

From above results of the analysis of the suitability of el Niño and La Niña for predicting future wind power generation with the help of linear regression models it can be concluded, that using the indices for prediction in most cases leads to better results (higher adjusted R-squared, lower MSEs), especially with a lag of two months. El Niño and La Niña should both be considered.

3.7 Correlations with water inflows

This section presents results from the analysis of correlations between water inflows and wind power generation in Brazil. Firstly, before calculation of correlations, the datasets need to be deseasonalised, as they show anticyclic seasonality. The seasonality of wind power has already been shown in Figure 6, that of water inflows is depicted in Figure 8. It can be clearly seen, that water inflows are lowest in Brazilian spring months, especially in September, and highest in autumn months, especially March, contrary to wind power generation, which is lowest in autumn and highest in spring months (see Figure 6). Therefore wind power generation, as well as water inflows are deseasonalised.

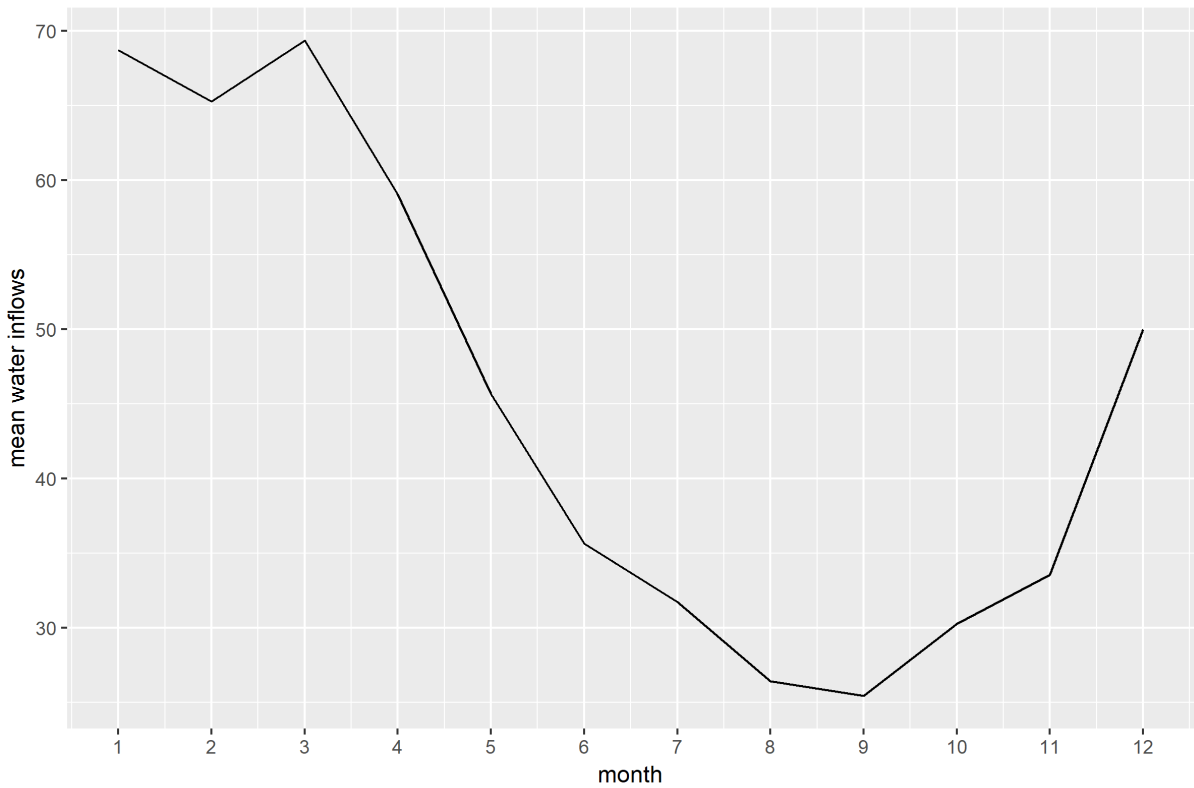

Figure 8: Mean of monthly water inflows into Brazilian hydropower plants (own depiction with data from [33])

Afterwards, correlations are calculated: Table 19 shows monthly and yearly correlations of deseasonalised wind power generation and water inflows for Brazil, its North-East and South. Monthly correlations without lag are not very high, between -3.8 % and 9.5 %. If lagging time series, correlations can be slightly increased. In Brazil the highest correlation of monthly time series occurs with a lag of two months at 10.6 % and it is also the only significant correlation in this comparison. In the North-East the highest correlation is even higher at 0.122 and with a lag of one month. Also the correlation with a lag of two months is significant, but lower at 0.096. In the South the highest absolute correlation is -0.131, thus higher, but negative. It occurs with a lag of four months and is the only correlation in the South, which is significant. This means, if there is an unusually high water event, it can be expected that wind power generation will slightly rise after two months, in the North-East maybe even a bit earlier, whereas in the South it is more likely to decrease four months later.

Also the yearly correlations are of interest, as they can indicate whether in years of higher water inflows and thus higher potential hydropower generation wind power generation is also higher or lower. The yearly correlations, which are presented in the right part of Table 19, are higher, especially in the South. They lie in a range of 14 % to 22.9 %, however they are not significant. Therefore, rainy or dry years seem not to have an impact on overall wind power generation.

Table 19: Monthly and yearly correlations of simulated wind power generation in Brazil (B), the North-East (NE) and the South (S) with lags between 0 and 12 months. Significant correlations are printed in bold (own representation)

	monthly			yearly		
lag	**B**	**NE**	**S**	**B**	**NE**	**S**
0	-0.038	0.013	0.095	0.039	0.017	0.216
1	0.092	**0.122**	0.038	0.052	-0.006	0.229
2	**0.106**	**0.096**	0.048	0.053	0.002	0.179
3	0.088	0.095	0.040	0.066	0.015	0.150
4	0.078	0.010	**-0.131**	0.100	0.022	0.155
5	0.012	-0.009	-0.065	0.120	0.027	0.129
6	0.029	0.000	-0.066	0.109	0.036	0.132
7	0.037	0.033	-0.066	0.066	0.044	0.097
8	0.021	0.039	0.013	0.059	0.068	0.078
9	-0.007	-0.001	-0.010	0.108	0.108	0.106
10	-0.038	-0.019	0.026	0.140	0.112	0.149
11	0.009	0.022	0.032	0.134	0.114	0.169
12	0.028	0.012	0.052	0.131	0.121	0.168

4 Discussion and Conclusion

In this thesis, a model, which originally was only generated for simulating wind power generation in the North-East of Brazil, is extended to the whole country. Furthermore, more precise data are applied, as well as several interpolation and wind speed correction methods are tested and analysed, to find the most accurate method for generating synthetic wind power time series. Finally, this simulation model is applied to analyse of the impacts of El Niño and La Niña as well as correlations with water inflows into Brazilian hydropower plants.

Results in previous chapters show that the developed model is capable of representing observed wind power generation. The quality of the simulated time series depends on the applied interpolation and wind speed correction methods and, more importantly, on the study region as well as on the level of aggregation. The results confirm, that wind power correction usually improves correlations, RMSEs and also differences in means – best improvements occur when no wind speed correction is applied beforehand, which fits the results from the first analysis only for the North-East of Brazil [23]. There, it was concluded, that at least one form of bias correction (wind power or wind speed correction) should be applied. For the whole of Brazil correlations between simulated and observed wind power generation are high, many above 0.99 and mostly above 0.98, for monthly as well as daily comparison. These are higher than in smaller regions as subsystems or single states. However, correlations in the North-East attain nearly as high levels. Compared to results from an analysis of wind power generation in Great Britain [41], which obtained hourly correlations of 96 % at high levels of aggregation, these are even better, which may be due to the larger area, and daily temporal aggregation, and thus higher smoothing effects in our model. In a simulation of European wind power generation [11], some similar, but also lower hourly correlations were achieved with values between 0.8 and 0.97 (with a few exceptions below). Although the simulation of Brazilian wind power generation seems to perform well, compared to other states, it has to be noted that these correlations are of hourly time series, but the ones presented in this work are daily and monthly correlations, as hourly wind power generation time series were not available for comparison.

Additionally, the performance of the developed model was analysed on smaller scale for the North-East and South of Brazil, where correlations between simulated and observed wind power time series are similar, though slightly lower (NE 0.978 – 0.994, S 0.957 – 0.990, for the daily comparison all – except one – are above 0.9). Some notable differences between the South and the North-East need to be mentioned: In the North-East correlations tend to be higher than in the South, relative RMSEs are lower in the North-East than in the South in the daily comparison, but higher in the monthly comparison. Deviations in means are lower in the North-East in the daily analysis, as well as in the monthly analysis before wind power correction, but after wind power correction they are higher in the

North-East than in the South. In general, for most simulations an improvement in terms of the analysed statistical parameters can be observed when applying wind power correction, however, this is more significant in the South than in the North-East, where wind power correction sometimes even worsens the results. It is furthermore worth mentioning that in the South, contrary to the North-East, wind speed correction has hardly any effect, which makes wind power correction even more important and more effective than in the North-East. This fact is most likely due to fewer places, where wind speed correction is applied: Figure 2 shows that in the South wind speed measurement stations usually are too far away or correlations between measured and reanalysis wind speeds are too low to apply wind speed bias correction. Especially in Paraná, which is small and has only little wind power capacity installed, for none of the installed wind parks wind speed correction is performed. Another difference between these subsystems is that in the North-East the simulation sometimes results in higher, sometimes in lower wind power generation than the actual one, whereas in the South wind power generation is usually overestimated.

Results of the simulation are investigated on an even lower level of aggregation, which is on the level of single states. Results vary significantly between states, but correlations are still above 0.9 after wind power correction (except in Santa Catarina and in the daily comparison neither in Pernambuco nor in Piaui). Also the effect of wind power correction cannot be identified, as in some states it leads to better results in terms of correlations, relative RMSEs and relative deviations in means of observed wind power generation, and in others it does not. In Paraná wind power correction results in significantly lower RMSEs in the monthly analysis; in daily analysis relative RMSEs are high in Rio Grande do Norte, but in general wind power correction lowers them. Santa Catarina generally shows some of the poorest results, which may be an indicator that wind speed correction is necessary to obtain good wind power simulations on a smaller scale. It is known that reanalysis data are not able to represent variations of wind speeds adequately [41] [42] [43], which is probably the reason why in a smaller region results are worse if not wind speed corrected. For larger areas these errors can be smoothed, as also reckoned by Goić et. al. [12], who recommend analysis of wind power generation over geographically spread areas to take advantage of this compensating effect. This balancing effect is of importance if a stable electricity supply is desired, as fluctuations in wind power generation are reduced if wind parks are interconnected [44].

An advantage of this simulation, compared to the previous one only for the North-East [23] may be that wind power correction is carried out for each state and not only for subsystems. This is an advantage as bias correction at lower spatial levels can improve correlations [14]. Another improvement compared to the first approach is the use of more precise data about wind farms. In particular geo-referenced locations are used, while in the previous work only the municipalities were known. The use of power curves of different wind turbines instead of only one standard turbine was considered, however, it was considered too much effort for the 50 different turbine types installed. However, this might not have a significant impact on the outcome: In the analysis of European wind power generation [14] different approaches were tried, one with detailed turbine type information and another with only one standard turbine type, as in the present work, and differences in results were only minimal.

Furthermore, for some simulations wind power generation is overestimated by up to more than 1.3-fold, whereas for others it is underestimated, similarly to results of a simulation of wind power in Europe [11], which showed that in some countries wind power is underestimated, whereas in others it is overestimated. However, for a few cases overestimation seems to dominate, especially in the South, where lower capacities are installed.

The comparison of different correction methods revealed, that Nearest Neighbour interpolation combined with wind power as well as wind speed correction, when at least long rows of 0 m/s wind speed are removed, performed the best. Surprisingly, more complex methods (Inverse Distance Weighting and Bilinear Interpolation) did not deliver better results. In the past work [23], it was observed for the North-East that time series of measured wind speeds sometimes contained obviously erroneous data, as wind speed was 0 m/s over longer periods of time. Removing these rows of 0 m/s wind speed apparently has a positive effect, as simulations where this selection of data is applied, result as the best.

The analysis of the impact of El Niño and la Niña is performed with this simulation and various results are obtained: Monthly and yearly correlations of wind power generation and El Niño and La Niña events reach a maximum of 9.8 % for the whole of Brazil in the yearly comparison, but none of the correlations is significant. Therefore, El Niño and La Niña events are examined separately and these results are more conclusive: For the comparison with La Niña events in the ONI maximum correlations of 16.1 % (South), 26.4 % (Brazil) and 29.8 % (North-East) are achieved, all of which are significant. With the one monthly Southern Oscillation Index correlations are lower and all negative, with lags of two to three months. With the ONI the correlation in the South is positive with a lag of 6 months, whereas for Brazil and its North-East correlations are negative with lags of three to four months. From these results it can be concluded that La Niña has a negative impact on wind power generation, meaning if one of these events occurs, it can be assumed that wind power generation will decrease after two to four months. In the South no conclusions can be drawn, as results are divergent. For El Niño events correlations are even higher, up to 31.1 % (in Brazil with ONI). Considering lags, they are also low, with two months, when compared to SOI, with ONI, however, lags of highest correlations are at six to seven months. To sum up, it can be seen that El Niño has a positive impact on wind power generation, meaning that after an event wind power generation is expected to rise. However, it is not conclusive with which lag this is most likely to happen, as they vary between two to seven months. In the South, results are less conclusive in general, as correlations are positive and negative when regarding La Niña, and with El Niño not significant when compared to the SOI. Therefore, in the North-East El Niño and La Niña events may help in predictability, whereas in the South, it may be more difficult to identify the impact of these events. This is underpinned also by the results from the analysis of the impacts on single states, where it was shown, that especially in the South results are ambiguous: Paraná shows a correlation of 23.2 % with SOI, whereas in Santa Catarina it is at -17.7 %, while both are located in the South. In the Project Ukko [45] wind is predicted and also skills are given for different prediction points: In the map available online it can be observed, that prediction skill is slightly lower in the South with a maximum of 23 %, than in the North-East, where values of up to 27 % are reached.

In previous studies mainly the impact of ENSO on rainfall has been investigated. Grimm et al. [46], for example, concluded that the impact of El Niño and La Niña is perceived especially in the South of Brazil, as well as in Argentina and Uruguay, which is contradictory with the results obtained in this work, where the impact in the South is smaller, if even significant, than in the North-East. Nogueira Lima et al. [8] investigate correlations between wind and El Niño and La Niña events, and find that there are positive correlations of nearly 4 %, which is lower than the correlations found here. They state, however, that the effect for wind power may be higher, as wind speed affects wind power with a power of three. In an analysis of the effects of El Niño on wind power in Chile [22] it was found that in most cases these events lead to a reduction of wind power generation, however depending on the site examined as well as on the month in which the event occurs. The results of this work indicate different behaviour: The highest correlations of wind power generation with El Niño events are positive, meaning that such an event will probably lead to an increase in wind power generation. The authors also point out the necessity of such knowledge for the prediction of future wind power output, which is crucial information for electrical system operators. Therefore, it was also attempted to build a linear regression model which predicts wind power generation when fed with past wind power generation as well as the index for El Niño and La Niña events. It was shown that for most regions of Brazil adjusted R-squared between data from the linear model and test data is higher, when including El Niño and La Niña events, especially with a lag of two months. These results are very interesting, as they show that it can have an impact – although rather low – on the prediction of future wind power output from historical data if ENSO events are included in the prediction model. Therefore, when planning for energy systems it can be recommended considering also El Niño and La Niña in forecasts.

Another aspect that has been investigated in this work are the correlations or anti-correlations of wind power and hydropower. The analysis shows that after deseasonalisation there is some relation between wind power and water inflows into Brazilian hydropower plants, however absolute correlations are not very high, with a maximum of -13.1 % in the South. Regarding yearly analysis, no significant correlations were determined. This fits the observation made by Bezerra et al. [47], who found no inter-annual complementarity between hydro inflows and availability of wind. Chade Ricosti and Sauer [48] investigated the relationship of wind power and rain in the North-East and found that in years of lower precipitation wind power generation was higher, which contradicts the present results, as positive correlations were found for this region.

For the future the question remains whether it is possible to improve the present simulation of wind power generation. It is not assumed that more complex spatial interpolation methods will increase quality, as the simplest Nearest Neighbour method proved to yield the best results. What can be improved, however, are underlying data, especially wind speed measurement data: Data are erroneous in several cases and also wind speed measurement stations close to wind parks or locations of interest for wind power could be installed to provide better data for bias correction. Also reanalysis data with a higher spatial resolution might be of interest. It will be task of future research, to find out whether the results of the analysis of the impact of El Niño and La Niña can be represented in events to come.

References

[1] J. Schmidt, R. Cancella and A. O. Pereira Junior, “An optimal mix of solar PV, wind and hydro power for a low-carbon electricity supply in Brazil,” *Renewable Energy,* pp. 137-147, January 2016.

[2] J. Schmidt, R. Cancella and A. O. Pereira Junior, “The effect of windpower on long-term variability of combined hydro-wind resources: The case of Brazil,” *Renewable and Sustainable Energy Reviews,* pp. 131-141, March 2016.

[3] IRENA - International Renewable Energy Agency, “Renewable Energy Market Analysis. Latin America,” IRENA, Abu Dhabi, 2016.

[4] FBDS - Fundação Brasileira para o Desenvolvimento Sustentável, “Climate change and extreme events in Brazil,” LLOYDS, Rio de Janeiro, 2010.

[5] J. D. Makholm, “El Niño’s Uneven Disruption of World’s Electricity Systems,” *International Energy,* pp. 29-32, April 2016.

[6] P. E. Bett and H. E. Thornton, “The climatological relationships between wind and solar energy supply in Britain,” *Renewable Energy,* pp. 96-110, 2016.

[7] J. Schmidt, R. Cancella and A. O. J. Pereira, “The role of wind power and solar PV in reducing risks in the Brazilian hydro-thermal power system,” *Energy,* pp. 1748-1757, November 2016.

[8] C. N. Nogueira Lima, C. A. Coelho Fernandes, G. Borges França and G. Gonçalves de Matos, “Estimação do Impacto do El Niño/La Niña na Intensidade dos Ventos do Nordeste Brasileiro,” *Anuário do Instituto de Geociências,* pp. 232-240, February 2014.

[9] M. Barasa and A. Aganda, “Wind power variability of selected sites in Kenya and the impact to system operating reserve,” *Renewable Energy,* pp. 464-471, 2015.

[10] I. González-Aparicio and A. Zucker, “Impact of wind power uncertainty forecasting on the market integration of wind energy in Spain,” *Applied Energy,* pp. 334-349, 16 September 2015.

[11] I. González-Aparicio, F. Monforti, P. Volker, A. Zucker, F. Careri, T. Huld and J. Badger, “Simulating European wind power generation applying statistical downscaling to reanalysis data,” *Applied Energy,* pp. 155 - 168, 9 May 2017.

[12] R. Goić, J. Krstulović and D. Jakus, “Simulation of aggregate wind farm short-term production variations,” *Renewable Energy,* pp. 2602-2609, 1 May 2010.

[13] C. E. Hoicka and I. H. Rowlands, “Solar and wind resource complementarity: Advancing options for renewable electricity integration in Ontario, Canada,” *Renewable Energy,* pp. 97-107, 2011.

[14] F. Monforti, T. Huld, K. Bódis, L. Vitali, M. D'Isidoro and R. Lacal-Arántegui, “Assessing complementarity of wind and solar resources for energy production in Italy. A Monte Carlo approach,” *Renewable Energy,* pp. 576-586, 2014.

[15] INPE/CPTEC, “El Niño e La Niña – CPTEC/INPE,” INPE - Instituto Nacional de Pesquisas Espaciais, CPTEC - Centro de Previsão de Tempo e Estudos Climáticos, 12 March 2018. [Online]. Available: http://enos.cptec.inpe.br/. [Accessed 26 March 2018].

[16] P. Thomson, “El Niño is back, and global temperature records are in danger,” PRI - Public Radio International, 4 July 2014. [Online]. Available: https://www.pri.org/stories/2015-05-13/el-ni-o-back-and-global-temperature-records-are-danger. [Accessed 5 March 2018].

[17] M. McPhaden, “Predicting El Niño Then and Now,” NOAA - National Oceanic and Atmospheric Administration, 3 April 2015. [Online]. Available: https://www.climate.gov/news-features/blogs/enso/predicting-el-ni%C3%B1o-then-and-now. [Accessed 5 March 2018].

[18] A. Cook, A. B. Watkins, B. Trewin and C. Ganter, “El Niño is over, but has left its mark across the world,” The Conversation UK, 25 May 2016. [Online]. Available: https://theconversation.com/el-nino-is-over-but-has-left-its-mark-across-the-world-59823. [Accessed 5 March 2018].

[19] P. Medina Uribe, “Update: El Niño's Latin American Impact,” Americas Society/Council of the Americas, February 2016. [Online]. Available: https://www.as-coa.org/articles/update-el-ni%C3%B1os-latin-american-impact. [Accessed 5 March 2018].

[20] E. Leister, “Brazil Drought: El Nino Impacts and Political Unrest,” AccuWeather, Inc., 12 June 2014. [Online]. Available: https://www.accuweather.com/en/weather-news/brazil-drought-el-nino-world-cup/28146693. [Accessed 5 March 2018].

[21] D. Rife, “DNV GL study: El Niño not cause of 2015 “Wind Drought”,” DNV GL, 15 February 2016. [Online]. Available: https://www.dnvgl.com/news/dnv-gl-study-el-nino-not-cause-of-2015-wind-drought--57798. [Accessed 5 March 2018].

[22] D. Watts, P. Durán and Y. Flores, “How does El Niño Southern Oscillation impact the wind resource in Chile? A techno-economical assessment of the influence of El Niño and La Niña on the wind power,” *Renewable Energy,* pp. 128-142, 2017.

[23] K. Gruber, *Simulation of Synthetic Wind Power Time Series in the North-East of Brazil,* Vienna: University of Natural Resources and Life Sciences, Vienna, 2017.

[24] Github Repository, “MERRABin,” [Online]. Available: https://github.com/joph/RE_EXTREME/tree/master/scripts.

[25] G. Cruz, F. Estrela, B. Junior and M. Lima, “Shapefiles do Brasil para download,” CodeGeo, 16 April 2013. [Online]. Available: http://www.codegeo.com.br/2013/04/shapefiles-do-brasil-para-download.html. [Accessed 05 Janaury 2018].

[26] Brasil Governo Federal, “INMET Instituto Nacional de Meteorologia,” [Online]. Available: http://www.inmet.gov.br/portal/. [Accessed 11 February 2018].

[27] P. Michaël, “The Wind Power. Wind Energy Market Intelligence,” The Wind Power, [Online]. Available: https://www.thewindpower.net/. [Accessed 11 February 2018].

[28] ENERCON, *ENERCON Produktübersicht,* Aurich, Germany: ENERCON GmbH, 2015.

[29] ONS, “GERAÇÃO DE ENERGIA,” ONS, 2018. [Online]. Available: http://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/geracao_energia.aspx. [Accessed 05 January 2018].

[30] Climate Prediction Center, “Cold & Warm Episodes per Season,” NOAA Center for Weather and Climate Prediction , 06 January 2018. [Online]. Available: http://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php. [Accessed 10 December 2017].

[31] National Climatic Data Center, “Southern Oscillation Index (SOI),” National Oceanic Atmospheric Administration, 06 January 2018. [Online]. Available: https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/. [Accessed 10 December 2017].

[32] National Climatic Data Center, “Southern Oscillation Index (SOI),” National Oceanic Atmospheric Administration, 06 January 2018. [Online]. Available:

https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/data.csv. [Accessed 10 December 2017].

[33] J. Schmidt, "daily inflows 4regions 1979 2014," 20 April 2017. [Online]. Available: https://homepage.boku.ac.at/jschmidt/COPA/daily_inflows_4regions_1979_2014.csv. [Accessed 07 January 2018].

[34] J. Li and A. D. Heap, "A Review of Spatial Interpolation Methods for Environmental Scientists," Geoscience Australia, Canberra, 2008.

[35] R. Sluiter, "Interpolation methods for climate data," KNMI, De Bilt, 2009.

[36] E. Svensson, *Performance of Long Term Wind Estimation Method at Wind Power Development,* Göteborg: Chalmers University of Technology, 2012.

[37] M. Keskin, A. Ozgur Dogru, F. Bektas Balcik, C. Goksel, N. Ulugtekin and S. Sozen, "Comparing Spatial Interpolation Methods for Mapping Meteorological Data in Turkey," *Energy Systems and Management,* pp. 33-42, 26 March 2015.

[38] M. Bilala, G. Arayab and Y. Birkelund, "Preliminary assessment of remote wind sites," in *The 7th International Conference on Applied Energy – ICAE2015*, Abu Dhabi, 2015.

[39] H. Wickham, "A Layered Grammar of Graphics," *Journal of Computational and Graphical Statistics,* pp. 3-28, 01 January 2012.

[40] ETH Zürich, "Test for Association/Correlation Between Paired Samples," [Online]. Available: http://stat.ethz.ch/R-manual/R-devel/library/stats/html/cor.test.html. [Accessed 22 April 2018].

[41] D. J. Cannon, D. J. Baryshaw, J. Methven, P. J. Coker and D. Lenaghan, "Using reanalysis data to quantify extreme wind power generation statistics: A 33 year case study in Great Britain," *Renewable Energy,* pp. 767-778, March 2015.

[42] J. Olauson and M. Bergkvist, "Modelling the Swedish wind power production using MERRA reanalysis data," *Renewable Energy,* pp. 717-725, April 2015.

[43] S. Pfenninger and I. Staffell, "Using bias-corrected reanalysis to simulate current and future wind power output," *Energy,* pp. 1224-1239, November 2016.

[44] F. J. Santos-Alamillos, D. Pozo-Vázquez, J. A. Ruiz-Arias, V. Lara-Fanego and J. Tovar-Pescador, "A methodology for evaluating the spatial variability of wind energy resources: Application to assess

the potential contribution of wind energy to baseload power," *Renewable Energy,* pp. 147-156, 2014.

[45] Barcelona Supercomputing Center, Future Everything, "Project Ukko," Barcelona Supercomputing Center, Future Everything, 2018. [Online]. Available: http://project-ukko.net/. [Accessed 4 April 2018].

[46] A. M. Grimm, S. E. T. Ferraz and J. Gomes, "Precipitation Anomalies in Southern Brazil Associated with El Niño and La Niña Events," *ournal of Climate,* pp. 2863-2880, November 1998.

[47] B. C. G. Bezerra, P. B. L. Àvila, M. Carvahlho and M. Pereira, "Análise do percentual máximo para inserção de energia eólica na matriz elétrica brasileira sob a ótica energética," in *XXII Seinário de produção e transmissão de energia elétrica*, Brasilia, 2013.

[48] J. F. Chade Ricosti and I. L. Sauer, "An assessment of wind power prospects in the Brazilian hydrothermal system," *Renewable and Sustainable Energy Reviews,* pp. 742-753, March 2013.

Appendix

1 Additional Results

This section presents several more results additionally to the ones in the Results chapter, as there were too many to show all of them there. For the results already presented above, also figures are presented and for the comparison on the level of states, tables with the calculated values are displayed as well as some graphs.

1.1 Monthly wind power generation per state

This chapter presents the results of the comparison of simulated and observed monthly wind power generation per state. Table 20 compares correlations between observed and simulated wind power generation: In some states (Bahia, Ceará, Pernambuco, Piaui, Rio Grande do Norte and Rio Grande do Sul) monthly correlations are very high (above 0.9) already before wind power correction. In other states (Paraíba, Rio de Janeiro, Paraná and Sergipe), they are mostly significantly increased with wind power correction. Only in Santa Catarina the correlations are also quite low after wind power correction (between 0.677 and 0.727). The lowest correlations before wind power correction can be observed in Paraíba and Sergipe at 0.538 and 0.390, respectively, whereas other, higher correlations reach up to 0.994.

In Table 21, relative RMSEs for eleven states of Brazil are compared. Values vary significantly between different states: Before wind power correction, lower relative RMSEs are found in Bahia, Pernambuco, Rio Grande do Sul and Rio de Janeiro (between 0.155 and 0.685), but higher values are achieved in Paraíba, Paraná, Rio Grande do Norte, Santa Catarina and Sergipe, where relative RMSEs up to 4.050 are achieved (in Paraná). After wind power correction, relative RMSEs are decreased to between 0.050 and 0.348 (without considering Rio Grande do Norte). The most significant reduction is observed in Paraná, where the highest relative RMSE before wind power correction is 4.050 and after correction 0.101. Wind speed correction in some cases leads to lower RMSEs, but in other cases the relative RMSE is higher when wind speed correction is applied. It cannot be determined which method results in the lowest relative RMSE as this varies from state to state.

Table 20: Correlations of monthly wind power generation time series for eleven states of Brazil (own representation)

		correlations							correlations				
		x	r	m	rm	nINc			x	r	m	rm	nINc
Bahia	NN	0.987	0.987	0.984	0.987	0.985	Rio de Janeiro	NN	0.778	0.917	0.892	0.917	0.778
	BLI	0.987	0.987	0.988	0.987	0.983		BLI	0.857	0.927	0.900	0.857	0.857
	IDW	0.987	0.983	0.979	0.987	0.981		IDW	0.849	0.849	0.916	0.928	0.849
	NNc	0.994	0.994	0.988	0.994	0.993		NNc	0.945	0.936	0.928	0.936	0.945
	BLIc	0.994	0.994	0.993	0.994	0.992		BLIc	0.957	0.943	0.939	0.957	0.957
	IDWc	0.994	0.988	0.986	0.994	0.992		IDWc	0.955	0.955	0.940	0.941	0.955
Ceará	NN	0.956	0.956	0.872	0.956	0.958	Rio Grande do Norte	NN	0.979	0.975	0.977	0.975	0.977
	BLI	0.957	0.957	0.943	0.961	0.958		BLI	0.977	0.975	0.977	0.965	0.977
	IDW	0.971	0.976	0.890	0.971	0.956		IDW	0.977	0.986	0.977	0.976	0.977
	NNc	0.977	0.977	0.950	0.977	0.978		NNc	0.980	0.985	0.979	0.985	0.979
	BLIc	0.981	0.981	0.959	0.980	0.981		BLIc	0.979	0.984	0.979	0.984	0.979
	IDWc	0.980	0.983	0.950	0.980	0.981		IDWc	0.979	0.990	0.979	0.983	0.979
Paraiba	NN	0.882	0.883	0.579	0.883	0.843	Rio Grande do Sul	NN	0.990	0.990	0.987	0.990	0.989
	BLI	0.896	0.896	0.826	0.905	0.826		BLI	0.991	0.991	0.986	0.991	0.990
	IDW	0.890	0.662	0.538	0.890	0.835		IDW	0.990	0.977	0.986	0.991	0.990
	NNc	0.976	0.976	0.981	0.976	0.986		NNc	0.991	0.991	0.988	0.991	0.990
	BLIc	0.971	0.971	0.985	0.970	0.985		BLIc	0.991	0.991	0.988	0.991	0.991
	IDWc	0.974	0.976	0.979	0.974	0.985		IDWc	0.991	0.983	0.989	0.991	0.991
Paraná	NN	0.793	0.793	0.793	0.793	0.793	Santa Catarina	NN	0.670	0.670	0.655	0.670	0.655
	BLI	0.728	0.728	0.728	0.728	0.728		BLI	0.705	0.705	0.693	0.693	0.693
	IDW	0.700	0.700	0.700	0.700	0.700		IDW	0.692	0.681	0.681	0.692	0.681
	NNc	0.911	0.911	0.911	0.911	0.911		NNc	0.693	0.693	0.677	0.693	0.677
	BLIc	0.911	0.911	0.911	0.911	0.911		BLIc	0.727	0.727	0.715	0.715	0.715
	IDWc	0.907	0.907	0.907	0.907	0.907		IDWc	0.724	0.715	0.715	0.724	0.715
Pernambuco	NN	0.952	0.952	0.960	0.952	0.951	Sergipe	NN	0.417	0.612	0.790	0.612	0.790
	BLI	0.952	0.952	0.948	0.951	0.949		BLI	0.390	0.632	0.746	0.659	0.746
	IDW	0.950	0.958	0.936	0.950	0.948		IDW	0.395	0.734	0.749	0.630	0.749
	NNc	0.987	0.987	0.985	0.987	0.986		NNc	0.871	0.879	0.913	0.879	0.913
	BLIc	0.985	0.985	0.985	0.986	0.984		BLIc	0.900	0.903	0.925	0.903	0.925
	IDWc	0.987	0.984	0.983	0.987	0.985		IDWc	0.894	0.914	0.922	0.898	0.922
Piaui	NN	0.940	0.942	0.982	0.942	0.975							
	BLI	0.939	0.942	0.984	0.974	0.976							
	IDW	0.950	0.977	0.985	0.952	0.977							
	NNc	0.989	0.989	0.989	0.989	0.983							
	BLIc	0.989	0.989	0.990	0.985	0.984							
	IDWc	0.991	0.985	0.991	0.991	0.985							

Table 21: Relative RMSEs of monthly wind power generation time series for eleven states of Brazil (own representation)

		relative RMSEs							relative RMSEs				
		x	r	m	rm	nINc			x	r	m	rm	nINc
Bahia	NN	0.164	0.164	0.355	0.164	0.207	Rio de Janeiro	NN	0.289	0.299	0.558	0.299	0.289
	BLI	0.207	0.207	0.166	0.155	0.189		BLI	0.262	0.315	0.401	0.262	0.262
	IDW	0.224	0.685	0.431	0.224	0.189		IDW	0.259	0.259	0.394	0.254	0.259
	NNc	0.129	0.129	0.182	0.129	0.160		NNc	0.142	0.153	0.161	0.153	0.142
	BLIc	0.126	0.126	0.160	0.135	0.162		BLIc	0.127	0.145	0.148	0.127	0.127
	IDWc	0.131	0.202	0.198	0.131	0.167		IDWc	0.129	0.129	0.147	0.147	0.129
Ceará	NN	0.349	0.349	0.468	0.349	0.847	Rio Grande do Norte	NN	0.475	0.754	0.900	0.754	0.900
	BLI	0.332	0.332	0.335	0.329	0.859		BLI	0.482	0.681	0.873	0.592	0.873
	IDW	0.375	0.334	0.458	0.375	0.864		IDW	0.503	1.006	0.885	0.639	0.885
	NNc	0.237	0.237	0.344	0.237	0.232		NNc	0.486	0.373	0.480	0.373	0.480
	BLIc	0.207	0.207	0.301	0.211	0.217		BLIc	0.494	0.408	0.480	0.402	0.480
	IDWc	0.222	0.219	0.348	0.222	0.221		IDWc	0.497	0.225	0.480	0.418	0.480
Paraíba	NN	0.385	0.377	0.248	0.377	1.678	Rio Grande do Sul	NN	0.399	0.399	0.362	0.399	0.432
	BLI	0.257	0.252	1.645	0.210	1.645		BLI	0.334	0.334	0.298	0.336	0.362
	IDW	0.312	0.301	0.281	0.306	1.672		IDW	0.343	0.659	0.294	0.343	0.368
	NNc	0.065	0.065	0.058	0.065	0.050		NNc	0.166	0.166	0.207	0.166	0.176
	BLIc	0.072	0.072	0.052	0.073	0.052		BLIc	0.169	0.169	0.196	0.169	0.179
	IDWc	0.068	0.066	0.060	0.068	0.052		IDWc	0.173	0.229	0.201	0.173	0.183
Paraná	NN	3.265	3.265	3.265	3.265	3.265	Santa Catarina	NN	0.597	0.597	0.612	0.597	0.612
	BLI	3.651	3.651	3.651	3.651	3.651		BLI	0.928	0.928	0.942	0.942	0.942
	IDW	4.050	4.050	4.050	4.050	4.050		IDW	1.368	1.381	1.381	1.368	1.381
	NNc	0.097	0.097	0.097	0.097	0.097		NNc	0.216	0.216	0.221	0.216	0.221
	BLIc	0.097	0.097	0.097	0.097	0.097		BLIc	0.206	0.206	0.209	0.209	0.209
	IDWc	0.101	0.101	0.101	0.101	0.101		IDWc	0.207	0.209	0.209	0.207	0.209
Pernambuco	NN	0.260	0.259	0.184	0.259	0.201	Sergipe	NN	0.802	0.731	0.725	0.731	0.725
	BLI	0.272	0.272	0.195	0.219	0.209		BLI	0.674	0.557	0.783	0.542	0.783
	IDW	0.274	0.268	0.307	0.274	0.214		IDW	0.702	2.811	0.776	0.595	0.776
	NNc	0.100	0.099	0.104	0.099	0.098		NNc	0.157	0.148	0.117	0.148	0.117
	BLIc	0.104	0.103	0.105	0.099	0.102		BLIc	0.129	0.126	0.107	0.126	0.107
	IDWc	0.103	0.114	0.116	0.103	0.102		IDWc	0.133	0.114	0.109	0.130	0.109
Piaui	NN	0.360	0.354	0.136	0.354	0.277							
	BLI	0.376	0.368	0.131	0.213	0.255							
	IDW	0.407	0.242	0.127	0.400	0.242							
	NNc	0.123	0.122	0.113	0.122	0.146							
	BLIc	0.119	0.117	0.106	0.148	0.138							
	IDWc	0.116	0.135	0.105	0.115	0.135							

Figure 9 shows the RMSEs of monthly wind power generation per state. In Paraíba, Paraná, Pernambuco, Rio de Janeiro, Rio Grande do Sul, Santa Catarina and Sergipe the RMSEs are quite low, even before wind power correction. The highest RMSEs are observed in Rio Grande do Norte, with up to more than 200 GWh RMSE with wind speed correction with one adaptation. In Ceará, simulations without any bias correction are also high, as well as in Bahia with Inverse Distance Weighting, when only wind speed correction with removal of long rows of 0 m/s wind speed is applied.

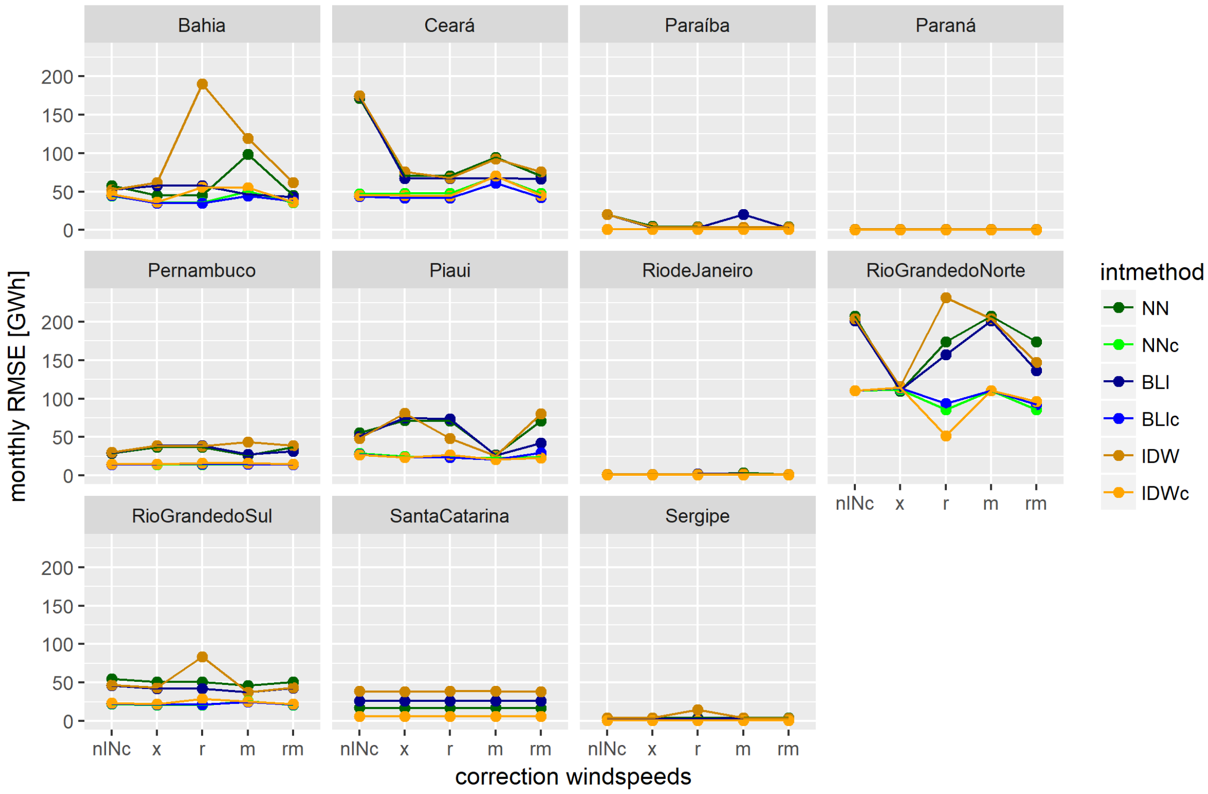

Figure 9: RMSEs between observed and simulated monthly wind power generation in eleven states of Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method (own depiction)

Table 22 represents the normalised differences between monthly means of observed and simulated wind power generation in eleven states of Brazil, only for not wind power corrected methods, as with wind power correction the differences are always 0, due to the method. In most cases the relative differences in means are quite small, slightly above 0. Only in Santa Catarina and in Paraná, the normalised differences are considerably higher, especially when Inverse Distance Weighting is applied. In the first state, relative deviations in means are one and a half times as high for Inverse Distance Weighting and in the latter, even four times as high as observed wind power generation. Also in Sergipe, the simulation overestimates observed wind power generation for about the threefold with Inverse Distance Weighting and wind speed correction with removal of long rows of 0 m/s winds speed. Wind speed correction reduces relative deviations in means for some cases, but not for all, which is similar to the results of comparison of RMSEs.

Table 22: Relative absolute differences in means of monthly wind power generation time series for eleven states of Brazil (own representation)

		rel. abs. difference in means				
		x	**r**	**m**	**rm**	**nINc**
Bahia	**NN**	0.011	0.011	0.186	0.011	0.148
	BLI	0.067	0.067	0.088	0.023	0.104
	IDW	0.083	0.624	0.244	0.083	0.089
Ceará	**NN**	0.066	0.066	0.050	0.066	0.754
	BLI	0.081	0.081	0.033	0.075	0.752
	IDW	0.156	0.283	0.053	0.156	0.756
Paraíba	**NN**	0.352	0.344	0.060	0.344	1.614
	BLI	0.212	0.206	1.579	0.162	1.579
	IDW	0.274	0.079	0.126	0.267	1.606
Paraná	**NN**	3.242	3.242	3.242	3.242	3.242
	BLI	3.624	3.624	3.624	3.624	3.624
	IDW	4.017	4.017	4.017	4.017	4.017
Pernam-buco	**NN**	0.186	0.186	0.072	0.186	0.086
	BLI	0.202	0.202	0.057	0.121	0.097
	IDW	0.202	0.203	0.212	0.202	0.102
Piaui	**NN**	0.251	0.246	0.010	0.246	0.117
	BLI	0.263	0.257	0.018	0.046	0.103
	IDW	0.310	0.090	0.034	0.304	0.090

		rel. abs. difference in means				
		x	**r**	**m**	**rm**	**nINc**
Rio de Janeiro	**NN**	0.021	0.131	0.484	0.131	0.021
	BLI	0.105	0.203	0.318	0.105	0.105
	IDW	0.084	0.084	0.312	0.099	0.084
Rio Grande do Norte	**NN**	0.348	0.279	0.689	0.279	0.689
	BLI	0.347	0.208	0.668	0.107	0.668
	IDW	0.372	0.518	0.675	0.172	0.675
Rio Grande do Sul	**NN**	0.204	0.204	0.140	0.204	0.218
	BLI	0.146	0.146	0.097	0.147	0.157
	IDW	0.148	0.443	0.088	0.148	0.158
Santa Catarina	**NN**	0.530	0.530	0.544	0.530	0.544
	BLI	0.879	0.879	0.892	0.892	0.892
	IDW	1.333	1.345	1.345	1.333	1.345
Sergipe	**NN**	0.759	0.696	0.652	0.696	0.652
	BLI	0.607	0.502	0.712	0.482	0.712
	IDW	0.642	2.777	0.703	0.546	0.703

The comparison of means of monthly wind power generation per state can be seen in Figure 10. In Paraíba, Paraná, Rio de Janeiro and Sergipe monthly means of wind speed are very low and close to 0 GWh, as there is just little capacity installed. In most of the other states, the means of simulated monthly wind power generation are close to observed monthly wind power generation, except for certain methods in Bahia, Ceará and Rio Grande do Norte. When wind speed correction with removal of long rows of 0 m/s wind speed is applied, some deviation of the monthly mean of simulated wind power generation from observed wind power generation can be seen, especially in Bahia. In Ceará and also in Rio Grande do Norte also the simulations without wind speed correction differ considerably from the mean of observed monthly wind power generation. The means of wind power corrected simulations are always the same as the ones of observed wind power generation, due to the method.

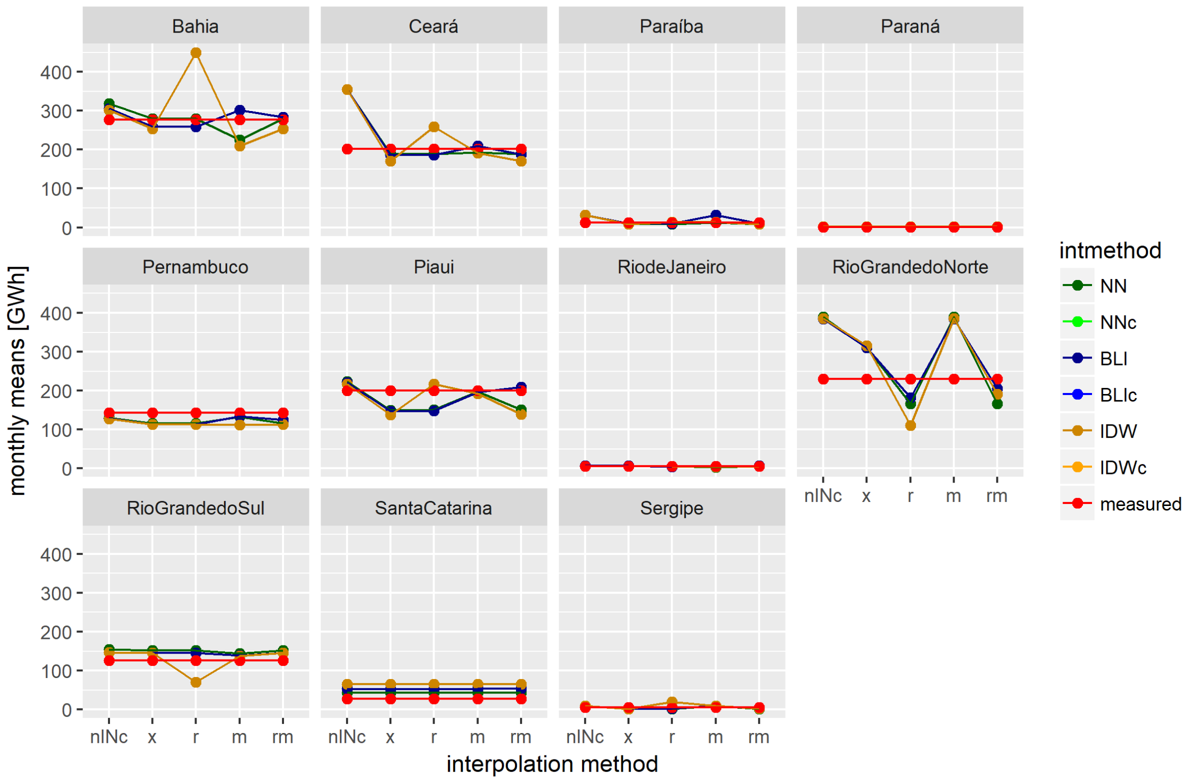

Figure 10: Means of monthly wind power generation in eleven states of Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method and the observed values (own depiction)

Figure 11 shows the absolute differences between simulated and observed monthly wind power generation in the states per interpolation method. Before wind power correction, highest deviations from observed wind power generation occur when applying Inverse Distance Weighting, where simulated wind power generation differs from observed wind power generation up to 115 GWh (without outliers). The Nearest Neighbour method and Bilinear Interpolation are very close regarding differences, which reach up to about 100 GWh, disregarding outliers. After wind power correction absolute differences are reduced and more than 75 % of the time, simulated wind power generation differs from observed wind power generation less than 50 GWh. No differences between various interpolation methods are visible.

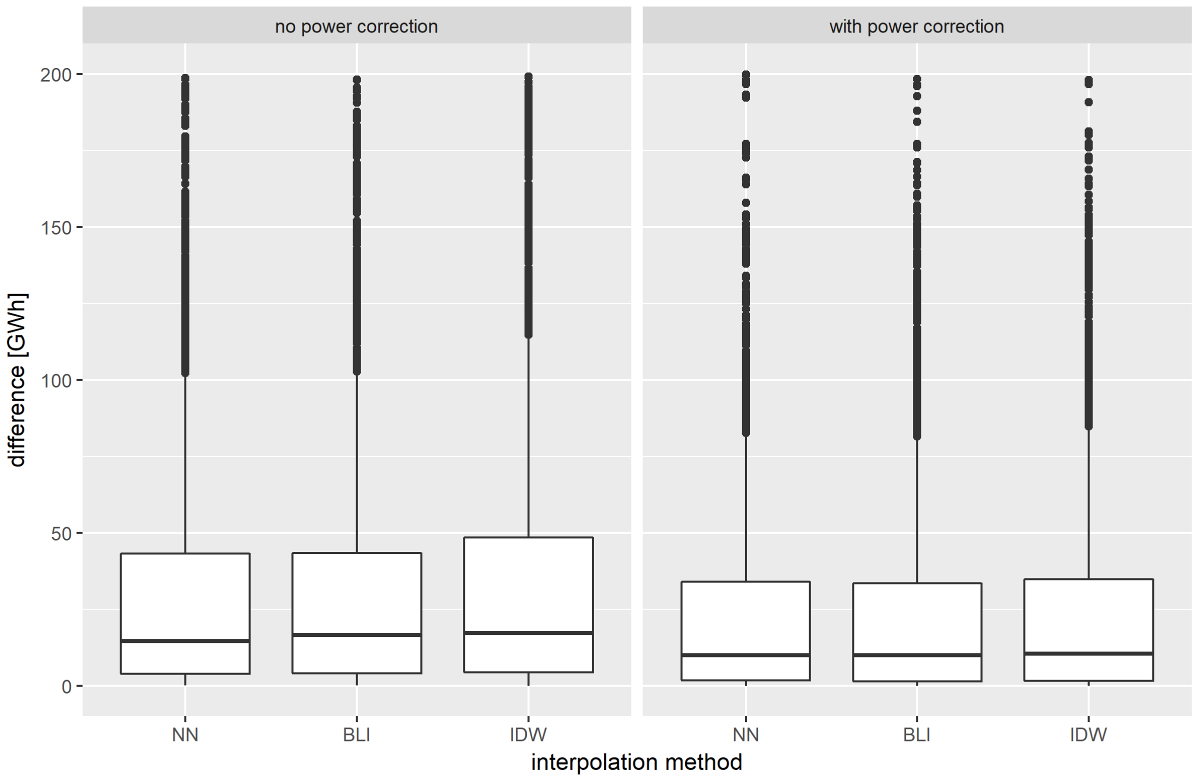

Figure 11: Absolute monthly differences in wind power generation per interpolation method in eleven states of Brazil. On the x-axis different interpolation methods are shown (own depiction)

In Figure 12, the differences between simulated and observed monthly wind power generation per wind speed correction method are shown. Before wind power correction, the method without wind speed correction seems worst adapted regarding differences, as they can be as high as 150 GWh, whereas when wind speed correction is applied, most of the absolute monthly differences are below 110 GWh (disregarding outliers). The lower 50 %, however, are in a similar range of up to 20 GWh. After wind power correction, the results can be improved regarding differences in simulated and observed wind power generation. The smallest improvement is observed when wind speed correction with adaptation of means is applied and the most significant improvement is seen when no wind speed correction is performed. The lowest differences occur when normal wind speed correction without adaptations is applied, closely followed by the methods where at least removal of long rows of 0 m/s wind speed is applied during wind speed correction. Furthermore, it is notable, that after wind power correction, the largest differences are not obtained without wind speed correction, but with wind speed correction with approximation of means.

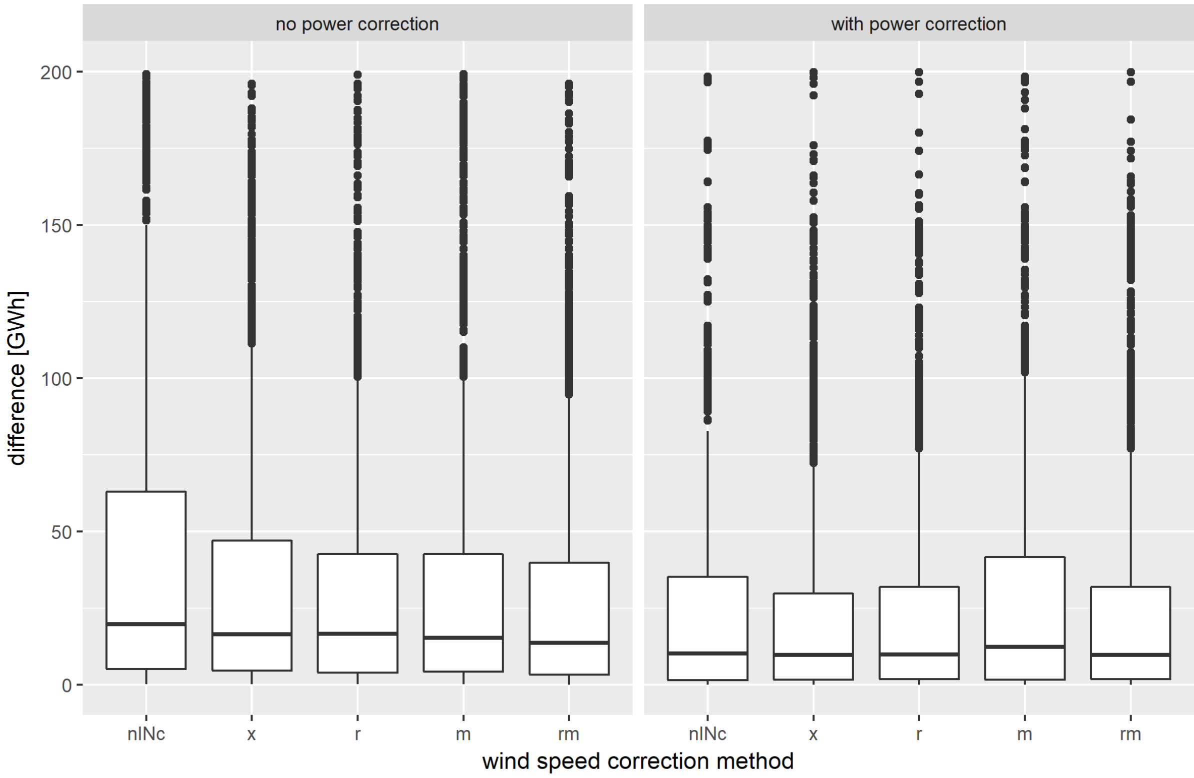

Figure 12: Absolute monthly differences in wind power generation per wind speed correction method in eleven states of Brazil. On the x-axis different wind speed correction methods are shown (own depiction)

It can be seen from the above plots and tables, that simulation of wind power generation without any correction can differ significantly from observed monthly wind power generation, at least for some states. Wind power correction always improves results in terms of correlation, comparison of means, RMSEs and also differences between observed and simulated wind power generation. For single interpolation and wind speed correction methods, results are ambiguous and it cannot be determined which method should be preferred.

1.2 Monthly wind power generation per subsystem

This section provides some additional results to the ones in chapter 3.1 in graphical form. The first parameter which is used to compare monthly wind power generation between the two subsystems is the RMSE. Figure 13 shows the RMSEs between observed monthly wind power generation and different simulations. In the South, the RMSEs are lower than in the North-East, due to the lower installed capacities and therefore lower wind power generation. It can be clearly seen, that RMSEs are lower when applying wind power correction. The highest RMSEs appear when no bias correction is applied. In the North-East the results of comparison of RMSEs of monthly wind power generation are more complex: When wind speed correction is applied without any adaptations, wind power correction has hardly any impact on the RMSEs, regardless of the interpolation method used. In other cases (no wind speed correction, wind speed correction with at least removal of long rows of 0 m/s wind speed), wind power correction lowers the RMSEs between simulated and observed monthly wind power generation. The highest decrease in RMSE is observed when no wind speed correction is applied before wind power correction.

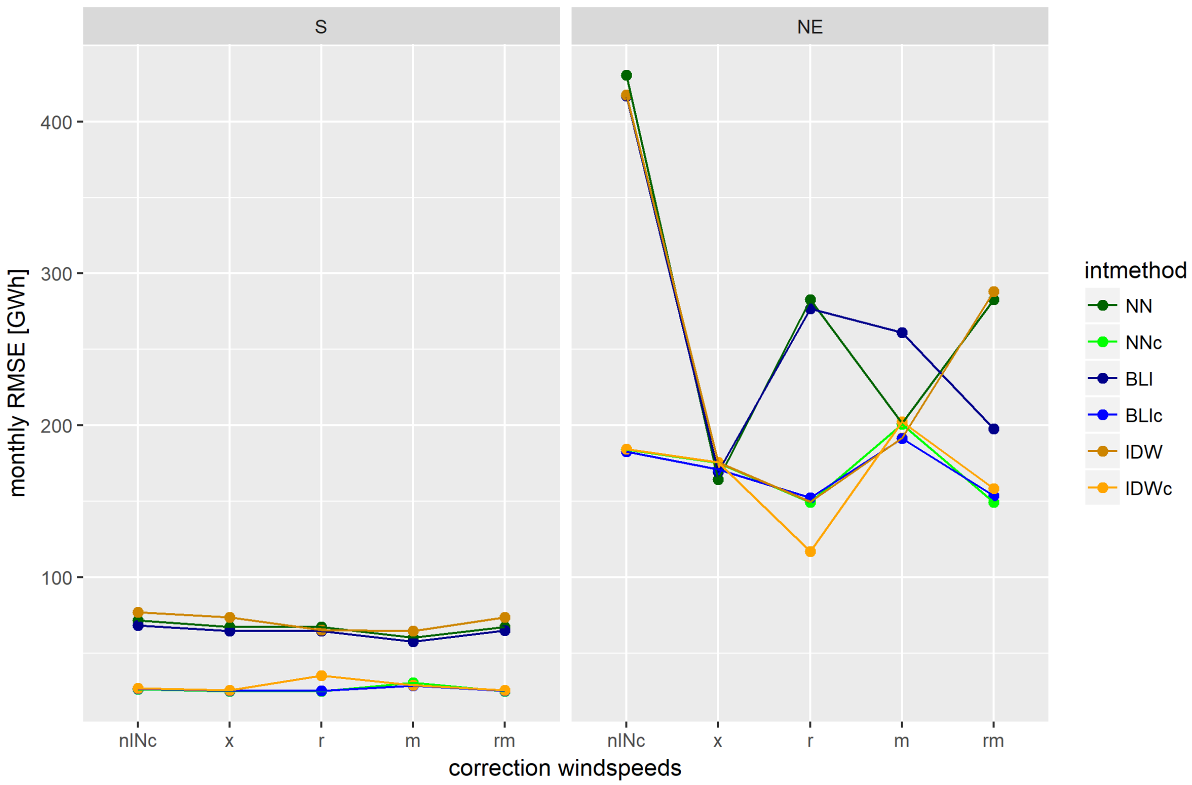

Figure 13: RMSEs between observed and simulated monthly wind power generation in the North-East (NE) and South (S) of Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method (own depiction)

The following graph (Figure 14) shows a comparison of means of simulated and observed monthly wind power generation in the North-East and in the South of Brazil. In the South, the monthly means of wind power generation are in general lower than in the North-East, as the installed capacity is smaller. However, in the South, the simulations seem to estimate observed wind power generation quite well and even better after wind power correction. In the North-East, means of monthly wind power generation vary more. Also in the North-East, the means of monthly wind power generation fit the observed mean of monthly wind power generation better after wind power correction and also wind speed correction usually leads to means closer to that of observed wind power generation.

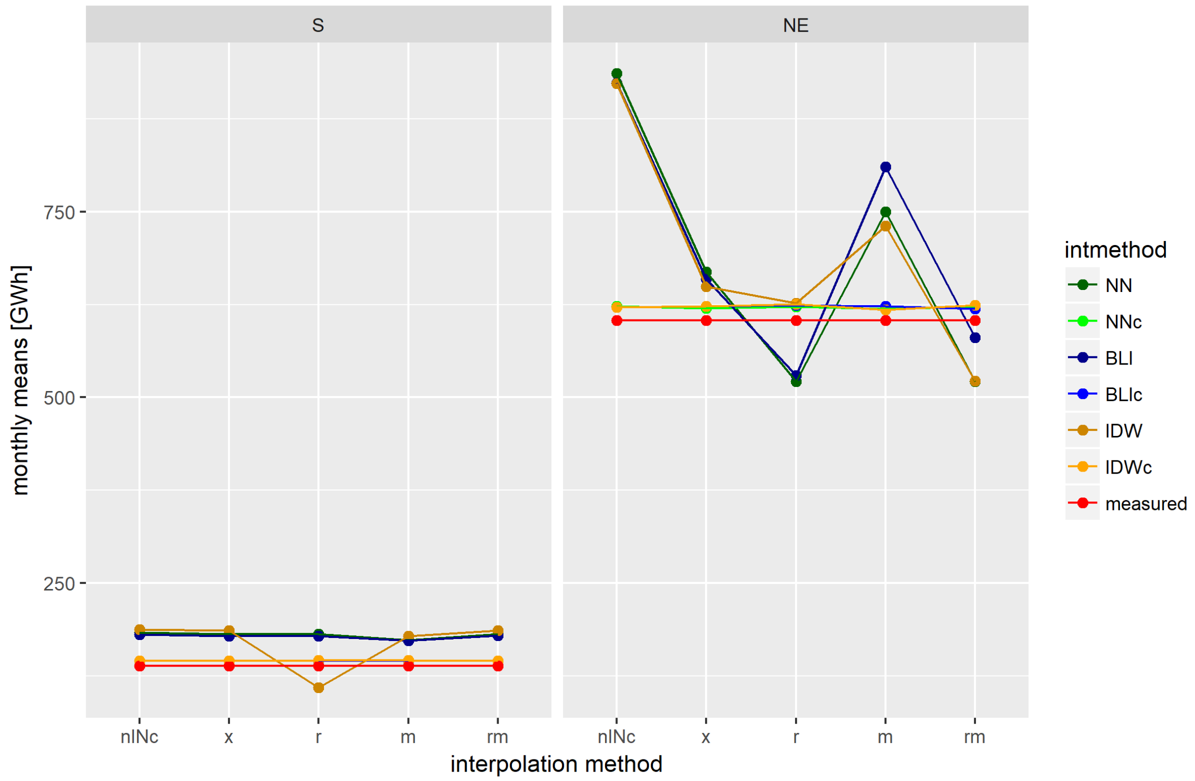

Figure 14: Means of monthly wind power generation in the North-East (NE) and South (S) of Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method and the observed values (own depiction)

Another measure for comparing time series of simulated and observed wind power generation are absolute differences between wind power generations over time, which are depicted in Figure 15 for different interpolation methods. Due to lower wind power generation capacities, the differences are lower in the South than in the North-East. Although differences in the South are quite small already before wind power correction (below 250 GWh), a decrease to about half can be perceived after correction. In the South, the deviation between simulated and observed wind power generation before wind power correction is highest with Inverse Distance Weighting, with maxima at about 250 GWh difference, very closely followed by the other two methods. In the North-East differences between simulated and observed monthly wind power generation are bigger and are also reduced with wind power correction, however not as much as in the South. For the North-East it can be seen that the highest differences between simulation and observed generation occur when using Bilinear Interpolation; Inverse Distance Weighting and the Nearest Neighbour method show similar results. Here, higher differences of up to nearly 1000 GWh can be observed for a few cases, which is reduced to mostly values below 600 GWh by wind power bias correction. Less than 25 % of the differences between simulated and observed wind power generation are above 250 GWh after wind power correction, which is not the case before bias correction.

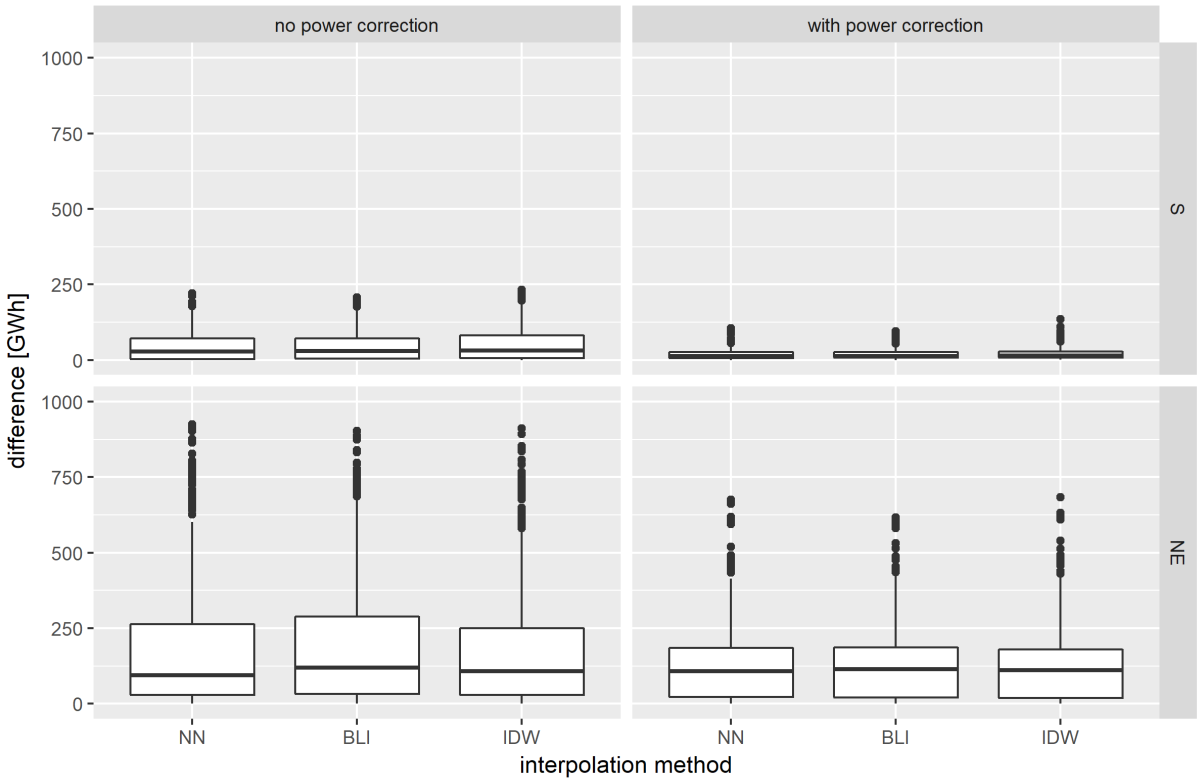

Figure 15: Absolute monthly differences in wind power generation per interpolation method in the North-East (NE) and South (S) of Brazil. On the x-axis different interpolation methods are shown (own depiction)

The following boxplot (Figure 16) also shows differences between simulated and observed monthly wind power generation in the South and North-East of Brazil, but per wind speed correction method. In the South, the smallest differences before wind power correction occur when wind speed correction with adaptation of means is applied; for the other wind speed correction methods no clear differences can be perceived. After wind power correction, the differences are reduced and it is not possible to determine the one with the largest or smallest differences. In the North-East, differences are considerably larger, especially when no wind speed correction is applied, but also when both adaptations are applied during wind speed correction. The other simulations seem to perform not that badly, compared to the South. After wind power correction, it can be observed, that the simulation with wind speed correction and removal of long rows of 0 m/s wind speed results in the lowest differences in monthly wind power generation and also that only the two aforementioned methods with the highest differences improve significantly with wind power correction.

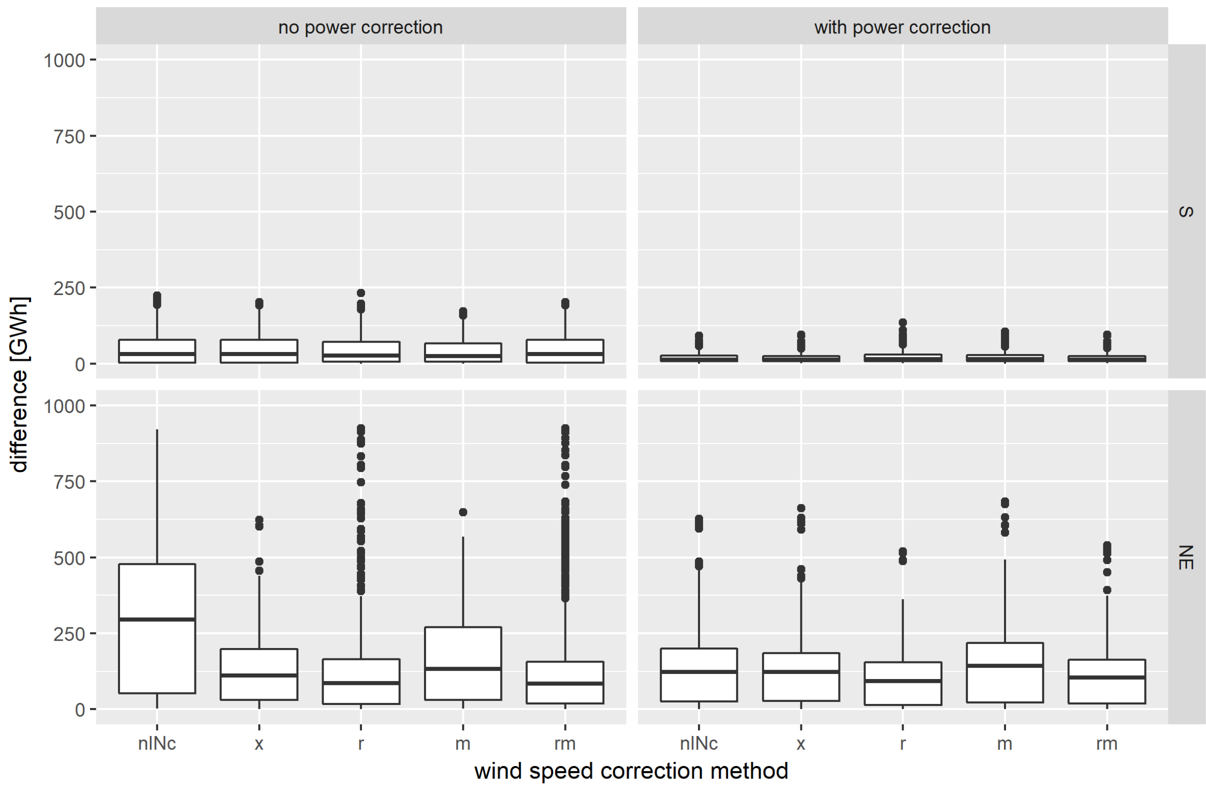

Figure 16: Absolute monthly differences in wind power generation per wind speed correction method in the North-East (NE) and South (S) of Brazil. On the x-axis different wind speed correction methods are shown (own depiction)

1.3 Monthly wind power generation in Brazil

Also for all of Brazil, some additional results are provided. The RMSEs between different simulations and observed wind power generation in Brazil are compared in Figure 17. This shows more clearly, what has already been determined from Table 8: The highest RMSEs are seen for the simulations without bias correction, followed by the simulation with Bilinear Interpolation combined with mean approximation during wind speed correction. The lowest RMSEs can be observed with Inverse Distance Weighting when wind power correction combined with wind speed correction with removal of long rows of 0 m/s wind speed is applied. Compared to RMSEs from the North-East, those of all of Brazil are not significantly higher, but of course the ones in the South are lower, due to small installed capacities.

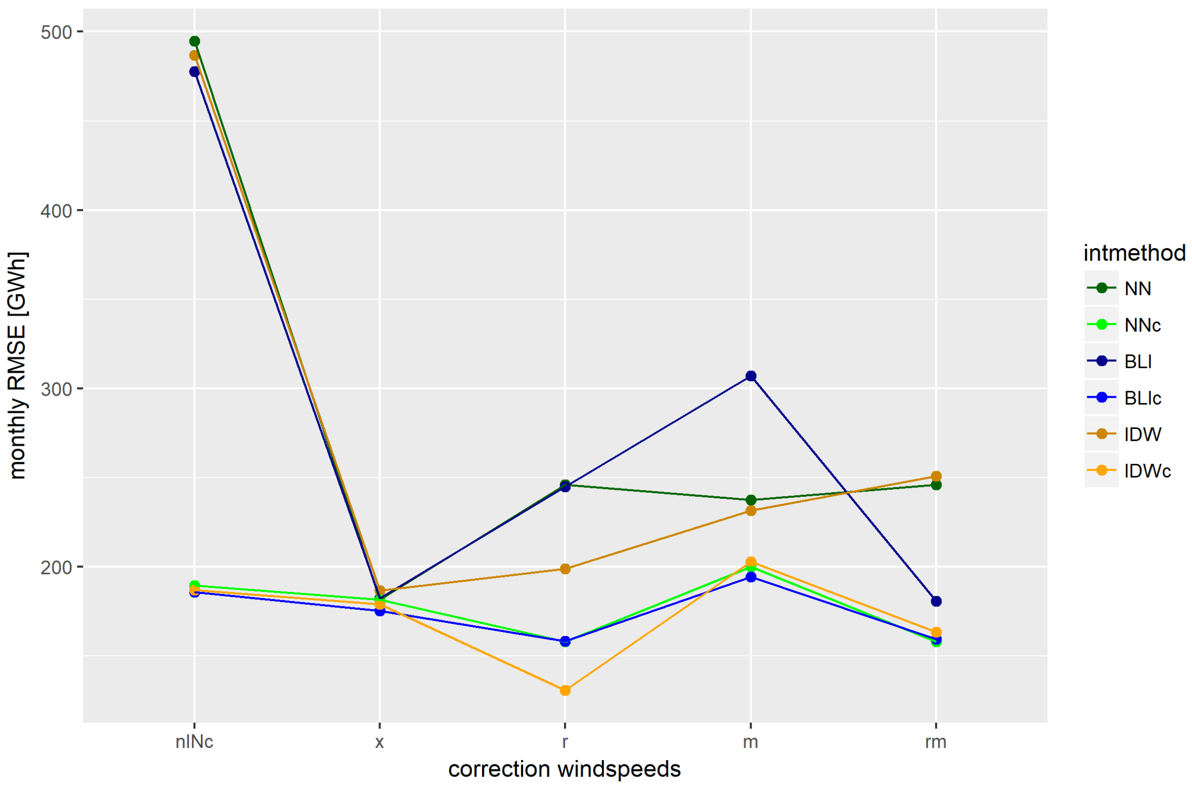

Figure 17: RMSEs between observed and simulated monthly wind power generation in Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method (own depiction)

In addition, the means of simulated and observed monthly wind power generation in Brazil are compared (see Figure 18). It can be clearly seen, that means of simulated wind power generation without any bias correction are much higher than the mean of recorded wind power generation, followed by simulations without wind power correction and wind speed correction with adaptation of means. From the graph it can be seen that the mean of the simulation which is closest to the mean of observed wind power generation, is the one with Inverse Distance Weighting without wind power correction when long rows of 0 m/s wind speed are removed. After wind power correction, the differences to the mean of observed wind power generation are very similar for all interpolation methods.

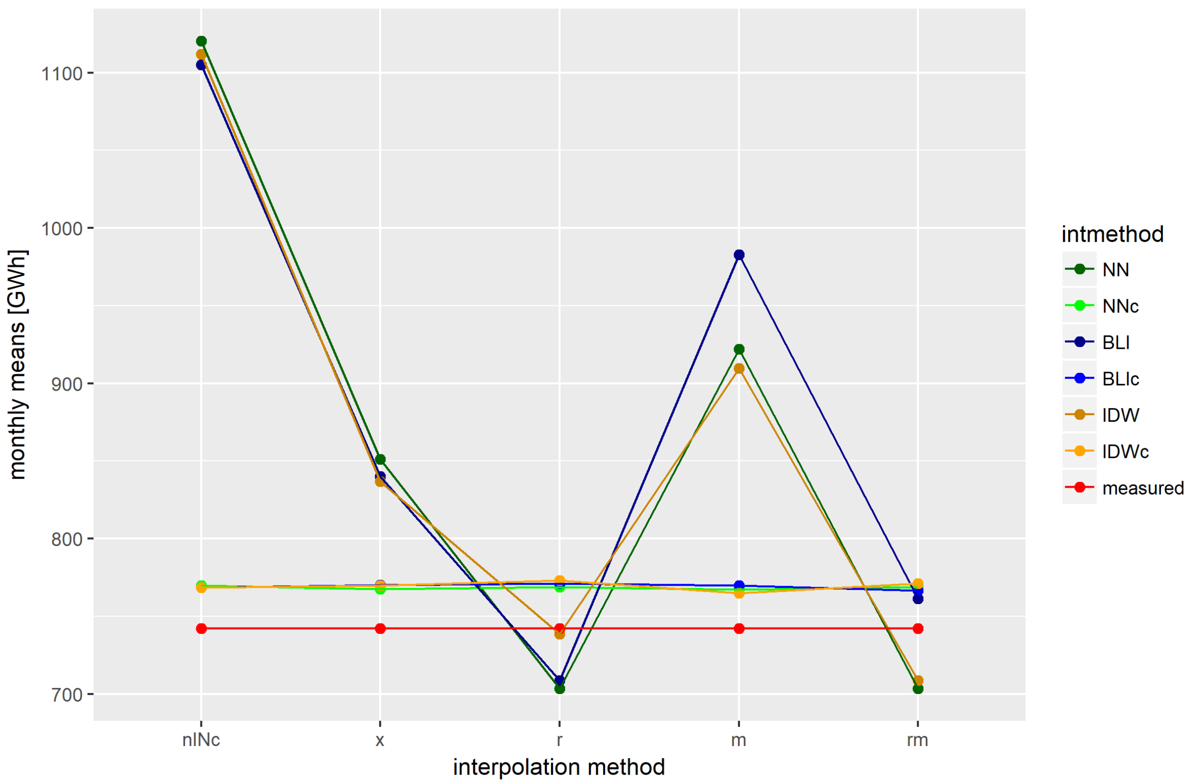

Figure 18: Means of monthly wind power generation in Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method and the observed values (own depiction)

For all of Brazil, also the differences between simulated and observed wind power time series are calculated and depicted in Figure 19 per interpolation method. The largest deviations from observed wind power generation occur when using Bilinear Interpolation, closely followed by the other interpolation methods. Half of the time the simulations over- or underestimate observed monthly wind power generation more than 125 GWh, before as well as after wind power correction. The most obvious improvement happens with Bilinear Interpolation, as before wind power correction more than 50 % of the time differences are higher than 125 GWh but after more than 50 % are below 125 GWh difference. For the other two interpolation methods only for the upper 50 % of deviation from observed monthly wind power generation an improvement can be observed, as the median is at about the same height and only the highest 50 % of differences are observed in a lower range after wind power correction. The lowest differences occur after wind power correction, irrespective of the interpolation method, as more than 50 % of differences are below 125 GWh after wind power correction.

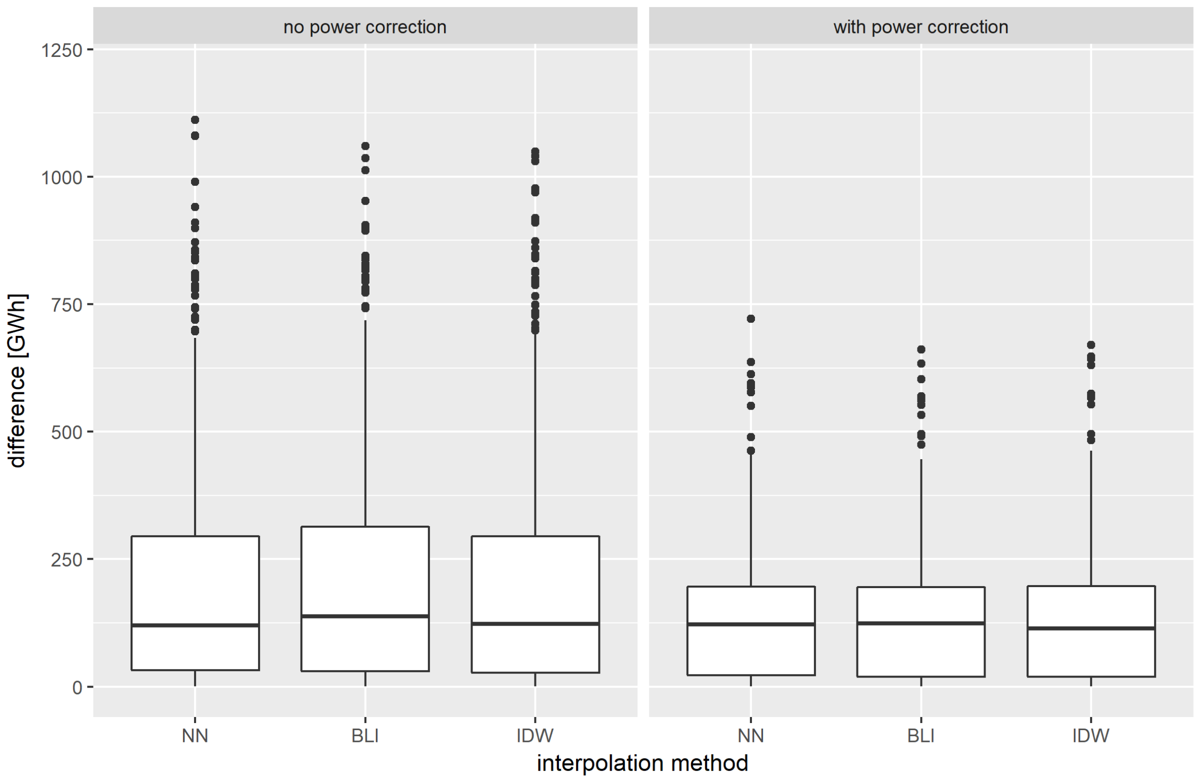

Figure 19: Absolute monthly differences in wind power generation per interpolation method in Brazil. On the x-axis different interpolation methods are shown (own depiction)

Figure 20 shows the absolute differences in monthly wind power generation per wind speed correction, where it can be clearly seen, that differences before wind power correction are highest without wind speed correction, but also higher than with other wind speed correction methods when only adaptation of means is applied. The lowest differences occur when normal wind speed correction or only removal of long rows of 0 m/s wind speed are applied. After wind power correction, the latter method shows the best results (lowest differences) and the simulation differs most from observed monthly wind power generation when applying wind speed correction with mean adaptation. The most significant improvement occurs in the simulation without wind speed correction: Before wind power correction nearly 50 % of the differences are above 375 GWh, and after wind power correction more than 75 % are below 250 GWh.

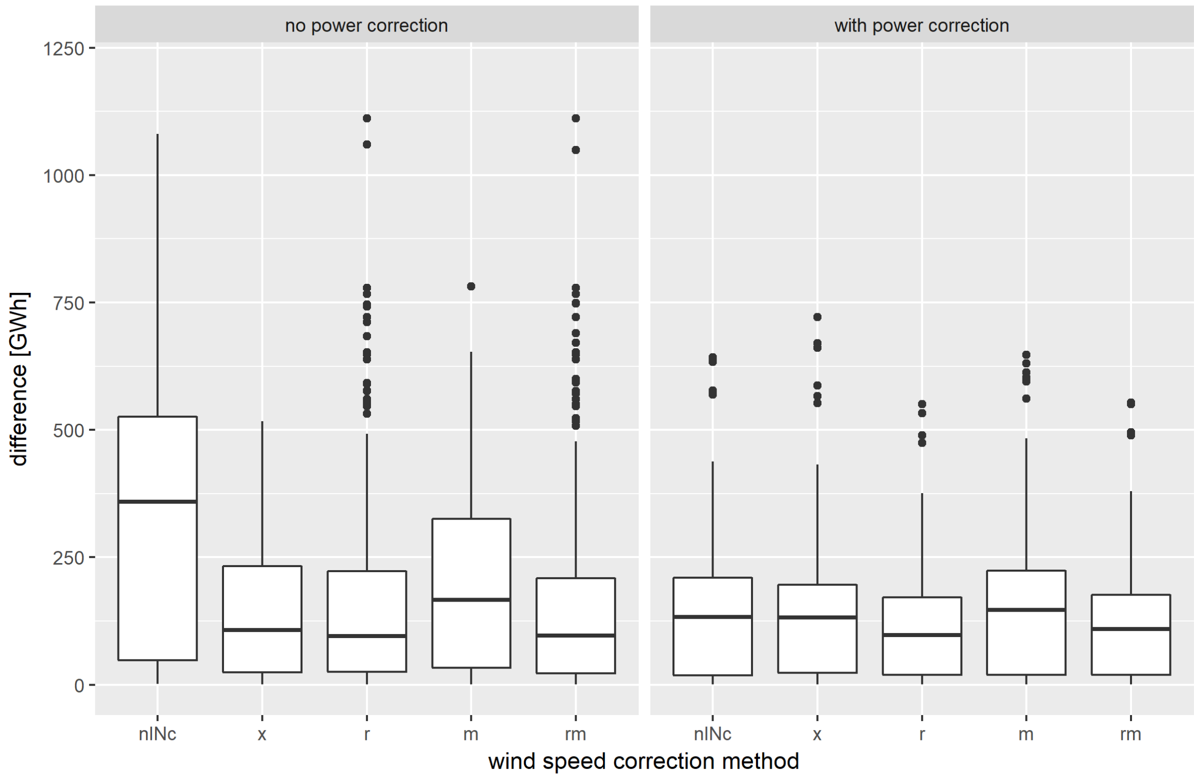

Figure 20: Absolute monthly differences in wind power generation per wind speed correction method in Brazil. On the x-axis different wind speed correction methods are shown (own depiction)

1.4 Daily wind power generation per state

This chapter analyses the results from daily wind power generation on the level of seven states of Brazil. First of all, correlations between simulated and observed daily wind power generation are compared in the left part of Table 23. In most states the correlations are quite high, mostly above 0.9. Only Santa Catarina, Pernambuco and Piaui show lower correlations, the latter two between 0.845 and 0.929 and the first is even lower in a range of 0.692 to 0.746. This is most likely due to the fact that only a few wind power plants are installed in these states (15 in Santa Catarina and 21 in Pernambuco), as they are comparatively small (see Figure 2) and thus the installed capacities are low, too (242.5 MW in Santa Catarina and 596.13 MW in Pernambuco, compared to 2270 MW in Bahia). In some cases (for example in Ceará or Piaui) wind power correction leads to better correlations, but in others the increase is only minimal. Wind speed correction with monthly mean approximation sometimes results in lower correlations, especially in Ceará, and it can also be observed that omitting wind speed correction does not necessarily yield worse correlations; mainly in Bahia and Pernambuco slightly lower correlations are observed when no wind speed correction is applied.

In the middle of Table 23, relative RMSEs of daily wind power generation in seven states of Brazil are displayed. In most cases the relative RMSEs are quite small between 0.2 and 0.6, with a few exceptions: In Rio Grande do Norte, before wind power correction relative RMSEs are in a range of 0.583 to 1.035 and after correction they are lower, but still higher than in most of the other states at around 0.5 for most values. Also in Santa Catarina higher relative RMSEs are seen especially for Inverse Distance Weighting. In Ceará, values of more than 1 are achieved if no bias correction is performed. The other relative RMSEs are below 1 and often even below 0.5. Wind speed correction not always leads to lower relative RMSEs.

Table 23: Correlations, relative RMSEs and relative absolute differences in means of monthly wind power generation time series for seven states of Brazil (own representation)

		correlations					relative RMSEs					rel. abs. difference in means				
		x	r	m	rm	nINc	x	r	m	rm	nINc	x	r	m	rm	nINc
Bahia	NN	0.962	0.962	0.956	0.962	0.960	0.270	0.270	0.410	0.270	0.332	0.040	0.040	0.160	0.040	0.184
	BLI	0.962	0.962	0.962	0.961	0.957	0.287	0.287	0.284	0.268	0.311	0.038	0.038	0.122	0.055	0.139
	IDW	0.961	0.947	0.952	0.961	0.955	0.297	0.769	0.478	0.297	0.309	1.206	0.665	0.222	0.055	0.124
	NNc	0.966	0.966	0.959	0.966	0.964	0.255	0.255	0.290	0.255	0.262	0.028	0.028	0.031	0.028	0.028
	BLIc	0.966	0.966	0.965	0.964	0.962	0.252	0.252	0.261	0.255	0.267	0.029	0.029	0.029	0.028	0.028
	IDWc	0.965	0.950	0.957	0.965	0.961	0.255	0.342	0.301	0.255	0.272	0.029	0.027	0.029	0.029	0.028
Ceará	NN	0.917	0.917	0.854	0.917	0.934	0.350	0.350	0.446	0.350	1.061	0.006	0.006	0.043	0.006	0.878
	BLI	0.919	0.919	0.915	0.920	0.935	0.337	0.337	0.351	0.336	1.084	0.020	0.020	0.106	0.014	0.878
	IDW	0.928	0.951	0.867	0.928	0.933	0.355	0.465	0.433	0.355	1.089	0.100	0.372	0.010	0.100	0.882
	NNc	0.943	0.943	0.923	0.943	0.953	0.283	0.283	0.341	0.283	0.259	0.067	0.067	0.067	0.067	0.069
	BLIc	0.946	0.946	0.933	0.946	0.955	0.274	0.273	0.311	0.275	0.252	0.070	0.070	0.070	0.070	0.070
	IDWc	0.942	0.961	0.923	0.942	0.955	0.283	0.243	0.342	0.283	0.255	0.069	0.069	0.067	0.069	0.070
Pernambuco	NN	0.858	0.858	0.868	0.858	0.853	0.339	0.339	0.309	0.339	0.337	0.103	0.102	0.016	0.102	0.002
	BLI	0.855	0.855	0.888	0.852	0.848	0.347	0.346	0.282	0.336	0.342	0.121	0.120	0.037	0.035	0.010
	IDW	0.852	0.865	0.868	0.852	0.845	0.350	0.336	0.329	0.350	0.345	0.120	0.126	0.133	0.120	0.015
	NNc	0.893	0.893	0.891	0.893	0.885	0.339	0.339	0.306	0.339	0.325	0.102	0.102	0.099	0.102	0.097
	BLIc	0.890	0.890	0.900	0.885	0.882	0.347	0.347	0.300	0.331	0.330	0.102	0.102	0.105	0.099	0.097
	IDWc	0.889	0.887	0.894	0.889	0.881	0.347	0.354	0.315	0.346	0.332	0.102	0.102	0.107	0.102	0.097
Piaui	NN	0.862	0.864	0.918	0.864	0.928	0.420	0.416	0.308	0.416	0.416	0.166	0.159	0.110	0.159	0.260
	BLI	0.852	0.854	0.917	0.921	0.923	0.441	0.436	0.308	0.356	0.405	0.182	0.175	0.101	0.180	0.243
	IDW	0.865	0.925	0.917	0.866	0.925	0.456	0.392	0.301	0.451	0.392	0.229	0.229	0.083	0.222	0.229
	NNc	0.920	0.920	0.927	0.920	0.928	0.323	0.322	0.303	0.322	0.313	0.132	0.132	0.129	0.132	0.135
	BLIc	0.917	0.917	0.927	0.929	0.927	0.323	0.322	0.302	0.315	0.311	0.127	0.127	0.126	0.137	0.131
	IDWc	0.918	0.928	0.927	0.918	0.928	0.325	0.309	0.302	0.324	0.309	0.131	0.131	0.126	0.131	0.131
Rio Grande do Norte	NN	0.966	0.962	0.966	0.962	0.966	0.581	0.755	1.035	0.755	1.035	0.386	0.257	0.737	0.257	0.737
	BLI	0.965	0.961	0.965	0.953	0.965	0.583	0.687	0.876	0.610	1.007	0.385	0.185	0.715	0.081	0.715
	IDW	0.965	0.975	0.965	0.963	0.965	0.607	1.000	1.022	0.645	1.022	0.410	0.498	0.723	0.147	0.723
	NNc	0.967	0.971	0.965	0.971	0.965	0.515	0.423	0.513	0.423	0.513	0.027	0.033	0.027	0.033	0.027
	BLIc	0.966	0.969	0.965	0.971	0.965	0.523	0.451	0.514	0.444	0.514	0.027	0.031	0.027	0.031	0.027
	IDWc	0.966	0.979	0.965	0.969	0.965	0.525	0.337	0.514	0.458	0.514	0.027	0.042	0.027	0.031	0.027
Rio Grande do Sul	NN	0.953	0.953	0.938	0.953	0.952	0.645	0.645	0.610	0.645	0.672	0.240	0.240	0.174	0.240	0.255
	BLI	0.953	0.953	0.930	0.953	0.953	0.587	0.587	0.570	0.588	0.609	0.181	0.181	0.129	0.182	0.193
	IDW	0.951	0.912	0.930	0.951	0.950	0.605	0.829	0.571	0.605	0.624	0.183	0.426	0.121	0.183	0.194
	NNc	0.953	0.953	0.939	0.953	0.952	0.435	0.435	0.497	0.435	0.441	0.028	0.028	0.033	0.028	0.030
	BLIc	0.953	0.953	0.932	0.953	0.953	0.442	0.442	0.508	0.442	0.447	0.029	0.029	0.031	0.029	0.030
	IDWc	0.951	0.913	0.931	0.951	0.951	0.455	0.603	0.514	0.455	0.460	0.030	0.034	0.031	0.030	0.031
Santa Catarina	NN	0.696	0.696	0.692	0.696	0.692	1.008	1.008	1.024	1.008	1.024	0.631	0.631	0.646	0.631	0.646
	BLI	0.736	0.736	0.733	0.733	0.733	1.360	1.360	1.376	1.376	1.376	1.000	1.000	1.014	1.014	1.014
	IDW	0.741	0.738	0.738	0.741	0.738	1.782	1.797	1.797	1.782	1.797	1.478	1.491	1.491	1.478	1.491
	NNc	0.701	0.701	0.697	0.701	0.697	0.521	0.521	0.523	0.521	0.523	0.067	0.067	0.067	0.067	0.067
	BLIc	0.741	0.741	0.738	0.738	0.738	0.478	0.478	0.481	0.481	0.481	0.064	0.064	0.064	0.064	0.064
	IDWc	0.746	0.743	0.743	0.746	0.743	0.446	0.448	0.448	0.446	0.448	0.064	0.064	0.064	0.064	0.064

Figure 21 shows the RMSEs of daily wind power generation per state. In two states, Ceará and Rio Grande do Norte, the RMSEs are comparatively high when wind speed correction is not performed. Rio Grande do Norte in general shows quite high RMSEs for not wind power corrected simulated wind power generation, which is also due to higher capacities installed there. Also in Bahia a high RMSE occurs when Inverse Distance Weighting without wind power correction and with wind speed correction combined with removal of long rows of 0 m/s wind speed is applied. It also stands out that after application of wind power correction, RMSEs are lower in all states. The lowest RMSEs occur in Santa Catarina, probably due to the small capacity installed in that state.

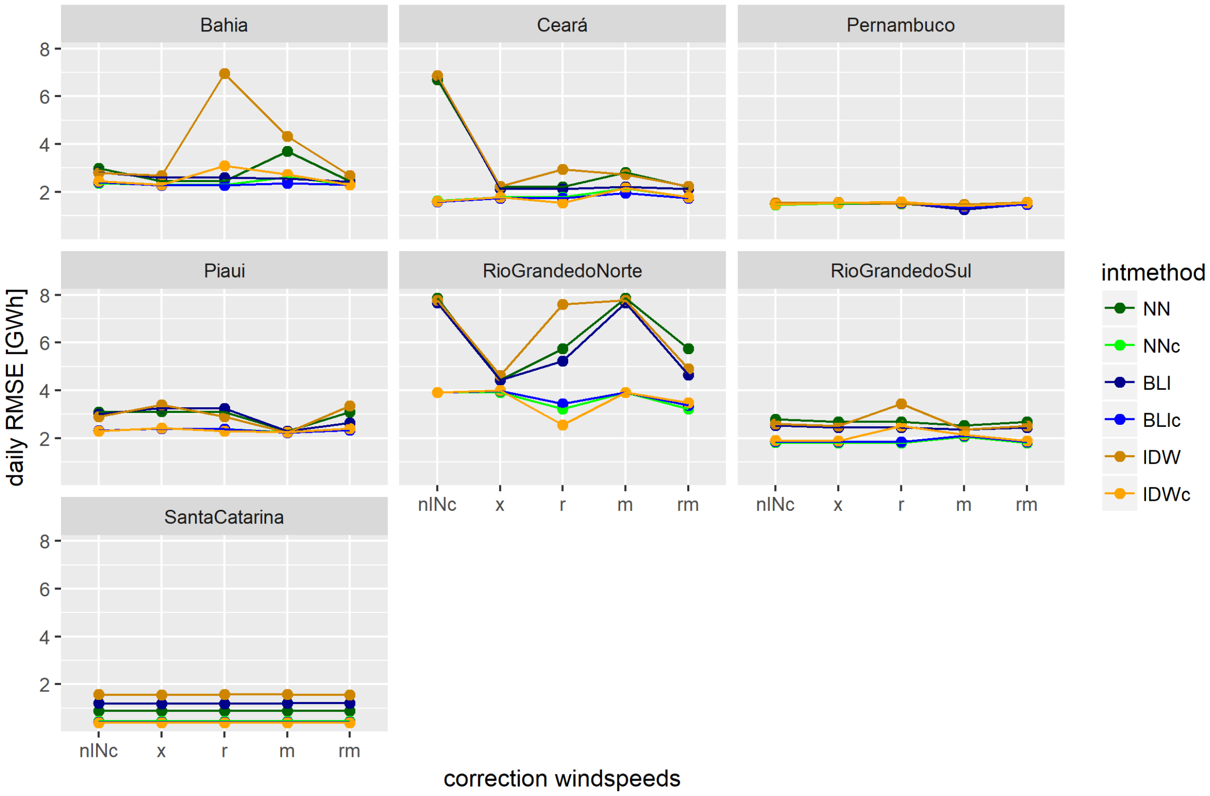

Figure 21: RMSEs between observed and simulated daily wind power generation in seven states of Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method (own depiction)

Another measure which is used to compare simulated to observed wind power generation is the mean daily wind power. Figure 22 shows the daily means for each wind power and wind speed correction method and also per each state. From the graphs it can be determined that in general simulated daily wind power generation either over- or underestimates observed wind power generation, depending on the simulation. It can be observed, that for each state, the means are generally closer to the mean of observed daily wind power generation after application of wind power correction. The highest deviations can be seen in Ceará as well as Rio Grande do Norte, and the highest observed mean wind power generation is in Bahia, which is why the higher deviation in means cannot be linked to a higher capacity for certain. The lowest observed mean wind power generation can be seen in Santa Catarina and as it is so low, the deviation of means of simulated wind power generation is relatively high for simulations that are not wind power corrected; this can be determined even better in the right part of Table 23, which shows the absolute values of relative deviations of means of different simulations from observed wind power generation. In the special case of Santa Catarina, it may seem unlikely that there is no difference in means between the different wind speed correction methods, but this is due to the fact, that the few wind power plants that are installed in this state are too far away from the nearest wind speed measurement station or correlations are too low for wind speed correction to be applied. The absolute value of relative normed daily differences in means for other states are usually lower, especially in Piaui, Pernambuco and Rio Grande do Sul.

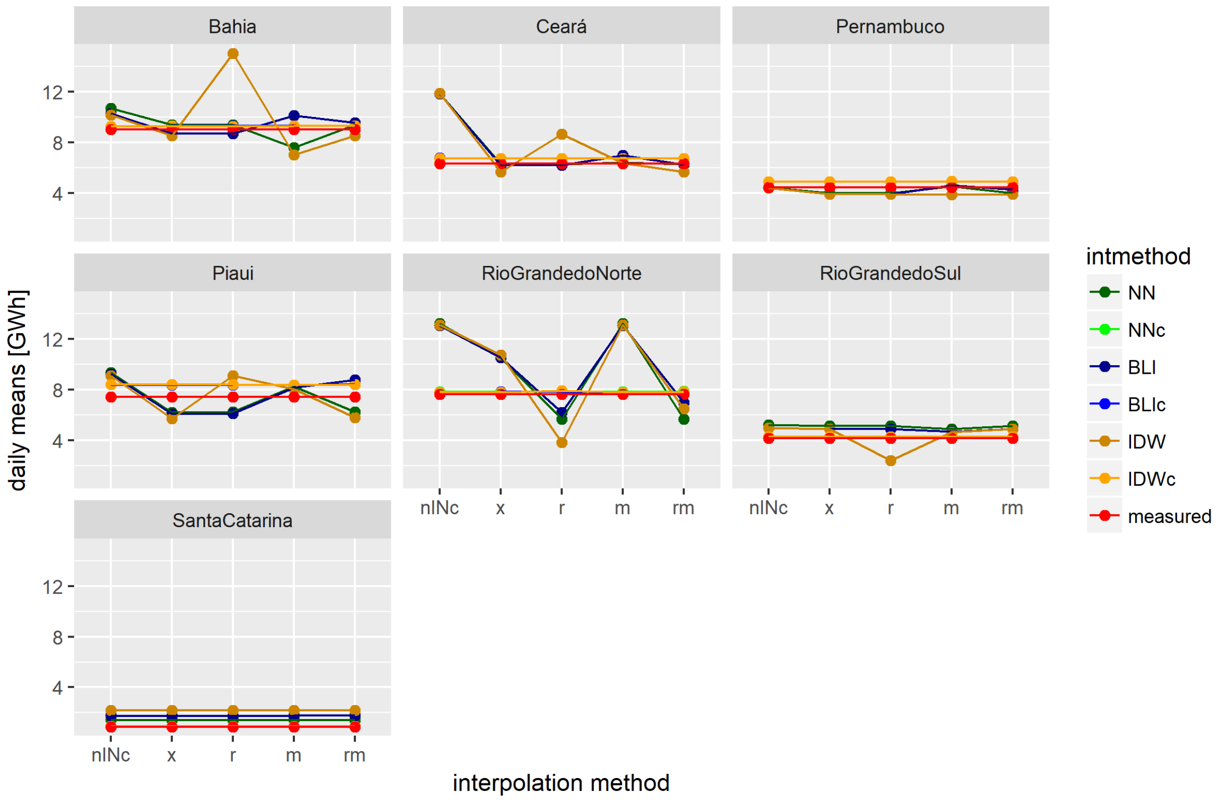

Figure 22: Means of daily wind power generation in in seven states of Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method and the observed values (own depiction)

After looking at the means, RMSEs and correlations of simulated and observed wind power generation, also the differences between these are examined. Figure 23 shows the absolute differences of simulated daily wind power generation and observed daily wind power generation per interpolation method before and after wind power correction for all states. There are no striking differences in the bias of simulated and observed wind power data between interpolation methods, however, the results from Inverse Distance Weighting show slightly larger differences than the other methods, followed by the Nearest Neighbour method and then by Bilinear Interpolation, which seems to fit observed wind power generation best. The disparities between interpolation methods decrease after applying wind power bias correction, when differences are reduced to less than 5 GWh daily, disregarding outliers.

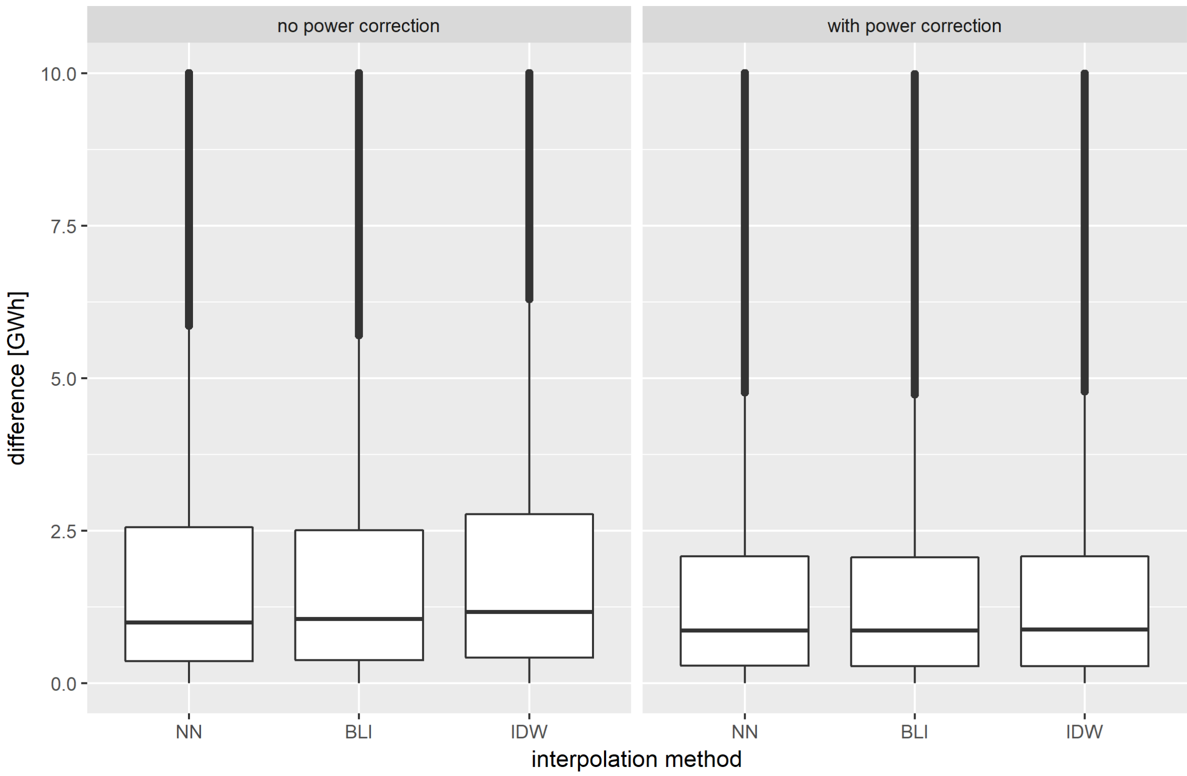

Figure 23: Absolute daily differences in wind power generation per interpolation method in seven states of Brazil. On the x-axis different interpolation methods are shown (own depiction)

The absolute daily differences between simulated and observed wind power generation per wind speed correction method are shown in Figure 24. The simulation deviates most from the observed wind power generation before wind power correction if no wind speed correction or wind speed correction at least with mean approximation are applied. After wind power correction, which decreases differences for all wind speed correction methods, results are very similar, with only the wind speed corrected simulation with mean approximation standing out as higher. From the remaining simulations, the one with removal of long rows of 0 m/s wind speed shows the best results regarding the differences between simulated and observed wind power generation time series, very closely followed by the other simulations.

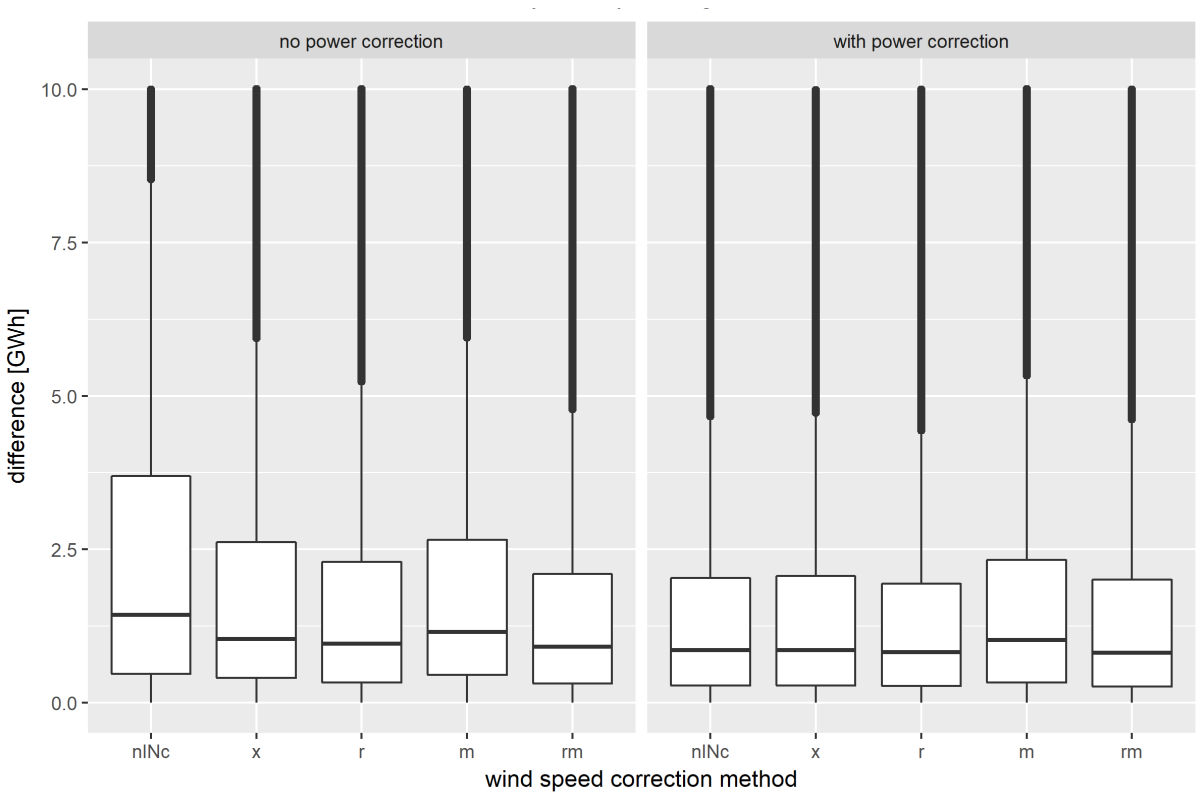

Figure 24: Absolute daily differences in wind power generation per wind speed correction method in seven states of Brazil. On the x-axis different wind speed correction methods are shown (own depiction)

Above results show, that wind power correction significantly improves the quality of simulations: RMSEs and differences to observed data become smaller, the means of simulated time series are usually closer to the means of observed wind power generation and correlations rise after correction. No clear differences between single states can be discerned, only Santa Catarina issues comparably bad results regarding correlations and means.

1.5 Daily wind power generation per subsystem

Additionally to the results from daily wind power generation on subsystem level in section 3.3, some supplementary graphs are displayed in this chapter. The first parameter that is compared are RMSEs of daily wind power generation in the North-East and South (Figure 25). In the South, RMSEs are lower (always below 5 GWh) than in the North-East, which is due to smaller installed capacities there. However, in the North-East the RMSEs are not very high either, usually below 10 GWh, except with Bilinear Interpolation combined with mean approximation or when no type of bias correction is applied: Then RMSEs of more than 15 GWh are reached. For most of the simulations, RMSEs of wind power corrected time series are lower than those of uncorrected wind power generation, with one exception in the North-East: When Inverse Distance Weighting is applied together with wind speed correction and removal of long rows of 0 m/s wind speed, the RMSE is lower than that of the wind power corrected Bilinear Interpolation method, but still higher than that of the wind power corrected Inverse Distance Weighted wind power generation with wind power correction. Usually the lowest RMSEs are obtained with the wind power corrected Nearest Neighbour method, except in the North-East when mean adaptation is applied during wind speed correction; then Bilinear Interpolation yields the lowest RMSEs.

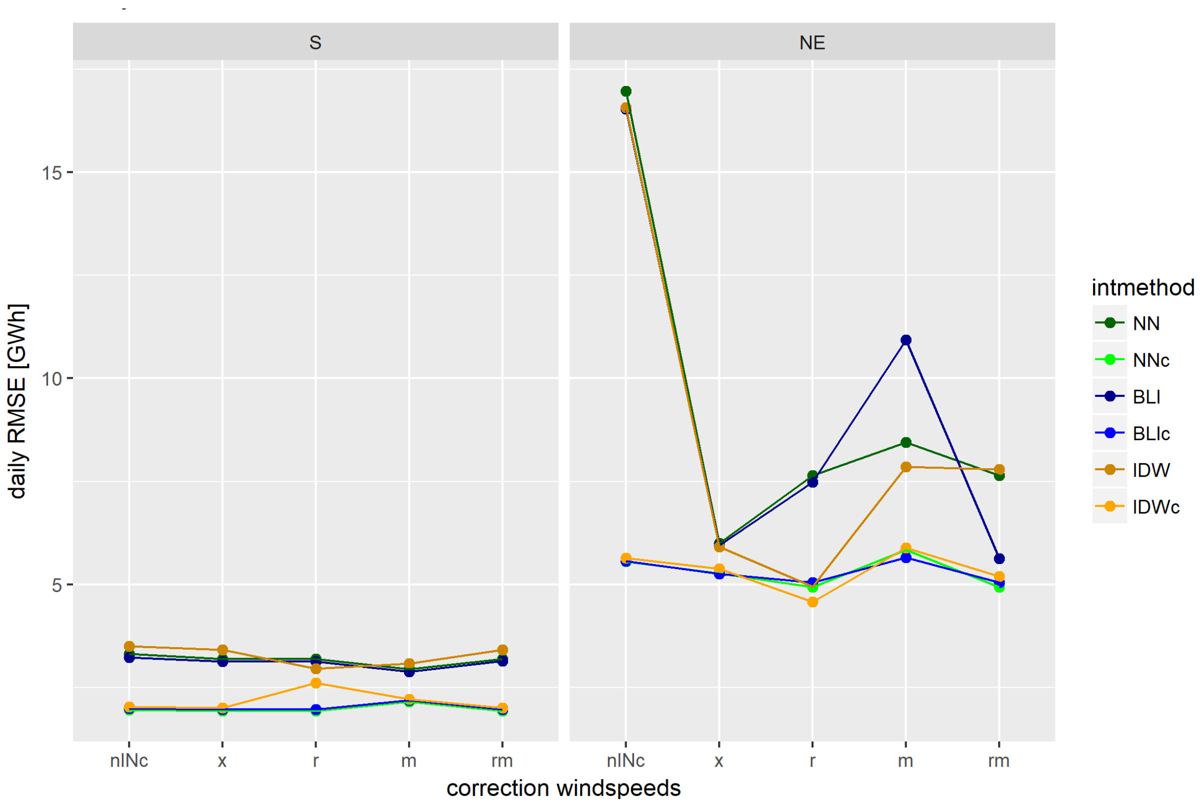

Figure 25: RMSEs between observed and simulated daily wind power generation in the North-East (NE) and South (S) of Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method (own depiction)

In Figure 26, the means of simulated and observed daily wind power generation in the South and North-East of Brazil can be seen. In the South, simulated as well as observed wind power generation is quite low, due to small installed capacities. The simulation always overestimates observed wind power generation or is about the same, except for the Inverse Distance Weighting method when long rows of 0 m/s wind speed are removed during wind speed correction. In the North-East, simulations also mostly overestimate observed wind power generation, only when at least long rows of 0 m/s wind speed are removed during wind speed correction, sometimes observed daily wind power generation is underestimated, but closest to the mean of observed daily wind power. The largest deviation from observed daily wind power generation occurs when no wind speed correction is performed.

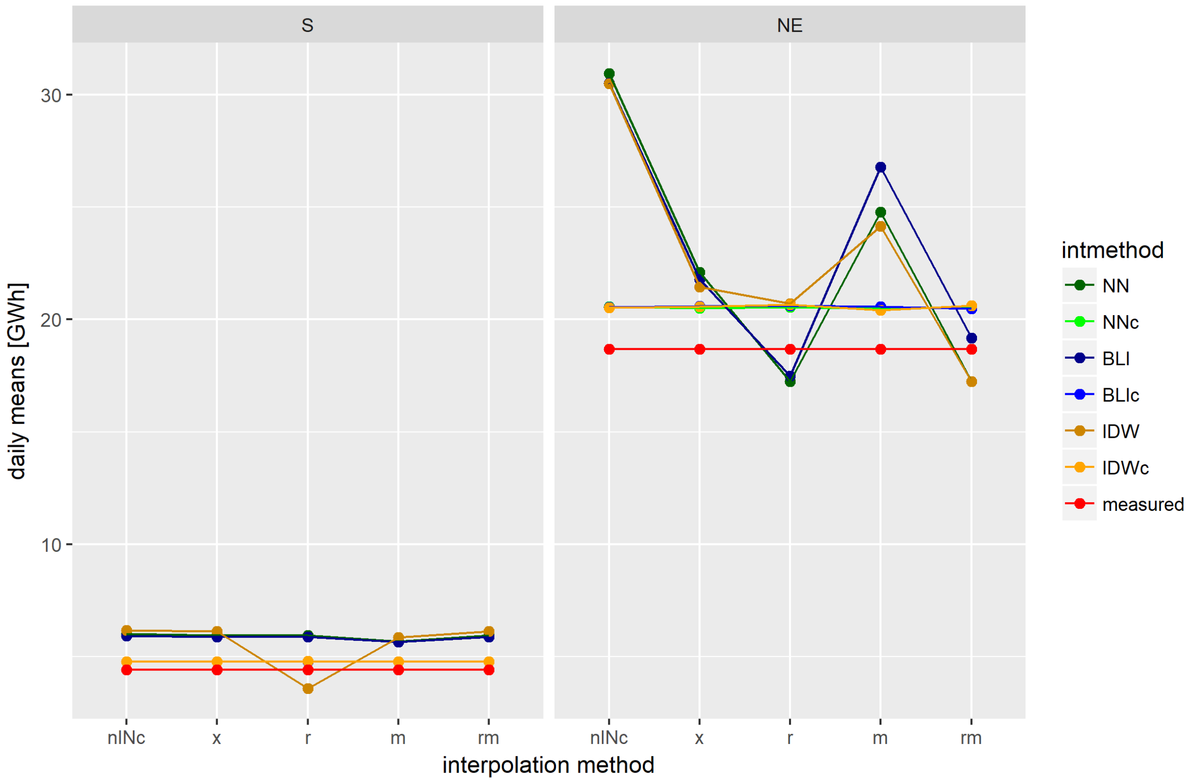

Figure 26: Means of daily wind power generation in the North-East (NE) and South (S) of Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method and the observed values (own depiction)

For the daily wind power generation in the North-East and South of Brazil, also differences between simulated and observed time series are compared. In Figure 27, which compares absolute differences per interpolation method, it is visible, that wind power correction reduces differences between simulated and observed wind power generation. In the South, before as well as after wind power correction, the results of different interpolations in the South are about the same (apart from a few outliers). The absolute differences are bigger in the North-East, due to higher wind power generation there. After wind power correction, differences are reduced and nearly 75 % instead of only about 50 % are below 5 GWh difference in daily wind power generation. In the North-East, no variation in differences between the three interpolation methods can be noticed, apart from the median of absolute differences of the Bilinear Interpolation being slightly higher than from the other interpolation methods, before as well as after wind power correction.

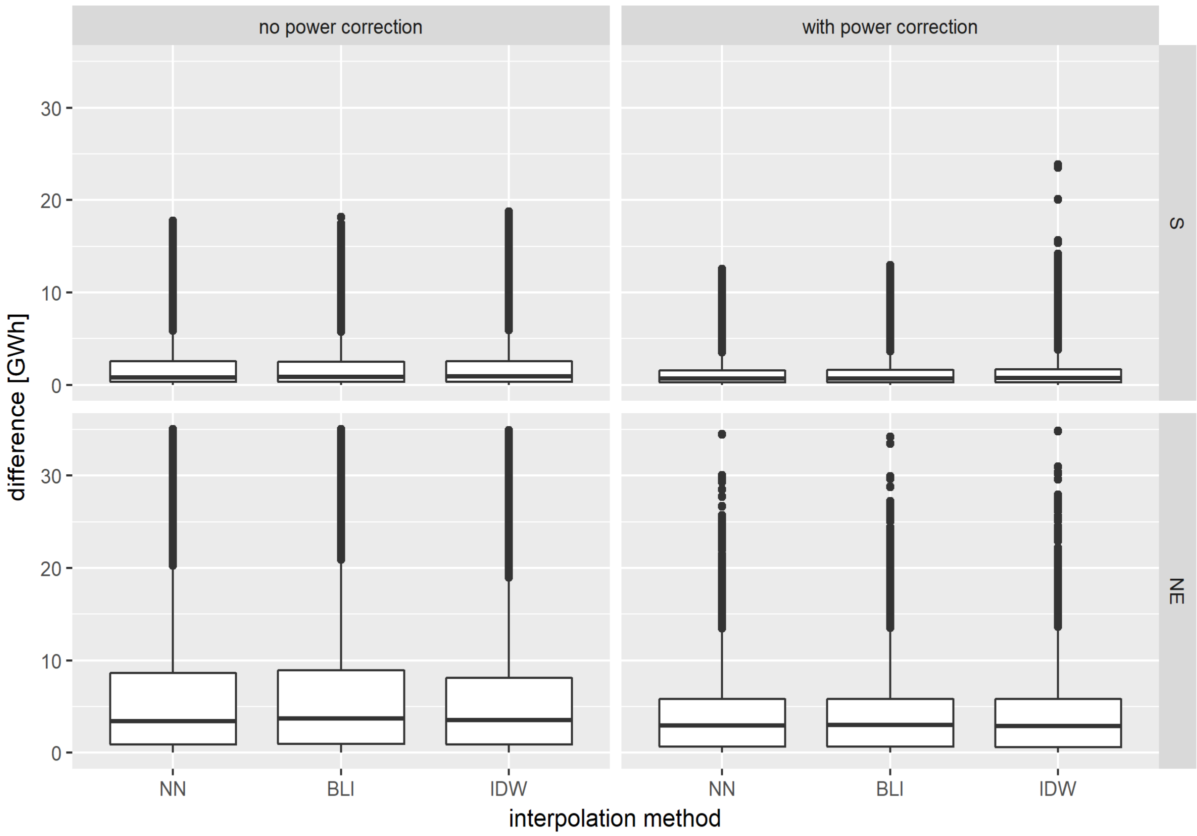

Figure 27: Absolute daily differences in wind power generation per interpolation method in the North-East (NE) and South (S) of Brazil. On the x-axis different interpolation methods are shown (own depiction)

The absolute differences between observed and simulated daily wind power generation in the North-East and South of Brazil are also displayed per wind speed correction method in Figure 28. Due to lower installed capacities in the South, differences there are smaller than in the North-East. In the South, differences are about the same for different wind speed correction methods, before as well as after wind power correction. A small decrease in absolute differences between observed and simulated wind power time series can be seen after wind power correction. In the North-East, clearer gaps between absolute differences of various wind speed correction methods are visible: The most significant deviation from observed wind power generation is observed when no bias correction is performed with differences of more than 40 GWh, followed by only wind speed correction with adaptation of means with up to more than 20 GWh absolute difference in daily wind power generation. Before wind power correction, the simulations that seem to be best adapted to observed daily wind power generation in terms of differences are the ones where wind speed correction with at least removal of long rows of 0 m/s wind speed is performed. After wind power correction, some improvement can be perceived, except for the two latter methods, where the boxplots look about the same. However, these methods still are the ones which seem to deliver the best results in terms of differences. The most significant improvement is observed when no wind speed correction is performed.

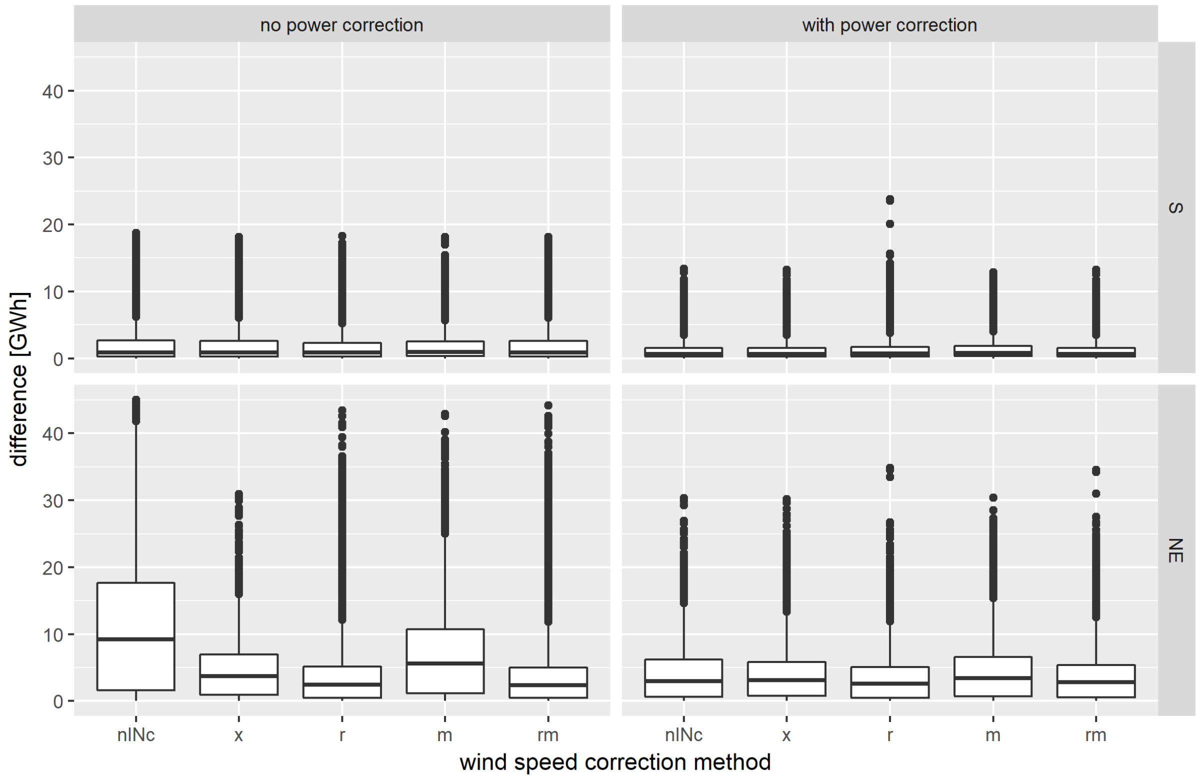

Figure 28: Absolute daily differences in wind power generation per wind speed correction method in the North-East (NE) and South (S) of Brazil. On the x-axis different wind speed correction methods are shown (own depiction)

1.6 Daily wind power generation in Brazil

This section provides some additional results from the comparison of daily wind power generation time series in Brazil. First of all, RMSEs between observed wind power generation and different simulations are compared in Figure 29. The highest values, occur when no bias correction is applied. Furthermore, it stands out that after wind power correction, RMSEs are always lower than before, but there are hardly any differences in RMSE after wind power correction. The lowest RMSE is achieved when performing wind power correction as well as wind speed correction with removal of long rows of 0 m/s wind speed. Whether wind power correction is applied or not, methods with removal of long rows of 0 m/s wind speed during wind speed correction yield better results in terms of RMSEs, compared to other methods.

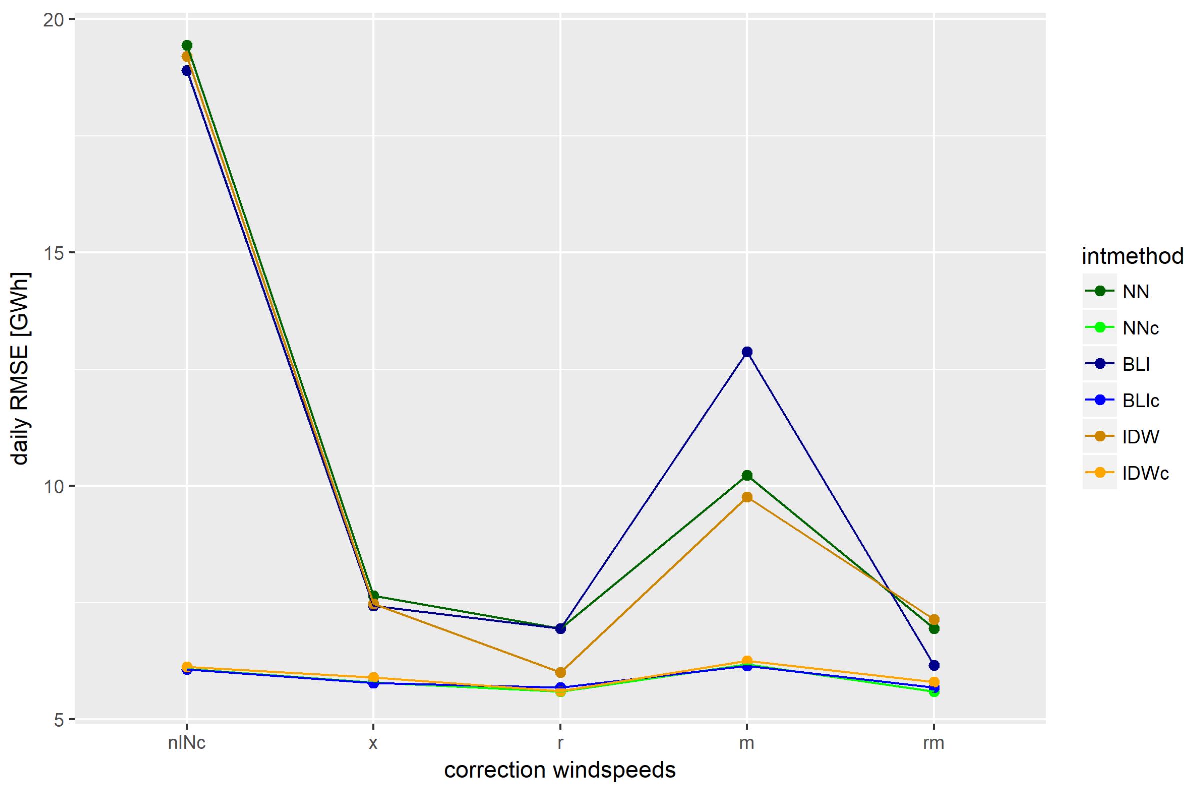

Figure 29: RMSEs between observed and simulated daily wind power generation in Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method (own depiction)

Figure 30 compares means of observed and simulated daily wind power generation. For most cases, observed wind power generation is overestimated. After wind power correction means of simulated daily wind power generation in Brazil are about the same for all methods, but not necessarily closer to the mean of observed daily wind power generation. The highest means and therefore highest overestimations occur when no bias correction is applied. The means of simulated daily wind power generation closest to the means of observed wind power are observed when the Nearest Neighbour method is combined with wind speed correction and at least removal of long rows of 0 m/s wind speed, with Bilinear Interpolation with wind speed correction and removal of long rows of 0 m/s wind speed or with the Inverse Distance Weighting method when both adaptations are applied during wind speed correction, in any of these cases when no wind power correction is applied.

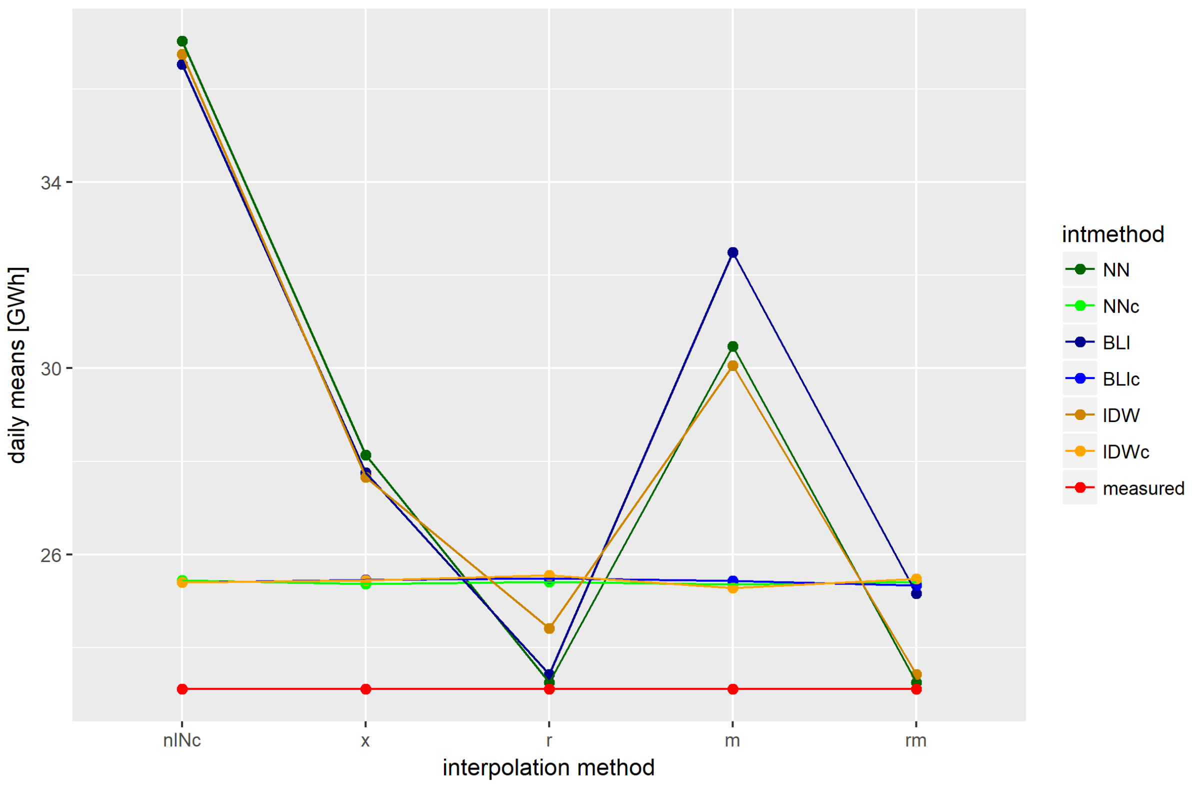

Figure 30: Means of daily wind power generation in Brazil. On the x-axis different wind speed correction methods are shown. The colours indicate the interpolation method and the observed values (own depiction)

In the comparison of absolute differences between observed and simulated daily wind power generation in Brazil (see Figure 31) it can be observed, that simulations with distinct interpolation methods hardly differ in their differences to observed daily wind power generation, before as well as after wind power correction. Nevertheless, Bilinear Interpolation yields slightly higher differences than other interpolation methods before wind power correction. Before bias correction, differences range up to about 25 GWh (disregarding outliers) with medians of about 8 GWh, which is improved to maximum differences of around 15 GWh (not considering outliers) with medians of about 3 GWh. Furthermore, the figure shows that before wind power correction about 75 % of the differences between observed and simulated daily wind power generation are below 10 GWh, which is reduced to at least 75 % of differences being below 7.5 GWh after wind power correction.

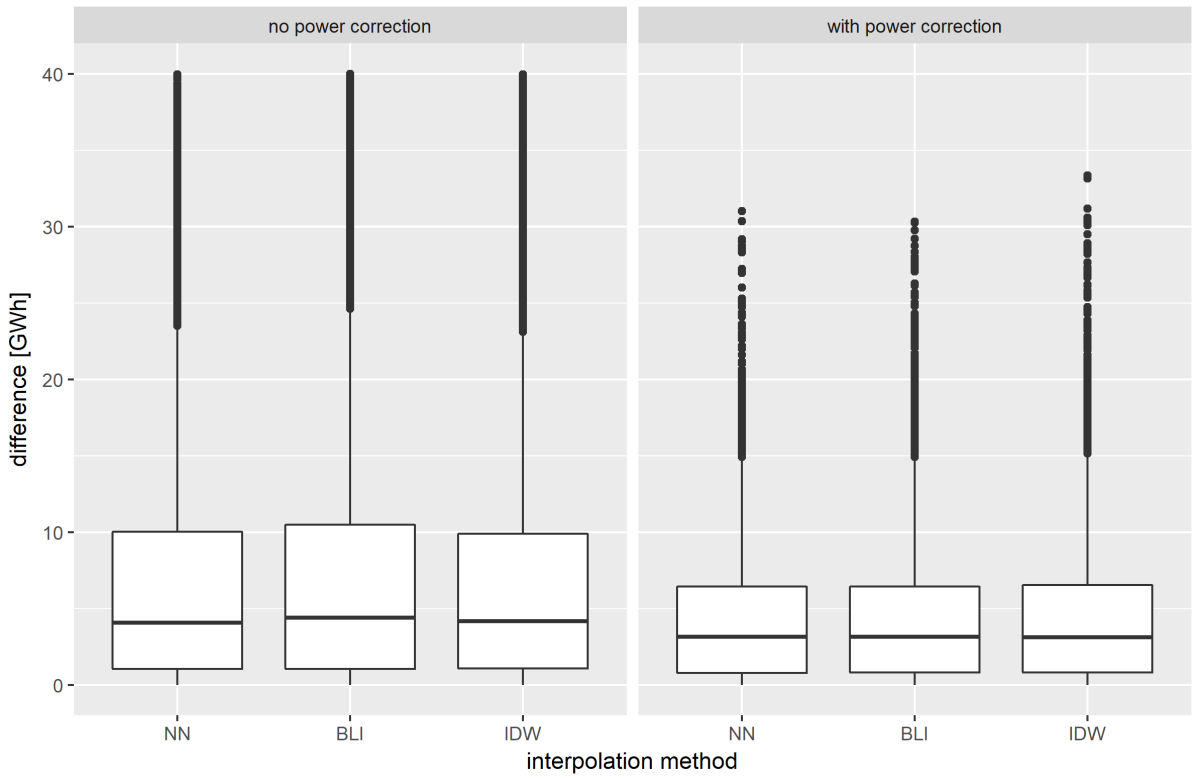

Figure 31: Absolute daily differences in wind power generation per interpolation method in Brazil. On the x-axis different interpolation methods are shown (own depiction)

Absolute differences are also compared between different wind speed correction methods in Figure 32. For most wind speed correction methods, differences can be reduced by wind power correction. When at least long rows of 0 m/s wind speed are removed, however, differences between simulated and observed daily wind power generation in Brazil seem not to change. The biggest improvement in terms of differences is observed when no wind speed correction is performed: Before wind power correction, more than 50 % of the daily wind power generation show differences of more than 20 GWh, whereas after wind power correction the upper 50 % are only above 7 GWh. The simulations which seem to perform best regarding differences in daily wind power generation are the ones where at least long time series of 0 m/s wind speed are removed during wind speed correction, however, only before wind power correction differences of simulated and observed wind power generation differ significantly between different methods. The smallest differences are obtained when wind speed correction is combined at least with removal of long rows of 0 m/s wind speed, whereas only approximation of means results in the highest differences after wind power correction. Results are similar to those of the analysis of daily wind power generation in subsystems, but differ from those in the monthly analysis.

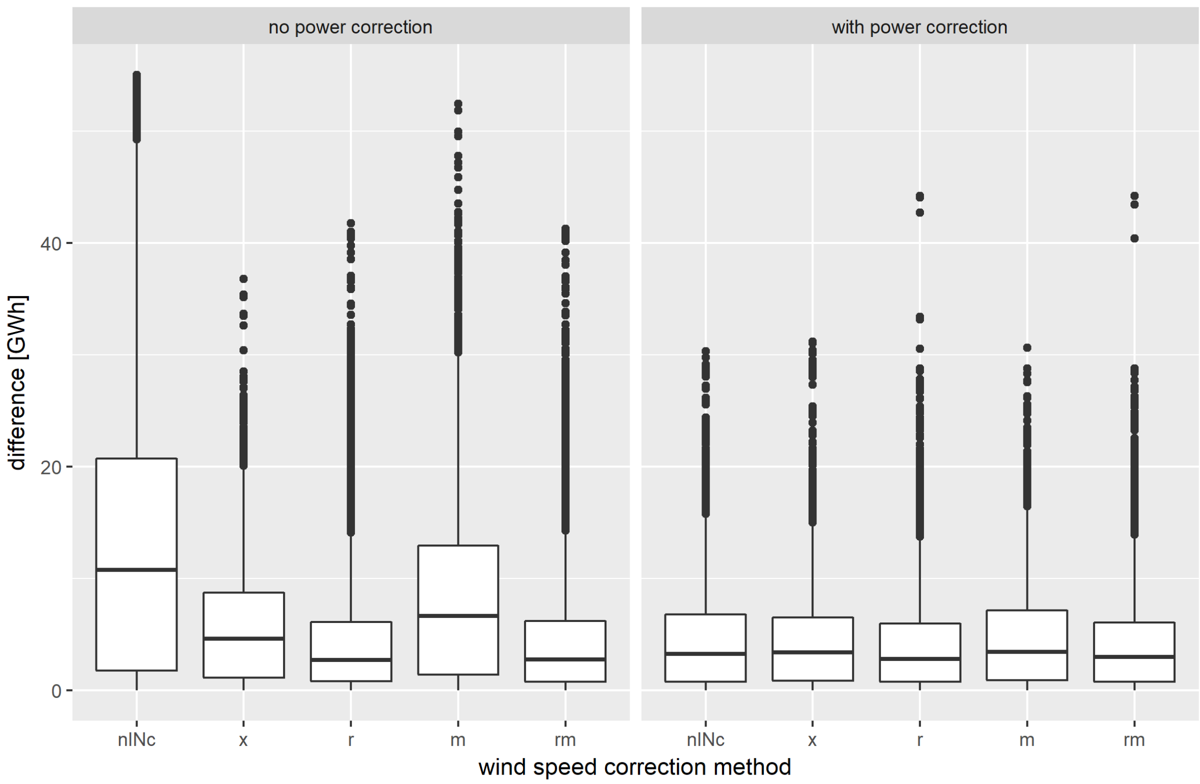

Figure 32: Absolute daily differences in wind power generation per wind speed correction method in Brazil. On the x-axis different wind speed correction methods are shown (own depiction)

2 Links for completing wind park data

This part of the appendix lists links, where information, which was missing in the list of wind parks, was found to be added to the data.

2.1 Locations wind parks

The following links helped to find the locations of wind parks, where the longitude and latitude were not given on The Wind Power websites.

1) Antônio Pimentel de Sousa: Fortaleza

https://www.thewindpower.net/owner_de_1462_antonio-pimentel-de-sousa.php

https://www.openstreetmap.org/search?query=-3.70292%2C-38.47351#map=17/-3.70292/38.47351

2) CGE Delta 3 II, CGE Delta 3 IV, CGE Delta 3 V, CGE Delta 3 VI: Paulino Neves

http://www.omegaenergia.com.br/noticias/primeira-eolica-do-maranhao-inicia-testes/

https://www.openstreetmap.org/search?query=-2.7269%2C-42.5862#map=15/-2.7269/-42.5862

3) Colonia: São Goncalo do Amarante

https://www.thewindpower.net/windfarm_es_20923_colonia.php

https://www.openstreetmap.org/search?query=-3.6078%2C-38.9664#map=14/-3.6078/-38.9664

4) Geraldo Júnior Cavalcante Lopes: Fortaleza

https://www.thewindpower.net/windfarm_en_20811_geraldo-junior-cavalcante-lopes.php

https://www.openstreetmap.org/search?query=-3.70292%2C-38.47351#map=17/-3.70292/-38.47351

5) IMT: Curitiba, Instituto Municipal de Turismo

https://www.thewindpower.net/windfarm_en_4124_imt.php

https://www.google.at/maps/place/Instituto+Municipal+de+Turismo/@-25.427805,-49.274068,19.95z/data=!4m5!3m4!1s0x0:0xc009967fdc27a532!8m2!3d-25.427643!4d-49.2737701

https://www.openstreetmap.org/search?query=-25.42770%2C-49.27363#map=19/-25.42770/-49.27363

6) Instituto Federal de Educação-Ciência-Tecnologia Sul: Pelotas

https://www.thewindpower.net/windfarm_en_8470_instituto-federal-de-educacao-ciencia-tecnologia-sul.php

https://www.google.at/maps/place/Instituto+Federal+de+Educa%C3%A7%C3%A3o+Ci%C3%AAncia+e+Tecnologia+Sul-Rio-Grandense/@-31.7585504,-52.333692,701m/data=!3m1!1e3!4m5!3m4!1s0x9511b5849706d1e5:0x798f23d47e04a1f6!8m2!3d-31.7576761!4d-52.3336171 https://www.openstreetmap.org/search?query=-31.75757%2C-52.33329#map=18/-31.75757/-52.33329

7) Lagoa Seca: Acarau

https://www.thewindpower.net/windfarm_es_18670_lagoa-seca.php

https://www.openstreetmap.org/search?query=-2.8715%2C-40.0901#map=15/-2.8715/-40.0901

8) Malhadinha: Ibiabin(b)a

https://www.thewindpower.net/windfarm_de_22999_malhadinha.php

https://www.openstreetmap.org/search?query=-3.9221%2C-41.0289#map=15/-3.9221/-41.0289

9) MEL 02: Praia de São Cristovão - Areia Branca

https://de.foursquare.com/v/parque-e%C3%B3lico-mel-ii--praia-de-s%C3%A3o-cristov%C3%A3o--areia-branca-rn/53dbab46498e62f1e9732991

https://www.google.at/maps/search/Praia+de+S%C3%A3o+Cristov%C3%A3o+areia+branca/@-4.4471175,-38.8202189,7.46z

https://www.openstreetmap.org/search?query=-4.9447%2C-36.9656#map=15/-4.9447/-36.9656

10) Miassaba III: Guamaré

https://www.thewindpower.net/windfarm_de_8597_miassaba-iii.php

https://www.openstreetmap.org/search?query=-5.1128%2C-36.3874#map=15/-5.1128/-36.3874

11) Pajeu do Vento: Caetité

https://www.thewindpower.net/windfarm_de_20838_pajeu-do-vento.php

https://www.openstreetmap.org/search?query=-14.1200%2C-42.5523#map=15/-14.1200/-42.5523

12) Pedra Cheirosa I – II: Itarema

http://www.pac.gov.br/obra/76907

https://www.openstreetmap.org/search?query=-2.9528%2C-39.8891#map=15/-2.9528/-39.8891

13) Pedro Pedron: Eusébio

https://www.thewindpower.net/windfarm_de_20812_pedro-pedron.php

https://www.openstreetmap.org/search?query=-3.8744%2C-38.3833#map=15/-3.8744/-38.3833

14) Santo Inácio: Icapuí

http://aliancaenergia.com.br/br/projeto-eolico-santo-inacio/

https://www.openstreetmap.org/search?query=-4.7710%2C-37.3082#map=15/-4.7710/-37.3082

15) Stela Maris Zambelli: Eusébio

https://www.thewindpower.net/windfarm_de_20813_stela-maris-zambelli.php

https://www.openstreetmap.org/search?query=-3.8744%2C-38.3833#map=15/-3.8744/-38.3833

16) Tarlene Guedes Bessa: Fortaleza

https://www.thewindpower.net/windfarm_de_20814_tarlene-guedes-bessa.php

https://www.openstreetmap.org/search?query=-3.70292%09-38.47351#map=16/-3.7029/-38.4735

17) Trari: Assumption: Trairi (Error -> fits number 4)

http://sistemabu.udesc.br/pergamumweb/vinculos/00001e/00001eeb.pdf

https://www.thewindpower.net/windfarm_es_20887_trari.php

https://www.thewindpower.net/windfarm_es_20886_trairi.php

18) Vento do Oeste: Acarau

https://www.thewindpower.net/windfarm_de_18672_vento-do-oeste.php

https://www.openstreetmap.org/search?query=-2.8707%2C-40.0902#map=15/-2.8707/-40.0902

19) Ventos de Bahia II: Mulungu do Morro

http://www.windpowerintelligence.com/article/Hv39yUYzNKI/2017/09/29/brazil_27mw_ventos_de_bahia_ii_grid_online/

https://www.openstreetmap.org/search?query=-11.9965%2C-41.5109#map=15/-11.9965/-41.5109

20) Ventos do Araripe III

https://www.openstreetmap.org/search?query=-7.3434%2C-40.5657#map=15/-7.3434/-40.5657

21) Vila Para I, Vila Para II, Vila Para III: Serra do Mel

https://www.thewindpower.net/windfarm_de_22675_vila-para-i.php

https://www.openstreetmap.org/search?query=-5.2115%2C-37.0242#map=15/-5.2115/-37.0242

2.2 Commissioning dates

The following links list where information on commission dates which were missing in the wind park data, were found.

1) Assurua: October 2017

https://renewablesnow.com/news/brazil-clears-128-mw-of-wind-farms-to-operate-as-ipps-in-bahia-473701/

2) Corredor do Senandes I: in construction

http://thehollywood-life.com/companies/odebrecht-energia

3) Eolica Sao Cristovao: 2014

http://www.fiduciario.com.br/uploads/docs/Relat%C3%B3rio_Anual_2013/Trustee/S%C3%83O%20CRIST%C3%93V%C3%83O.pdf

4) Olinda: in August 2015 since 7 years

https://blogdaoposicaodeolinda.wordpress.com/2015/08/31/turbina-eolica-de-olinda-completa-hoje-sete-anos-sem-funcionamento/

5) Primavera: August 2018

https://www.ambienteenergia.com.br/index.php/2017/06/primeiros-geradores-de-energia-eolica-de-sao-paulo-sao-colocados-em-operacao/31953

6) Santo Inácio: Start of operational test in the wind power complrx Santo Inácio on 19th of June 2017

http://aliancaenergia.com.br/br/projeto-eolico-santo-inacio/

7) Taiba: was added later

https://www.thewindpower.net/windfarm_en_3686_taiba.php

8) Ventos de São Benedito: was added later

https://www.thewindpower.net/windfarm_en_23536_ventos-de-sao-benedito.php

9) Ventos do Brejo A-6: project date 03/06/2011

http://cdmloanscheme.org/sites/default/files/pdd.1_1.pdf

3 Scripts

3.1 Download measured wind speeds from INMET

```
setwd("C:/Users/...")
library(RCurl)
library(parallel)
library(XML)
library(tidyverse)

url<- "http://www.inmet.gov.br/projetos/grafico/ema_html_pg.php"

stations<-
read.table("../stations_meta_data.csv",sep=";",header=T,stringsAsFactors=F)

for(station in 1:nrow(stations)){
  if(!file.exists(paste("data/",stations$name[station],".csv",sep=""))){

    s<-seq(ISOdate(1999,1,1), ISOdate(2017,1,1), "hour")
    final<-data.frame(matrix(nrow=length(s),ncol=9))
    names(final)<-c("date.time","temp","umi","po","pres","rad","pre","vdd","vvel")

    final[,1]<-as.numeric(s)

    print(paste("Dealing with",stations$name[station]))

    #station<-470
    #debug(downloadYear)
    #downloadYear(2002,stations=stations,station=station,url=url)

    no_cores <- detectCores() - 1

    cl <- makeCluster(no_cores)
    clusterEvalQ(cl, library("RCurl"))
    clusterEvalQ(cl, library("tidyverse"))
    clusterEvalQ(cl, sink(paste0("c:/temp/output", Sys.getpid(), ".txt")))

    a<-
parSapply(cl,1999:2016,downloadYear,stations=stations,station=station,url=url,simpl
ify=FALSE)

    stopCluster(cl)

    marker<-c("temp","umi","po","pres","rad","pres","vdd","vvel")

    for(j in 1:length(a)){
      df<-a[[j]]
```

```
      if(is.null(df)){
        next
      }

      for(i in 1:8){

        df1<-df[df$types==marker[i],]

        if(nrow(df1)==0){
          next
        }


        cc<-1:nrow(df1)
        cc1<-1:nrow(final)
        d1<-data.frame(df1[,1],cc)
        names(d1)<-c("date","cc")
        d2<-data.frame(final[,1],cc1)
        names(d2)<-c("date","cc1")
        merger<-merge(d1,d2,by=c("date"))
        final[merger[,3],i+1]<-df1[merger[,2],2]
        }
      }

    write.table(final,paste("data/",stations$name[station],".csv",sep=""),sep=";")
  }else{
    print(paste("file exists already:",stations$name[station]))
  }
}

downloadYear<-function(y,stations,url,station){

  print(y)
  df<-
data.frame(date.time=rep(0,8784*8),val=rep(0,8784*8),type=rep("",8784*8),stringsAsF
actors =FALSE)
  names(df)<-c("date.time","val","type")

  result <- postForm(url, mRelEstacao=stations$cod[station],
                     mRelAno=y,
                     btnProcesso=" Gera ")


  if(substr(result[1],1,21)=="Registro Inexistente."){
    print(paste(stations$cod[station],y," not found"))
    return(NULL)
  }

  result<-gsub("; \r","",result)
  result<-gsub("vlr = ","",result)
  result<-gsub(".push","",result)
  result<-gsub("dados_","",result)
  result<-gsub("var dt  =","",result)
  result<-gsub("\r","",result)
  result<-gsub("\\(\\[dt,vlr\\]\\);","",result)
  result<-gsub("var ","",result)
  result<-gsub(";","",result)
  result<-gsub(" ","",result)
  txtvec <- strsplit(result,'\n')[[1]]

  if(txtvec[22]=="<!---alert(temp)-->"){
    return(NULL)
  }

  is<-seq(22,length(txtvec),3)
  dates<-as.numeric(txtvec[is])/1000
  vals<-as.numeric(txtvec[is+1])
  types<-(txtvec[is+2])
```

```
  t<-tibble(dates,vals,types)
  t<-t[!is.na(t[,1]),]
  return(t)
  #t %>% spread(types,vals) %>% return()

}
```

3.2 Download ANEEL wind park data

```
library(RCurl)
library(htmltab)
library(XML)
library(dplyr)
library(tidyverse)
library(lubridate)

setwd("C:/Users/...")

url<-"http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/GeracaoTipoFase.asp"

result <- getForm(url,
                  tipo=7,
                  fase=3)

result<-gsub("\r","",result)
result<-gsub(" ","",result)
#result<-gsub("\r\n","",result)

txtvec<-strsplit(result,'\n')[[1]]

write.table(txtvec[3:length(txtvec)],
            "result.csv",
            quote=FALSE,
            row.names=FALSE,
            col.names=FALSE,
            fileEncoding="UTF-8")

doc<-htmlParse("result.csv",encoding="UTF-8")
tab<-readHTMLTable(doc,which=2,colClasses=rep("character",8))
names(tab)<-unlist(tab[1,])

###use only relevant rows (1st and last row are not useful)
tab<-tab[2:(nrow(tab)-1),]


###convert all to string
tab<-tab %>% as_tibble() %>%  mutate_all(as.character)
tab<-as.data.frame(tab)

###convert power to numeric
tab[,4]<-gsub("\\.","",tab[,4])
tab[,4]<-gsub("\\,",".",tab[,4])
tab[,4]<-as.numeric(tab[,4])

###convert date to date object
tab[tab[,3]=='-',3]<-NA
tab[,3]<-dmy(tab[,3])

###split municipio/state
res<-strsplit(tab[,8]," - ")
for(i in 1:length(res)){
  print(i)
  print(length(res[[i]]))
  if(length(res[[i]])>2){
```

```
    res[[i]]<-c(res[[i]][c(1,3)])
  }

}

tt<-t(data.frame(res[1:((length(res)))]))
tab[,8]<-tt[,1]
tab[,9]<-tt[,2]

names(tab)[9]<-"state"
tab<-as_tibble(tab)

###installed capacity in MW / state
tab %>% group_by(state) %>% summarize(sum=sum(`Potência Outorgada (kW)`)/1000)
```

3.3 Complete wind park data

```
load("C:/Users/.../windparkdata.RData")
load("C:/Users/.../barycentres_municipios.RData")
inst.cap <- read.csv2("C:/Users/.../installed_capacities.csv")

dt1 <-
data.frame(windparks[,1:7],commissioning=as.double(gsub("/","",windparks$commission
ing)))

###### missing coordinates #####
# find data where coordinates are missing
dt2 <- dt1[which(is.na(dt1$long)),]
# try to match data to data of installed capacities with municipios they are
located in from ANEEL
dt3 <- data.frame(dt2,match=match(dt2$name,inst.cap$Usina))
dt3 <- data.frame(dt3,munic=inst.cap$Município[dt3$match])
# find the barycentres of the municipios, if the name can be connected
# fill in according long and lat
dt3$long <- mun_bary$long[match(dt3$munic,mun_bary$municipio)]
dt3$lat <- mun_bary$lat[match(dt3$munic,mun_bary$municipio)]

# in some cases connection is not possible, due to special characters; therefore
manually (data from mun bary):
mun_bary_man <-
data.frame(long=c(mun_bary$long[1868],mun_bary$long[1733],mun_bary$long[1762]),lat=
c(mun_bary$lat[1868],mun_bary$lat[1733],mun_bary$lat[1762]),municipio=c("João
Câmara","Caldeirão Grande do Piauí","Curral Novo do Piauí"))
# match long and lat accordingly (without overwriting existing values)
dt3$long[which(is.na(dt3$long))] <-
mun_bary_man$long[match(dt3$munic[which(is.na(dt3$long))],mun_bary_man$municipio)]
dt3$lat[which(is.na(dt3$lat))] <-
mun_bary_man$lat[match(dt3$munic[which(is.na(dt3$lat))],mun_bary_man$municipio)]

# search still missing data manually and read from csv file
missing_lonlat <-
read.csv2("C:/Users/.../missing_coordinates_windparks.csv",header=F)
names(missing_lonlat)<-
c("x","name","state","cap","lat","long","turbines","on/offshore","date","y")
# fill in missing coordinates
dt3$long[which(is.na(dt3$long))] <- as.numeric(as.vector(missing_lonlat$long))
dt3$lat[which(is.na(dt3$lat))] <- as.numeric(as.vector(missing_lonlat$lat))

# fill in missing coordinates to complete windpark data
dt1$long[which(is.na(dt1$long))] <- dt3$long
dt1$lat[which(is.na(dt1$lat))] <- dt3$lat


# for some reason some of the coordinates have the wrong sign, so this needs to be
changed...
dt1$long[which(dt1$long>0)] <- dt1$long[which(dt1$long>0)]*(-1)
```

```
##### missing comissioning dates #####
# find out where comisssioning date is missing
dt4 <- dt1[which(is.na(dt1$commissioning)),]
# search missing data and read from csv file
missing_dates <- read.csv2("C:/Users/.../missing_dates_windparks.csv",header=T)
missing_dates <- missing_dates[2:11,]
# fill in to complete data frame
dt1$commissioning[which(is.na(dt1$commissioning))] <- missing_dates$commissioning




##### missing capacities #####
# two wind parks are missing capacities
# search for information and fill in:
dt1[which(is.na(dt1$cap)),]
dt1$cap[which(is.na(dt1$cap))] <- c(5000,28000)


##### missing states #####
# 5 wind parks are missing states
# search for information and fill in:
dt1[which(is.na(dt1$state)),]
dt1$state[which(is.na(dt1$state))] <- c("Ceará","Bahia",rep("Piaui",3))

##### faulty commissioning date #####
# Ventos de Santa Edwiges has obviously wrong commissioning date (217/06), correct:
dt1$commissioning[which(dt1$commissioning==21706)] <- 201706


# some points are outside Brazil, remove them:
# (boundary box for download of MERRA data)
lon1<--74.1
lat1<--36
lon2<--33
lat2<-5.5

dt1 <- dt1[which(dt1$long>=lon1 & dt1$long<=lon2),]
dt1 <- dt1[which(dt1$lat>=lat1 & dt1$lat<=lat2),]


day <- NULL
month <- NULL
year <- NULL
for(i in c(1:length(dt1$commissioning))){
  if(dt1$commissioning[i]>10000){
    year[i] <- dt1$commissioning[i] %/% 100
    month[i] <- dt1$commissioning[i] %% 100
    day[i] <- 15
  }else{
    year[i] <- dt1$commissioning[i]
    month[i] <- 7
    day[i] <- 1
  }
}




windparks <- data.frame(dt1,year,month,day)
setwd("C:/Users/...")
save(windparks,file="windparks_complete.RData")
```