Improving wind power predictions on very short-term scales by including wind speed observations in the power forecast

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
(Dr. rer. nat.)

am Leipziger Institut für Meteorologie (LIM)
genehmigt von der Fakultät für Physik und Geowissenschaften
der Universität Leipzig

vorgelegt von

M.Sc. Moritz Lochmann

geboren am 09.05.1995
in Cottbus

Abstract

This work investigates how to improve wind power predictions using observational wind speed data. Measurements from ultrasonic anemometers (sonics) are available from five of the 22 wind energy turbines at the analysed wind farm in Beeskow, Germany (52°11'48"N, 14°13'E). In addition, measurements from a vertically pointing Doppler lidar (DL) at the Meteorological Observatory Lindenberg - Richard Aßmann Observatory located at a distance of 6 km from the wind farm are evaluated. The LoadManager® tool, developed by LEM-Software, Leipzig, is used to perform wind power predictions based on different input data for forecasting horizons of 15 min and 30 min.

Though wind power predictions have consistently improved in the last decade, persistent reasons for remaining uncertainties are sudden large changes in wind speed, so-called ramp events. The occurrence of ramp events at the wind farm has been investigated. Results on the seasonality of ramp events and their diurnal cycle are presented for multiple ramp detection thresholds. Ramps were found to be most frequent in March and April and least frequent in November and December. Furthermore, ramp events occur more frequently during the day than during the night and for most ramp detection thresholds up-ramp events are slightly more frequent than down-ramp events.

For further analysis, the wind power prediction tool is fed with different wind velocity input data. The reference wind power predictions are based on data from a numerical weather prediction (NWP) model. Power predictions using observed wind speed data (DL, sonics) are compared to these reference predictions and evaluated according to two metrics: (i) the absolute difference between the observed and predicted power generation and (ii) the costs incurred due to necessary balancing services. It was found that, (i) the absolute power deviation can be reduced significantly compared to the reference by using power prediction setups based on sonic data. This improvement is even greater during ramp time steps. Power predictions based on the available DL data do not improve the absolute power deviation for the entire data set, albeit they do provide an improvement during down-ramp events. Considering (ii) incurred balancing costs, all power prediction setups based on observational data reduce the balancing costs compared to the reference. Sonic-based configurations yield 75-80 % lower balancing costs than the reference and the DL-based setup results in 20 % lower balancing costs.

Zusammenfassung

Diese Arbeit untersucht, wie sich Windleistungsprognosen mit Hilfe von Windmesswerten verbessern lassen. Messungen von Ultraschallanemometern (sonics) an Gondeln von fünf der 22 Windenergieanlagen des untersuchten Windparks Beeskow, Deutschland (52°11'48"N, 14°13'E), sind verfügbar. Weiterhin sind Messungen des vertikalgerichteten Doppler Lidars (DL) am Meteorologischen Observatorium Lindenberg - Richard Aßmann Observatorium des DWD verfügbar, welches sich in einer Entfernung von 6 km zum Windpark befindet. Das Programm LoadManager® der Leipziger Firma LEM-Software wird für Windleistungsprognosen mit verschiedenen Eingangsdaten für die Prognosezeiträume +15 min und +30 min verwendet.

Die Qualität von Windleistungsprognosen hat sich in den letzten zehn Jahren stetig verbessert. Unsicherheiten bleiben z.B. sogenannte Windrampen, schnelle, starke Änderungen der Windgeschwindigkeit. Das Auftreten von Windrampen am Windpark Beeskow wurde untersucht und die Ergebnisse werden für verschiedene Rampengrenzwerte vorgestellt. Am häufigsten treten Windrampen im März und April auf und am seltensten treten sie im November und Dezember auf. Außerdem wurden Windrampen häufiger tagsüber als nachts festgestellt. Für die meisten Rampengrenzwerte wurden etwas mehr Leistungsanstiege ("up-ramps") als Leistungsrückgänge ("down-ramps") gefunden.

Für weitere Untersuchungen wurden Windleistungsprognosen mit verschiedenen Windgeschwindigkeitsdatensätzen durchgeführt. Als Referenz gelten Windleistungsprognosen auf Basis von Daten numerischer Wettervorhersagemodelle. Windleistungsprognosen auf Basis von Messwerten (sonics, DL) werden mit dem Referenzmodell verglichen und entsprechend zweier Metriken bewertet: (i) der absoluten Abweichung zwischen der vorhergesagten und beobachteten Stromerzeugung und (ii) der für Abweichungen anfallenden Regelenergiekosten. Die Ergebnisse zeigen, dass (i) die absolute Abweichung verglichen mit der Referenz signifikant reduziert werden kann, in dem man Messwerte von sonics für die Leistungsprognose verwendet. Dabei ist die Verbesserung während Windrampen größer als für den gesamten Datensatz. Windleistungsprognosen auf Basis von DL-Daten zeigen keine Verbesserung der Abweichungen für den gesamten Datensatz, jedoch eine signifikante Verbesserung während Leistungsrückgängen. Betrachtet man (ii) die anfallenden Regelenergiekosten, resultieren alle auf Messwerten basierenden Leistungsprognosen in einer Reduktion der Kosten verglichen mit dem Referenzmodell. Windleistungsprognosen auf Basis der Gondelmessungen reduzieren die Regelenergiekosten um 75-80 % und Windleistungsprognosen auf DL-Basis ergeben im Mittel etwa 20 % niedrigere Regelenergiekosten.

Contents

1 Introduction

The world is in the midst of an ever intensifying, anthropogenic climate emergency. The changing climate impacts the biodiversity of ecosystems, global supply chains and the human society itself (*IPCC*, 2022). Scientists agree that the rising temperature can largely be attributed to the still rising human made carbon emissions and resulting feedback mechanisms. The Paris Agreement, negotiated at the 2015 United Nations Climate Change Conference COP 21, set the goal to limit the global temperature increase to at most 2 K above pre-industrial levels, but even that goal will be difficult to achieve by most country's near-term mitigation targets (*Schleussner et al.*, 2016).

A considerable contribution to the rise of global greenhouse gas concentrations is the world's large dependency on fossil energies with high carbon dioxide emissions. The largest contribution (25 %) of greenhouse gas emissions by economic sector is caused by electricity and heat production (*IPCC*, 2014). Since the end of the 20th century, renewable energy sources, especially wind and solar energy, have been drawing increasing attention for the purpose of electricity generation. In 2021 in Germany, renewable energy sources contributed 41.1 % (233 TWh of 567 TWh) to the gross electricity consumption (*Bundeswirtschaftsministerium*, 2022). Approximately half of that is generated from on- and offshore wind farms. However, in order to meet carbon emission reduction targets, the contribution of conventional fossil electricity generation will be reduced further in the course of the next years in Germany. In April 2023, the three remaining German nuclear power plants will phase out. In the 2030s, German lignite and hard coal power plants will be shut down. Therefore, the share of renewable energy of the electricity consumption has to increase as the total energy consumption of Germany is expected to reach around 600 TWh in 2030 (*Agora-Energiewende*, 2018).

In addition, another advantage of renewable energies gained attention recently. Even though the production of photovoltaic modules and wind energy turbines still depends on resources from other countries and intact global supply chains, the generation of electricity by renewable power plants is independent of other countries, as it happens locally and without additional resources. Thus, in the majority of the public and the political parties there is a consensus that the expansion of renewable energies in the next decades is necessary.

1.1 Motivation

During the last decades, the electricity generation by renewables experienced considerable technical advances. Additionally, renewably generated electricity is already cheaper than most other sources (*Kost et al.*, 2021). However, renewable power plants generate electricity distributed rather than centralised, which adds novel demands to the electrical grid infrastructure. In addition, to ensure the safe and continuous operation of the electric grid, the power consumption and power generation have to be actively balanced. Since the power generation of renewable energies depends on the weather, their deployment cannot be scheduled precisely. One major challenge of the integration of renewably generated electricity into the power grid is to accurately predict their power generation. The more accurate the wind power predictions are, the less efforts to balance the grid are needed. Balancing services include, e.g., regulating the operation of already running power plants, deploying storage facilities,

like pumped storage hydro power stations, large-scale batteries or gas turbines, and, if necessary, forced power plant shutdowns. The deployment of extra power generating or consuming facilities and regulations of power plants on short notice are usually more expensive than regular electricity generation (*SMARD*, 2021a). Consequently, most deviations from the power prediction cause additional costs. Improving wind power predictions, thus, reduces necessary balancing efforts and costs and helps to keep the electric grid stable.

Naturally, there are additional challenges for the transition to a more sustainable society. There is still opposition among citizens to planned wind and photovoltaics projects, application procedures are laborious and disputes over areas designated for renewable energy generation are lengthy. All of these reasons extend the planning phase of renewable energy installation projects to multiple years. However, the contribution of this thesis to the so-called "green transition" is to evaluate potential improvements to the efficiency of already existing wind farms, which is discussed subsequently.

1.2 Improving wind power predictions

Since the power generation through wind is mostly dependent on the highly variable wind speed, accurate wind power predictions require accurate wind speed predictions on multiple scales. Large-scale synoptic conditions, e.g., frontal zones, affect the wind power generation on a national or international scale. Small-scale conditions, e.g., local orography or turbulence patterns, affect the wind power generation of specific wind farms or even single turbines. Numerical Weather Prediction (NWP) models are a suitable source for wind speed predictions. These models perform well in predicting the general development of the wind speed up to a few days in advance and are even able to predict local phenomena like low-level jets to some extent (*Tay et al.*, 2021). In addition, NWP model data are available for the entire model area without much preparation or instrument setups. However, sudden changes in wind speed cannot always be predicted in detail by NWP models, e.g., when they occur between or over the course of only a few model time steps or when they are on sub-grid scales.

To improve the representation of short-term changes in power generated from wind, this study uses observational wind speed data measured by ultra sonic anemometers at the back of the wind energy turbines and by a vertically pointing Doppler Lidar (DL), operated in proximity to the wind farm. In contrast to this thesis, horizontally pointing DL (*Theuer et al.*, 2020; *Pichault et al.*, 2021b) or horizontally pointing Doppler radar (*Valldecabres et al.*, 2020) observations have already been used to predict wind power generation for wind farms offshore and onshore in moderately complex terrain with significant improvements compared to the benchmarks. The use of a horizontally pointing DL in complex terrain was also attempted but was not yet successful (*Würth et al.*, 2019). When pointed into upstream direction, a horizontally pointing DL has the advantage of observing a cone-shaped layer at preferably hub height and, therefore, evaluating the whole horizontal wind field before it arrives at the wind farm. However, it usually does not measure the vertical wind profile. A third, rather novel way to utilise lidar measurements for wind energy purposes are nacelle mounted Doppler lidars. These instruments need to be located at the top of the nacelle, similar to the ultra sonic anemometers and measure the wind field in a horizontal plane. For power prediction purposes nacelle DL are observing the wind field in front of the wind energy turbine (e.g., *Dolatabadi et al.*, 2022), while for studies of wind turbine wake effects they would observe the wind field behind the turbine (e.g., *Shin and Ko*, 2022). Since their measurement range is only a few hundred meters, power predictions based on nacelle DL provide forecasts with forecasting horizons of few seconds to a few minutes. Still, for these forecasting horizons power

predictions based on nacelle DL data look promising and show improvements compared to the reference models (*Dolatabadi et al.*, 2022). For this thesis, only vertically pointing DL measurements of the vertical wind profile are available. Thus, it will be used to investigate whether the profile data can achieve similar improvements of the wind power prediction as horizontally pointing DL do.

In addition to DL data, observations by ultra sonic anemometers are used to improve the representation of short-term changes in wind speed. Previously, *Wilczak et al.* (2019) achieved improvements of power forecasts by including in-situ observations, especially during the first six hours of the forecasting horizon. Since the sonic measurements are available at the exact location of the wind farm, they represent a very valuable data set.

In this thesis, wind power predictions are investigated for the wind farm Beeskow, Germany (52°11'48"N, 14°13'E), which consists of 22 wind energy turbines and has a total capacity of 44 MW. This wind farm was chosen, because it is situated in close proximity to the Meteorological Observatory Lindenberg - Richard-Aßmann Observatory (MOLRAO). At MOLRAO, which is located at a distance of approximately 6 km to the wind farm, the German Meteorological Service (Deutscher Wetterdienst - DWD) (DWD) conducts long-term observations of the atmosphere. This multi-year data set includes observations of the vertical wind profile by a DL and a wind-profiling radar. Nacelle anemometer data are available for five of the 22 wind energy turbine locations for a five year period. Based on the given information, the first research topic of this thesis is:

I. **Are the Doppler lidar observations representative for the location of the wind farm? How well do the NWP model data agree to the observations?**

Before these data sets are used for wind power prediction, they are checked for biases. For that purpose, the in-situ measurements at the wind farm are compared to the DL observations at MOLRAO. Additionally, a wind velocity data set of a meteorological tower in proximity to both the wind farm and MOLRAO is used for validation. Furthermore, the agreement of the NWP model data with the observational data set is of interest, as a potential bias of the NWP model will affect the power predictions.

Subsequently, the meteorological data need to be linked to the wind power generation. In this thesis, the power prediction component of the LoadManager® software will be used. The LoadManager® software was developed by the Ingenieurbüro Last- und Energiemanagement (LEM-Software), Leipzig, Germany, as a tool to forecast time series of energy generation, energy consumption and costs (*Zimmer et al.*, 2012). The LoadManager® yields a power prediction schedule for the investigated wind farm based on the input data. This leads to the second research question of this work:

II. **Which wind situations cause challenges for wind power predictions?**

Small deviations from the power prediction schedule are almost unavoidable. However, unexpected large changes in wind power generation, so called wind power ramp events, pose a challenge to the efficient and reliable operation of the power grid (*Potter et al.*, 2009). Ramp events are typically defined over temporal scales of a few tens of minutes to a few hours resulting in changes of the generated power between 10 % and 80 %. In this thesis, ramp events ranging from 30 min to 2 h with power changes of 30 % to 70 % are investigated. This work will also study the diurnal cycle and the seasonality of ramp events for the investigated wind farm Beeskow, of which results are currently being published in *Lochmann et al.* (2022).

The third research topic considers the evaluation of wind power predictions with different wind speed input data:

III. **Can wind power predictions be improved with the available observational data sets? Which instrument provides the best power predictions?**

Wind power predictions in this thesis are based on either NWP model wind speed data, nacelle sonic observational data or DL measurements. The goal of this thesis is to determine whether observational data can improve the quality of wind power predictions. Therefore, the agreement between the power prediction and the observed power generation is analysed. In one approach, the power prediction with the least absolute deviation to the observed power generation is sought. In a second approach, power prediction errors are converted to costs incurred for balancing services and the least costly power prediction configuration is determined.

The evaluation of balancing costs motivates the fourth research topic:

IV. **Does the benefit of using observational data outweigh the acquisition costs of the instruments? How long is the amortisation time?**

A cost-benefit estimation for the acquisition of instruments is pursued, especially for the power predictions based on DL data. Anemometer measurements are already mandatory on most wind energy turbines in Germany. Thus, any potential benefit of the sonic observation would come at almost no cost. Although DL observations provide more information on the wind profile than the in-situ measurement at the nacelles, DL instruments are considerably more expensive. Whether the acquisition of an instrument pays off depends on the reduction in balancing costs achieved through a better power prediction, if any.

The efficient use of information in the time-sensitive renewable energy business is important, especially for very short-term forecasting horizons. However, the awareness of the challenges is high and the use of weather related data is promoted (*Steiner et al.*, 2017). To be conveniently implementable, any changes to the power prediction process need to be straightforward.

The outline of this thesis is as follows. Chapter 2 provides the theoretical background on wind in the atmospheric boundary layer and describes relevant parts of the German electricity system. Chapter 3 introduces the instruments, the NWP model and the LoadManager®️ power forecasting tool used in this thesis. In Chapter 4, the study site is introduced and the various wind speed data sets are analysed. In addition, Chapter 4 describes the method of short-term wind speed prediction and the detection method for ramp events. Furthermore, Chapter 5 presents results on the occurrence of ramp events and evaluates the performance of the various wind power prediction configurations. Finally, Chapter 6 contains the summary, conclusions and the outlook of this thesis.

2 Theoretical background

In this Chapter, a quick overview of the atmospheric boundary layer is given in Sect. 2.1. Subsequently, Sect. 2.2 describes the wind field in the atmospheric boundary layer in more detail. In the end of this Chapter, Sect. 2.3 explores the structure of the German power grid, discusses reasons for the necessity of balancing energy and concludes with a brief summary of the calculation of balancing costs.

2.1 The atmospheric boundary layer

The Atmospheric Boundary Layer (ABL) is the lowest part of the troposphere where interactions between the surface and surrounding air take place. This layer reacts to forcings from the surface, e.g., the emission of thermal radiation or friction due to surface roughness, on time scales of a few hours or less. The unique nature of the ABL affects most meteorological variables, especially temperature, turbulence and wind. *Stull* (2017) describes the properties of the atmospheric boundary layer under different conditions in great detail. Subsequently, these properties are briefly summarised.

Figure 2.1a shows the typical ABL structure at daytime and during the night (Fig. 2.1b) with typical profiles for temperature T, potential temperature Θ, water vapour mixing ratio r and mean horizontal wind speed v. The characteristics of the wind speed profile will be discussed in more detail in Sect. 2.2. The daytime ABL is structured into three parts. At the bottom is the surface layer (SL, German: "Prandtlschicht"), which is most dominantly affected by the ground underneath. Above the surface layer is the mixed layer (ML), which most of the time constitutes the largest part of the ABL. Between the mixed layer and the free atmosphere (FA) is the entrainment zone (EZ).

The diurnal boundary layer evolution represents an example for a convective atmospheric boundary layer. During the day, solar radiation is partly absorbed by the surface, heating it up. Subsequently, thermal emission of the surface heats up the lowest layers of air in the surface layer. Heat and water vapour are transported through the surface layer and the temperature as well as the water vapour mixing ratio decrease with height. However, as the turbulent fluxes are approximately constant with height, the surface layer is also called constant flux layer. Above the surface layer, turbulence causes a constant distribution of potential temperature throughout the mixed layer. Here, the mixing ratio further decreases with height. The mixed layer is topped by a temperature inversion in the entrainment zone, where the mixing ratio decreases significantly.

The nocturnal boundary layer is an example for a stable boundary layer (SBL), where a cold surface is topped by relatively warmer air. When the Sun sets, radiative surface warming ceases and the lowest layers of air cool down, causing a stable surface layer. Through the course of the night this stable nocturnal boundary layer deepens in height, but does not reach the capping inversion (CI) aloft. On top of the nocturnal boundary layer remains the residual layer with similar properties for temperature and mixing ratio as the diurnal mixed layer.
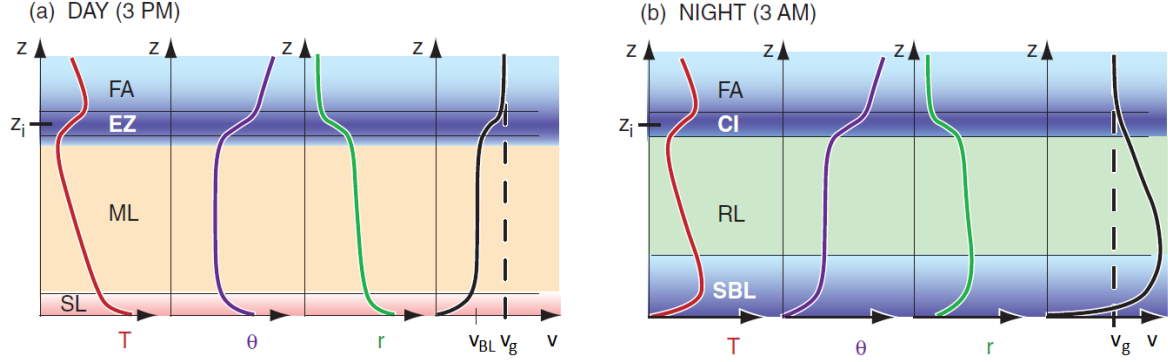
Figure 2.1: Typical vertical profiles of temperature T, potential temperature Θ, water vapour mixing ratio r, and wind speed v in the atmospheric boundary layer. The dashed line labeled v_g is the geostrophic wind speed. v_BL refers to the average wind speed in the atmospheric boundary layer. The shading describes the atmospheric stability in each layer, orange represents an unstable stratification, green is neutral and blue refers to a stable stratification (adapted from *Stull*, 2017).

2.2 Wind above and in the atmospheric boundary layer

Wind velocity is the main resource for the generation of wind energy. In the free troposphere above the atmospheric boundary layer, the horizontal wind speed is determined by the balance of forces between Coriolis force $\overrightarrow{F}_\mathrm{C,H}$ and horizontal pressure gradient force $\overrightarrow{F}_\mathrm{P,H}$:

$$\overrightarrow{F}_\mathrm{C,H} = -f\,\overrightarrow{k} \times \overrightarrow{v}_\mathrm{H}, \tag{2.1}$$

$$\overrightarrow{F}_\mathrm{P,H} = -\frac{1}{\rho}\nabla_\mathrm{H}p, \tag{2.2}$$

where f is the Coriolis parameter, $\overrightarrow{k}$ the vertical unit vector, $\overrightarrow{v}_\mathrm{H}$ the horizontal wind velocity, ρ the density of air and p the atmospheric pressure.

This approximation is based on scale analysis, which identifies $\overrightarrow{F}_\mathrm{C,H}$ and $\overrightarrow{F}_\mathrm{P,H}$ as dominant terms for the horizontal equations of motion. The equation of motion with only these dominant terms is reduced to:

$$-f\,\overrightarrow{k} \times \overrightarrow{v}_H = -\frac{1}{\rho}\nabla_H p. \tag{2.3}$$

Since the temporal derivative is missing, this approximation does neither contain advection nor local changes to the horizontal wind and, therefore, describes an equilibrium state (*Kraus*, 2004). This approximation of the horizontal wind is called geostrophic approximation and the resulting horizontal wind $\overrightarrow{v}_H$ is called the geostrophic wind $\overrightarrow{v}_g$:

$$\overrightarrow{v}_g = \frac{1}{\rho f}\,\overrightarrow{k} \times \nabla_H p. \tag{2.4}$$

Above the atmospheric boundary layer, the geostrophic wind can be considered as the actual wind and this approximation agrees very well to observations (*Kraus*, 2004). However, through all three sub-layers of the ABL the wind velocity is usually subgeostrophic because of frictional and turbulent drag. In the ABL, $\overrightarrow{v}_\mathrm{g}$ is not the actual wind velocity, but it can be considered as a measure of the pressure gradient force to simplify its effect in boundary layer equations.

During neutral conditions, the wind speed in the surface layer follows the logarithmic wind profile. At a height z_0 above ground, the wind speed is assumed to be 0. This height z_0 is called the roughness length and depends on the surface type. Upwards of z_0 the wind speed increases logarithmically within the surface layer, which typically extends to up to 100 m. The wind direction is approximately constant in the surface layer. The shape of the logarithmic wind profile depends on z_0:

$$v_\mathrm{H}(z) = \frac{u*}{k} ln\left(\frac{z}{z_0}\right) \text{ for } z \geq z_0.$$ (2.5)

where v_H is the horizontal wind speed, $u*$ the friction velocity, k=0.4 the von Karman constant and z the height above ground. For wind profiles above, e.g., forested areas, the logarithmic wind profile begins at the top of the canopy and, therefore, is shifted by the displacement height d upwards:

$$v_\mathrm{H}(z) = \frac{u*}{k} ln\left(\frac{z-d}{z_0-d}\right) \text{ for } z \geq z_0.$$ (2.6)

If the surface layer deviates from a neutral stratification, the logarithmic wind profile is not valid anymore. This is due to a simplification of turbulence in the derivation. For non-neutral conditions, the Monin–Obukhov similarity theory has to be applied (*Monin and Obukhov*, 1954). The logarithmic wind profile is adjusted with the Obukhov Length L, which describes the height at which more turbulence is created due to buoyancy than due to wind shear (*Stull*, 2017). The surface layer is stable for L>0, unstable for L<0 and $L \rightarrow$ inf for neutral stratification. More details on the Monin–Obukhov similarity theory can be found in *Kraus* (2008).

2.2.1 Wind in the stable boundary layer

For a stable stratification, the boundary layer is usually only a few hundred meters deep. The wind speed is lower near the ground and higher aloft than given by the logarithmic wind profile, causing a higher vertical wind speed gradient. A stable boundary layer is especially prone to the development of a low-level jet. Low-level jets are characterised by an elevated maximum of the wind speed within the atmospheric boundary layer, typically between heights of 100 m to 500 m (*Baas et al.*, 2009). They often occur during night, when radiative cooling causes a stable layer right above the ground. The air above becomes decoupled from the ground friction, which causes an imbalance in the forces resulting in a region with supergeostrophic wind.

2.2.2 Wind in the convective boundary layer

For situations with a high amount of thermal convection, the wind speed increases throughout the surface layer and then becomes constant at a short distance above ground (see Fig. 2.1a). The wind speed stays approximately constant throughout the mixed layer, which can extend from a few hundred meters to up to more than 2 km (*Kraus*, 2008). The extent of the mixed layer depends largely on the amount of turbulence generated and, therefore, predominately on the time of the day, the season and the synoptic boundary conditions. In the mixed layer the wind direction changes according to the Ekman spiral as the influence of surface roughness decreases with height. Figure 2.2 illustrates the deviation from the geostrophic wind $\overrightarrow{v}_\mathrm{g}$ for a height z. Above the atmospheric boundary layer, the horizontal wind speed is approximated by the balance of forces between the Coriolis force $\overrightarrow{F}_\mathrm{C,H}$ and the horizontal
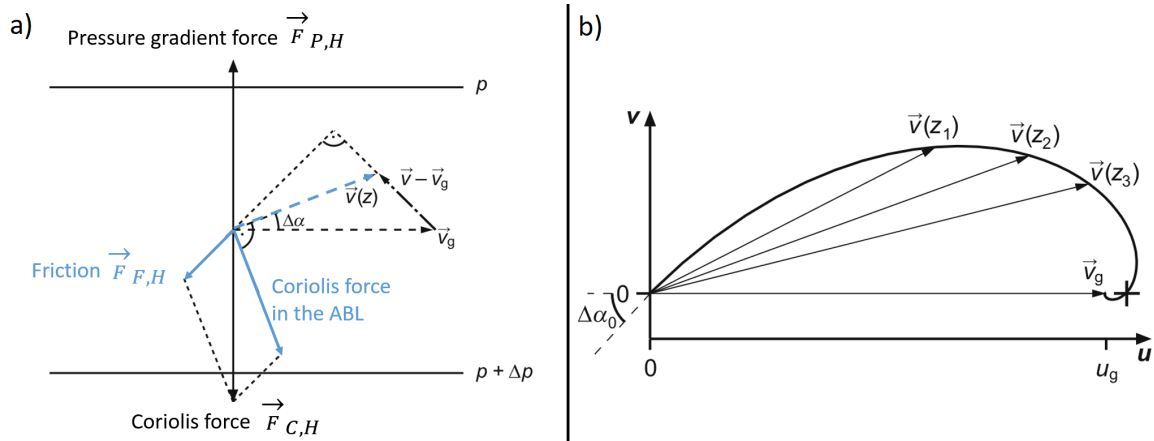
Figure 2.2: a) Balance of forces in the atmospheric boundary layer. The black arrows show the geostrophic approximation, the blue arrows represent the deviation from the geostrophic approximation in the atmospheric boundary layer caused by the additional friction. b) Change of the wind vector with height, for $z_1 < z_2 < z_3$, in the atmospheric boundary layer. The spiraling shape is eponymous for the so-called Ekman spiral (adapted from *Kraus*, 2008).

pressure gradient force $\vec{F}_{\mathrm{P,H}}$ (solid black lines in Fig. 2.2a). In the boundary layer, the friction $\vec{F}_{\mathrm{F,H}}$ causes a deceleration of the wind (blue in Fig. 2.2a). The difference between $\vec{v}(z)$ - $\vec{v}_{\mathrm{g}} = \vec{v}_{\mathrm{ag}}$ is called ageostrophic wind. The friction $\vec{F}_{\mathrm{F,H}}$ is directed perpendicularly to $\vec{v}_{\mathrm{ag}}$ and, therefore, not directly opposite to the wind speed, but slightly inclined. For this lower wind speed the absolute value of $\vec{F}_{\mathrm{C,H}}$ decreases (blue) and is no longer equal to $\vec{F}_{\mathrm{P,H}}$. Therefore, the horizontal wind vector $\vec{v}(z)$ (blue) is no longer parallel to the isobars, but angled into the low pressure and has a lower value. The effect of the frictional force is largest near the ground. With increasing height the influence of the frictional force on the balance of forces decreases and $\vec{v}(z)$ converges to the geostrophic wind vector $\vec{v}_{\mathrm{g}}$, as shown in Fig. 2.2b. In the free troposphere above the entrainment zone the wind is, according to the geostrophic approximation, fully geostrophic.

2.3 The German power grid

In this section, the integration of the generated wind power into the electrical grid will be discussed. To enable the implementation of results of this thesis on real-world applications, they need to fit to the real-world electricity system framework. Therefore, the transition from the observed or modeled wind data, via power generation by wind energy turbines to electrical power within a stable national grid has to be briefly explained. Subsequently, the structure of the grid will be introduced in Sect. 2.3.1. Since the wind power generation is variable and can deviate from the power prediction, a balancing system is needed to ensure the stability of the electrical grid. The balancing system is explained in Sect. 2.3.2. Section 2.3.3 summarises how these balancing costs are calculated.

However, this introduction into the German electricity system does not conclusively explain all its procedures and intricacies, as this is not the main topic of this thesis. This section is rather meant to provide background information to integrate the *cost*-oriented analyses (Sect. 5.3) into the larger context.
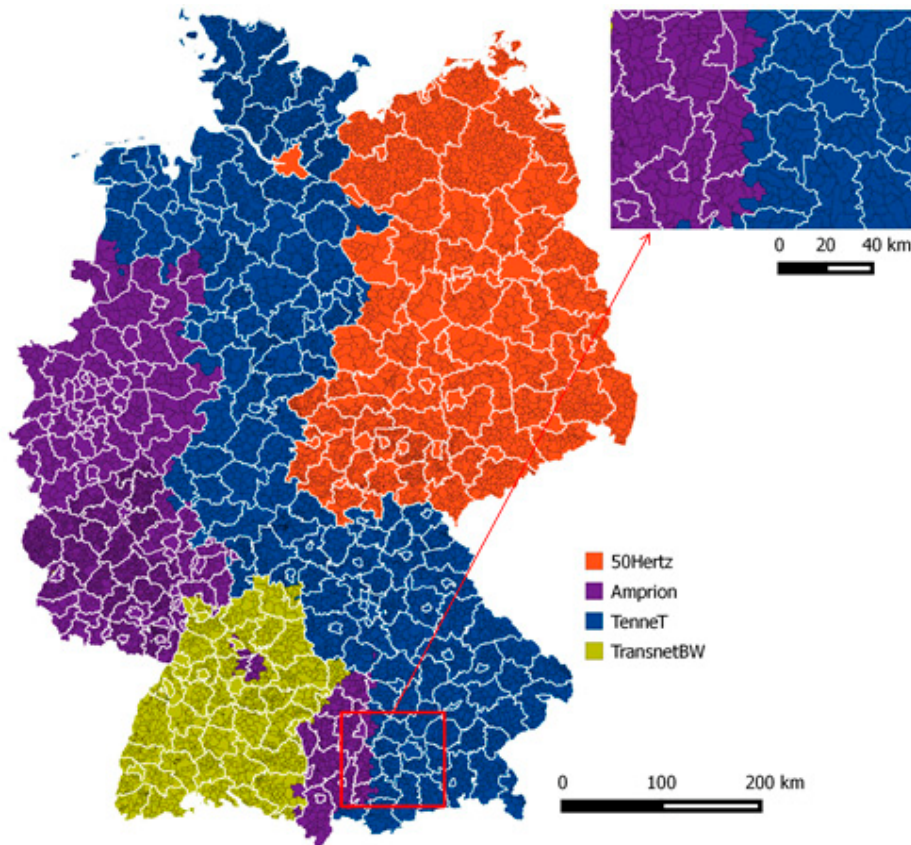
Figure 2.3: Transmission system operators of Germany. The cutout highlights that TSO area border do not necessarily coincide with state or district borders (*Robinius et al.*, 2017).

2.3.1 Structure of the German power grid

While the German electricity grid consists of many individual units, it is thoroughly connected over the whole country. The smallest unit is the balancing group ("Bilanzkreis"). Balancing groups are virtual energy quantity accounts managed by a balance responsible party and may include, e.g., the power stations of a power plant operator or the entire generation and demand of an energy utility (*SMARD*, 2021a). However, balancing groups are not coupled to a physical location but to its operator. For example, one person may be in a different balancing group than their neighbour, because they have different electricity providers. Similarly, one could be in the same balancing group as a power plant on the other side of Germany because one electricity provider is running both. Every balancing group is part of one of the four German Transmission System Operator (TSO) areas. Figure 2.3 shows the four German TSOs. Notably, the TSO area borders do not necessarily coincide with state or district borders. Balancing groups are connected within the TSO areas and the TSO areas are connected within the national power grid. Additionally, the German electrical grid is connected to its neighbouring countries. The goal of these connections is to minimise balancing efforts, explained in detail in the following section.

2.3.2 The power balancing system

The German power grid operates at a frequency of 50 Hz, which has to be maintained at all times to remain operational. To keep this frequency stable, the amount of electric energy fed

into the grid needs to equal the amount of electric energy taken out of the grid by consumers. If this balance is not achieved, the grid frequency deviates from 50 Hz. If there is too much power in the system, the grid frequency increases; if there is not enough power, the grid frequency decreases.

To avoid major deviations from the grid frequency, both the electricity generation and consumption are estimated in advance. All power plant operators have to register their planned daily power production estimates on the day before, called power schedule (*SMARD*, 2021a). This is especially important for renewable energy power plants, where the power generation is variable. With this power schedule and an estimated power consumption profile, balancing group operators create a net schedule for the following day. Subsequently, the TSOs plan the operation of power plants to provide exactly as much power as needed, based of the received power schedules. However, even with the most advanced power forecast and power load models, deviations from the schedule are unavoidable, causing a net deficit or excess of power in the electric grid. Since the electric grid cannot store electricity in a meaningful capacity, differences between power generation and power consumption have to be balanced carefully.

One method is given by the regulation of the power generation of the different power plants. In the case of power generation through wind energy, power generation can be decreased by turning the rotor blades of a number of wind energy turbines out of the wind. This process, called curtailment, reduces the power generation effectively and on short time scales. The opposite, i.e., the increase of the power generation of wind energy plants, only works if some turbines have previously been shut down and can be turned back on again. This procedure is called "Einspeisemanagement" in the German electricity system. It is usually applied for renewable energy generated from wind or photovoltaics power plants, since they can be turned on and off on a rather short notice. For other power plants it can be more difficult to regulate their power generation swiftly. Operators of large fossil or nuclear power plants face significant costs by turning their facilities down and back up again later. In most cases, throttling the generation to a technical minimum and accepting even negative electricity prices for a little while is cheaper than stopping the power generation completely. Another reason for continuous power generation through negative market prices is called *conventional generation base* (German: "konventioneller Erzeugungssockel"). Some conventional power plants are not only used for electric energy generation but are combined heat production and power plants, which provide heat for a large area or high number of consumers. In that case the continuous operation of the facility is necessary in order to provide both services. Furthermore, (conventional) power plants attached to industrial facilities are allowed to generate power for their own requirements, independent of the grid situation (*SMARD*, 2021a). Therefore, even with actions taken to reduce differences between power generation and consumption within the electrical grid, there will always be an imbalance of electricity in the grid, which needs to be compensated.

The connections between the different units within the German power grid and beyond serve the purpose to reduce balancing measures needed. If, for example, one balancing group is slightly oversupplied and another slightly undersupplied, both would initiate balancing efforts in different directions (e.g., turning off wind energy plants to reduce power generation in the oversupplied balancing group or activate a pumped storage hydro power station to increase generation in the undersupplied balancing group). However, with the interconnections they can balance their excess and deficit against each other. The value of the resulting imbalance is lower than the sum of imbalances before, even if they do not cancel each other out completely.

There are a few thousand balancing groups in Germany (28.10.2022: 3302 balancing groups in TenneT, Germany's largest TSO; *TenneT*, 2022), most of them with either an oversupply or an undersupply. To increase synergies between the German TSO areas, the German Grid Control Cooperation (GCC) was formed, consisting of four modules.

Module 1, called *Prevent counteracting control reserve activation*, formulates the aim to avoid activating opposite balancing services. A controlled and targeted energy exchange within Germany and over its borders reduces balancing costs simply by not activating control reserves with different polarities.

Module 2, namely *Common dimensioning of control reserve*, means that the four TSO areas join their control reserves across balancing groups in order to decrease their individual power reserves and, therefore, also decrease maintenance costs.

Module 3, the *Common procurement of secondary control reserve*, opens the balancing power market across the TSO areas. This way providers of balancing energy need to compete directly with providers from other areas, basically leveling prices for balancing energy across the four TSO areas.

Module 4, the *Cost-optimised activation of control reserve*, explains that balancing power is used based on merit order lists. Power generation from different sources is associated with different costs. Starting with the cheapest source, more energy sources are added until the current consumption is reached. The order of the power plants is determined by a merit order list.

Any imbalances remaining in the electrical grid need to be corrected as quickly as possible to assure a stable grid frequency. This correction is the responsibility of the TSOs. For that matter they distinguish between different response times, depending on how much power has to be balanced and how long the deviation persists, these reserves are triggered (*SMARD*, 2021a). The reserves are:

- frequency containment reserves (FCR, German "Primärregelreserve"), which must be fully available within 30 seconds,

- frequency restoration reserves with automatic activation (aFRR, German "Sekundärregelreserve") within five minutes and

- frequency restoration reserves with manual activation (mFRR, German "Tertiärregelreserve") within 15 minutes.

2.3.3 Balancing cost calculation

Activating any of the three balancing reserves costs money. While the TSOs are responsible for activating balancing reserves, the balancing group operators need to pay the balancing costs if their balancing group is deviating from the power schedule.

Module 3 of the GCC, establishing a German-wide market for balancing energy, constitutes a uniform balancing energy price for all balancing groups in Germany. The German *Bundesnetzagentur* is responsible for monitoring and managing the German power grid. It provides a detailed data set of balancing energy prices on their database *SMARD* (2021b). Included in this data set is the current status of the German electricity grid and the corresponding balancing price. Since the balancing price is determined by supply and demand, it can also be negative. If, for example, periods with high supply coincide with periods with low demand, the grid experiences an excess of electricity, which needs to be balanced. As explained in Sect. 2.3.2, some power plants will continue to operate during these times anyway, reinforcing the imbalance. At these times, power plant operators will get negative prices for

their excess power. That means they have to pay for the share of their generation, which surpasses their power schedule. On the other side of the situation are, e.g., storage operators, who are willing to take some power out of the grid and pay negative prices, meaning they are getting paid for relieving the pressure on the grid.

In Germany, operators of wind farms usually hire direct marketers (German: "Direktvermarkter"), who take over the responsibility for power forecasts for multiple power plants combined. In the business relation between wind farm operator and their direct marketer, usually the direct marketer is accountable for balancing costs. Therefore, it is mainly in the direct marketers interests to minimise deviations between power forecast and power generation to minimise their costs for balancing efforts. The larger and more spread the direct marketers portfolio, the lower their risk. Smaller direct marketer portfolios with fewer plants in only one region of Germany are more prone to higher risks - unfortunate periods on the "wrong" side of the nations grid situation can lead to balancing costs erasing weeks worth of revenue. Of course, there are equal chances to make larger profits due to being on the "right" side of the grid imbalance. Larger direct marketer portfolios with many power plants all over the country usually profit less, because more often than not they represent the general grid situation better and are thus generally on the "paying" side of the grid. This is also an advantage though, because outliers (which would cause very high balancing costs in smaller portfolios) are smoothed out over the large number of individual sources. Therefore, larger portfolios are less prone to higher risks.

As a simplified example, picture strong wind and, therefore, a large amount of wind power production in the North Sea and in Northern Germany. If that happens on a weekend, when there generally is lower power consumption, the German grid would probably be oversupplied. Despite precise wind power forecasts and the knowledge of a high production, the inevitability of low consumption and the inability to completely turn off the conventional generation base leads to an oversupply. However, synoptic conditions in a different part of Germany may be decoupled from the situation over the North Sea. A small balancing group with only one wind farm (or a small direct marketer portfolio) may generate less than originally predicted, due to various reasons (e.g., random errors in the forecast, delayed increases in wind speed in regards to the model or premature decreases in wind speed). This individual balancing group is undersupplied. That means its balancing demand goes in the opposite direction of the nationwide grid. While the majority of balancing group operators in Germany are looking to sell generated power exceeding their power schedule, the operator of this specific balancing group is looking to acquire additional power to equalise it. Naturally, the high supply and low demand in this situation causes energy prices to be very low or even negative, making it cheap for the operator of the specific balancing group to balance it. On another day in a similar situation, the singular balancing group may be oversupplied like the rest of Germany, due to the same random errors with contrary direction. Then the balancing group operator would have to pay to dispose of excess power like the majority of balancing group operators.

In conclusion, random errors may cause large swings in the balancing costs in both directions for single balancing groups. The larger the balancing group gets or the more balancing groups are united under the same direct marketer, the more these random effects average out. Very large operators will usually observe the same direction of over- or undersupply in their balancing area as in the nationwide power grid. This significantly reduces costs for balancing but in turn also reduces possible earnings for a differently directed supply state in their own area in relation to the nationwide grid.

However, there are different market prices for power, depending on the time of acquisition, which are shown in Fig. 2.4. Furthest in advance is the *forwards* market, where electricity
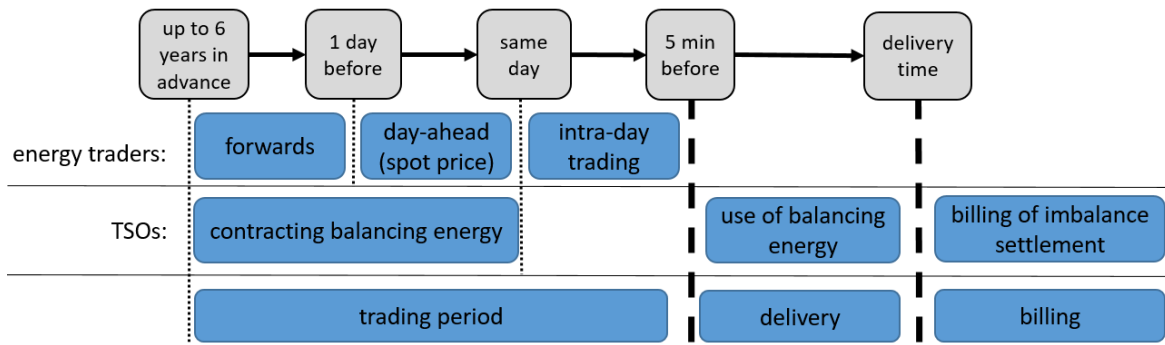
Figure 2.4: Chronological representation of the different markets for German electricity. Figure adapted from *SMARD* (2021a).

can be traded up to six years in advance. Along with the day-ahead power schedules for each balancing group, each power provider can trade their scheduled power on the market. Power consumers can similarly buy power to fit their scheduled consumption of the next day. Supply and demand form the *day-ahead* price, also called spot price, which can be traded on the day before the delivery.

Usually the previously given power schedule changes during the day, even if only sligthly. Additional power can then be traded for the *intra-day* price, again determined by supply and demand. If, for example, the wind farm is predicted to produce more power according to the latest power prediction runs, the direct marketer tries to sell the extra power as well. If the wind farm is predicted to generate less power than declared in the power schedule, the direct marketer needs to buy back some power from other sources in order to not have to pay balancing costs later. Intra-day trading is possible until five minutes before the delivery time.

The third price on the power market is the already discussed balancing price. After intra-day trading closes, any differences between actual power generation and the intra-day adapted forecast have to be paid for with balancing costs.

Depending on the individual contracts, the spot price is what is typically paid by the direct marketer to the wind farm operator for each generated MWh of power. Some contracts even include a flat rate for balancing services per generated MWh paid by the wind farm operator. Additional gains and losses from the variable sources are usually borne by the direct marketers. These do not *only* include balancing costs, but also, e.g., clever intra-day trading. That means, for the evaluation of the total financial gains and losses by different power predictions, balancing costs are only one contribution. However, for this study there are no financial information of direct marketers available. Therefore, balancing costs are the only metric evaluated in this thesis.

Figure 2.5 displays a two-dimensional (2D) frequency distribution of the balancing price and the corresponding balancing volume throughout Germany between 1 January 2017 and 31 December 2019. On the left column (quadrants i & iii) are situations with excess power in the grid, on the right column (quadrants ii & iii) situations with a power deficit. The top row (quadrants i & ii) consists of situations, when one would need to pay to get extra power into one's balancing group and get paid for excess power one distributes. On the bottom row (quadrants iii & iv) are situations, when, reversely, one would pay to distribute excess power and get paid for taking extra power to compensate one's balancing group. Most of the points are in quadrant (ii) and (iii), indicating a good correlation between the direction of a necessary balance and the sign of the balancing price. That means it is common to get paid

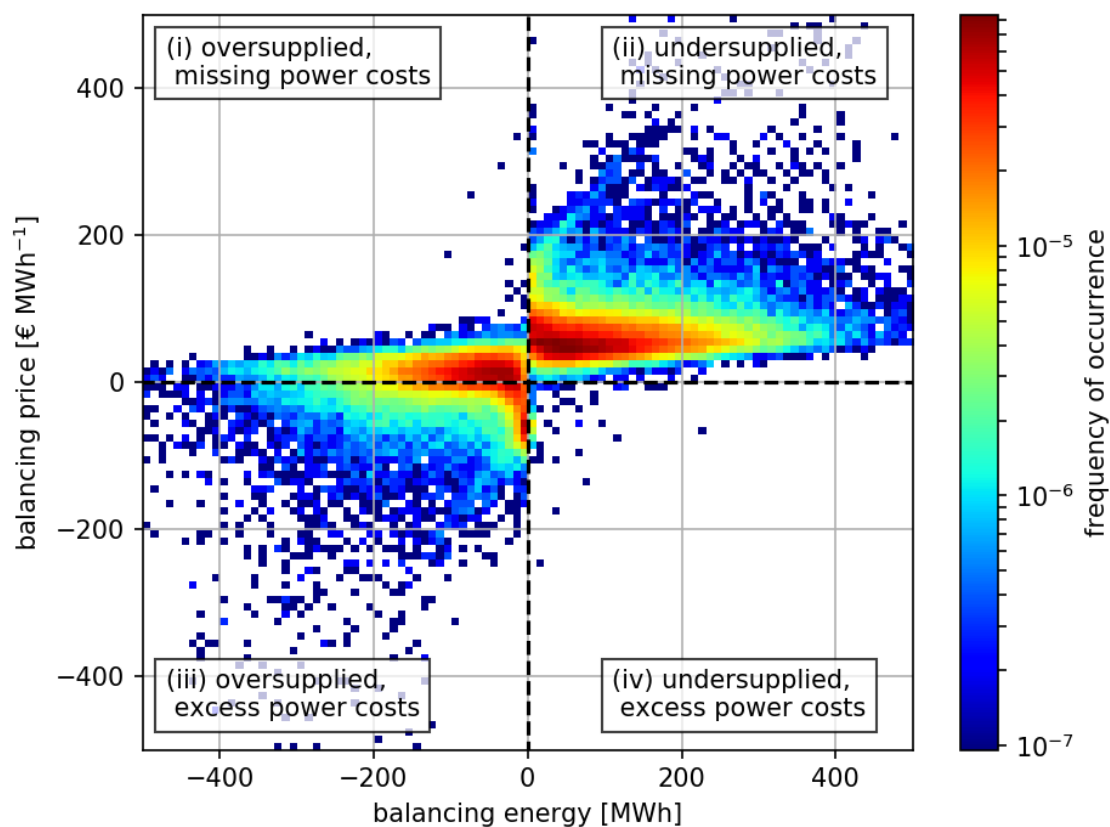Figure 2.5: Two-dimensional (2D) frequency distribution of balancing prices and corresponding balancing volume between 2017 and 2019 in Germany.

for extra power generated and to have to pay for missing power. However, the data points are skewed in the direction of positive prices. Even if operating storage facilities take power out of the grid, it is still a service, which costs money and never works at 100 % efficiency.

3 Instrumentation, weather prediction model and power forecasting tools

This Chapter is intended to summarise the instrumentation used in the acquisition of observational data and the development of wind power predictions for this thesis. In Sect. 3.1, the technical functionality and physical principle of power generation through wind is described. Section 3.2 describes multiple methods to observe the wind velocity in the troposphere and introduces the instruments used for this project. The Numerical Weather Prediction (NWP) model is introduced in Sect. 3.3. This models output serves as main input for the used wind power prediction tool LoadManager®, which is detailed in the end of this Chapter in Sect. 3.4.

3.1 Wind energy turbines

Wind energy turbines (WET) are the most prominent technique to generate power from wind, however, there is a wide range of different devices available. While most operational turbines on land still have hub heights of around $100\,\mathrm{m}$ above the surface, the latest developed platforms reach hub heights of up to $170\,\mathrm{m}$ with blade tips up to $300\,\mathrm{m}$ above the surface. However, the growth of WET has practical boundaries. The material requirements for higher platforms increase disproportionately and logistics for transporting the larger parts from the manufacturer to the wind farm site get more complicated (*Quaschning*, 2018). The increased efficiency remains beyond the necessary effort.

The power of the wind P_wind passing through an area A increases by the third power of the wind speed v:

$$P_\mathrm{wind} = \frac{1}{2} \cdot \rho \cdot A \cdot v^3, \tag{3.1}$$

where ρ is the density of the air. A fraction of that power is converted into electrical energy by the WET, thus slowing the wind down. The ratio of power taken from the wind (P_N) to the power contained in the wind (P_wind) is called coefficient of power c_P (*Quaschning*, 2016). The theoretical threshold of c_P is not 1 but $16/27 \approx 0.59$ (*Betz*, 1927). The rest of the power contained in the wind is required to push the air past the rotor blades. This means that even the most efficient WET cannot take more than $59\,\%$ from the power contained in the wind (*Betz's law*). While the *Betz's law* has been acknowledged as a universal rule for all wind turbines, one recent study points out unnecessary assumptions (*Strauss*, 2021). According to that, *Betz's law* is only applicable to WET with cowlings, since it assumes a constant internal velocity for the wind flow passing the WET. However, most commercial WET do not have cowlings and without that assumption, *Strauss* (2021) derive a higher c_P of $2/3 \approx 0.67$. While both theoretical possible values are not reached by commercial WET yet, efficient turbines typically have a power coefficient around 0.5 (*Quaschning*, 2016).

The drive of rotor blades via wind can be achieved in two ways. Firstly, by resistance: that means the wind blows directly at the rotor blade and pushes it away. Secondly, rotor blades can be moved by the lift effect, based on Bernoulli's principle. That means the air blows at the edge of the rotor blade and streams around it. For the lift effect the shape of
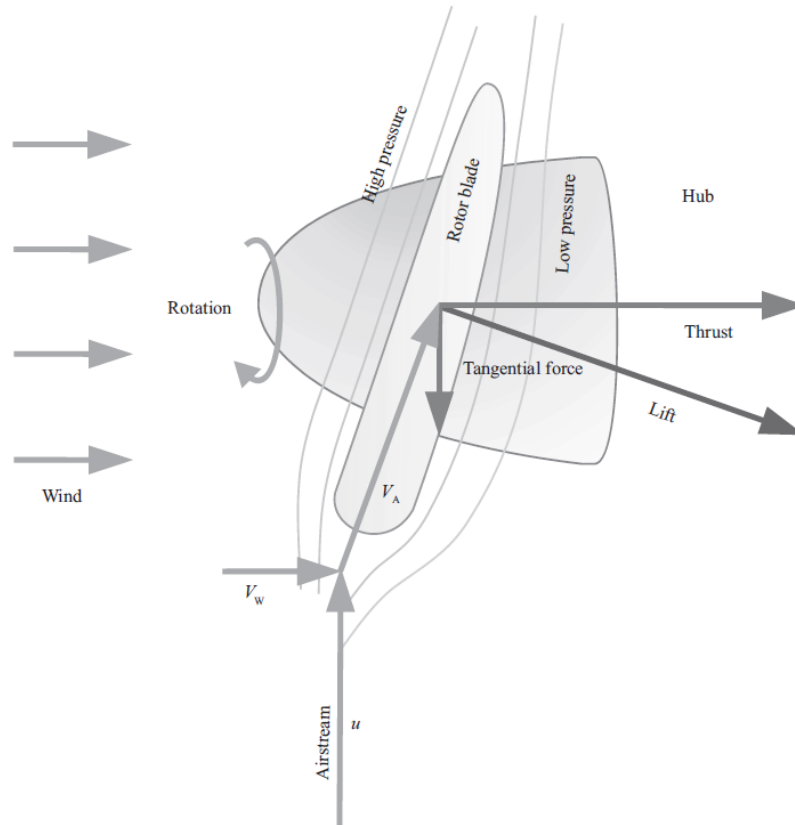
Figure 3.1: Wind velocities and resulting forces on a rotor blade (*Quaschning*, 2016).

the blade is essential: the distance for the air stream u around the blade is longer on the upper side than on the lower side (see Fig. 3.1). This creates lower pressure on the upper side and higher pressure on the lower side, resulting in a lift effect vertical to the rotor blade (*Quaschning*, 2016). *Betz* (1927) concluded that the resulting power of the WET is up to 54 times larger using the lift effect than using the resistance effect. Therefore, modern WET almost exclusively work with the former. During periods of power curtailment, when the power generation of wind turbines has to be decreased, the rotor blades of a number of wind energy turbines can be sligthly tilted. When the distance of the air stream around the blade is of equal length around both sides of the blade, the lift effect is prohibited and, thus, the power generation ceases. A more detailed description of the functionality of WET can be found in Chapter 6 of *Quaschning* (2016).

In Germany the measurement of wind velocity is required on every wind turbine and anemometers are usually located at the back of the nacelle. This serves the operators of the wind farm to identify situations when the WET has to be switched off. Wind turbines are only active in a certain wind speed window between cut-in wind speed and the cut-out wind speed. Table 3.1 shows an overview of the technical data of the Vestas V80-2.0 MW turbines used in the wind farm, which is considered in this project. The cut-in wind speed of the Vestas V80-2.0 MW is $4.0\,\mathrm{m\,s^{-1}}$. Increasing wind speeds increase the output power up to the so-called rated wind speed of $16\,\mathrm{m\,s^{-1}}$, where the WET reaches its maximum power generation. For further increasing wind speeds, the power output stays constant until the cut-out wind speed of $25\,\mathrm{m\,s^{-1}}$ is reached (see Fig. 3.2). Above the cut-out wind speed the rotor needs to be turned off or adjusted in position to avoid damage to the machine. Usually

Table 3.1: Excerpt of Vestas V80-2.0MW technical data (*Vestas*, 2011).

Operating data		Dimensions	
rated power	$2.0\,\mathrm{MW}$	hub height	$100\,\mathrm{m}$
cut-in wind speed	$4.0\,\mathrm{m\,s^{-1}}$	rotor diameter	$80\,\mathrm{m}$
rated wind speed	$16\,\mathrm{m\,s^{-1}}$	nominal revolutions	16.7 per min
cut-out wind speed	$25\,\mathrm{m\,s^{-1}}$	operational interval	10.8-19.1 per min
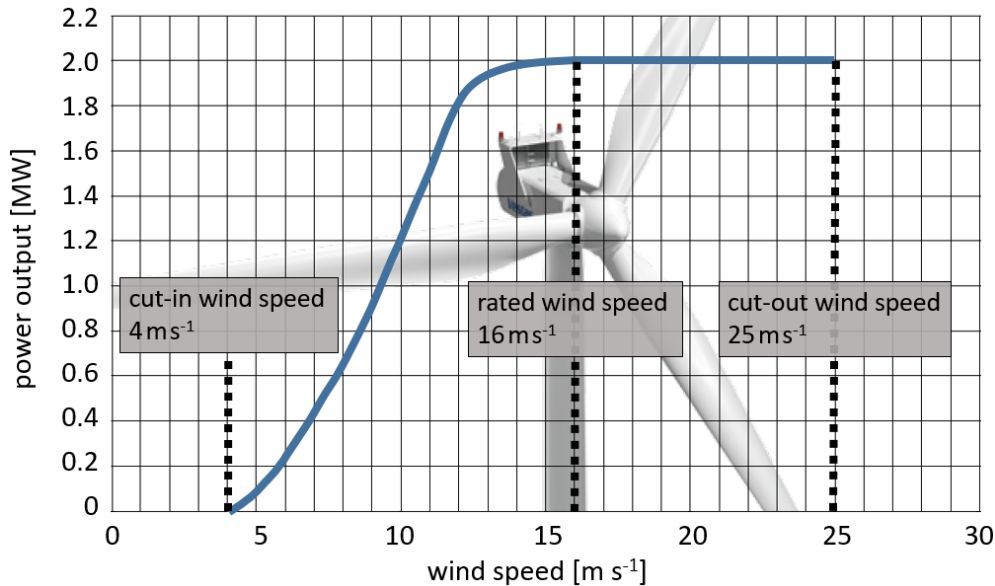
Figure 3.2: Power curve of the Vestas V80-2.0 MW wind turbine used in the wind farms Beeskow I & II (see Sect. 4.1.2) (adapted from *Vestas*, 2011).

it is sufficient to slightly adjust the pitch of the rotor blades so that the distance around both sides of the blade is similar and, therefore, the lift effect negated. The WET are then still oriented into the direction of the wind, but the reduced lift effect on the blades greatly reduces the thrust on the rotor and shuts down the energy generation.

The wind velocity measurements on the back of the nacelle further serve the administrative agencies to control external shutdown criteria. For example, one of these is to protect bats. During certain summer nights with mild temperatures and no precipitation, the rotors have to be turned off for wind speeds below $5\,\mathrm{m\,s^{-1}}$ (Ministry for Agriculture, Environment and Climate Protection Brandenburg, 2010). Turn-off periods, including the observed wind velocity time series, have to be documented well and handed over to federal authorities in the fall of each year.

Since wind velocity measurements are routinely done on every wind turbine and wind farm operators have easy access to their observations it further encourages the evaluation of their use to this project.
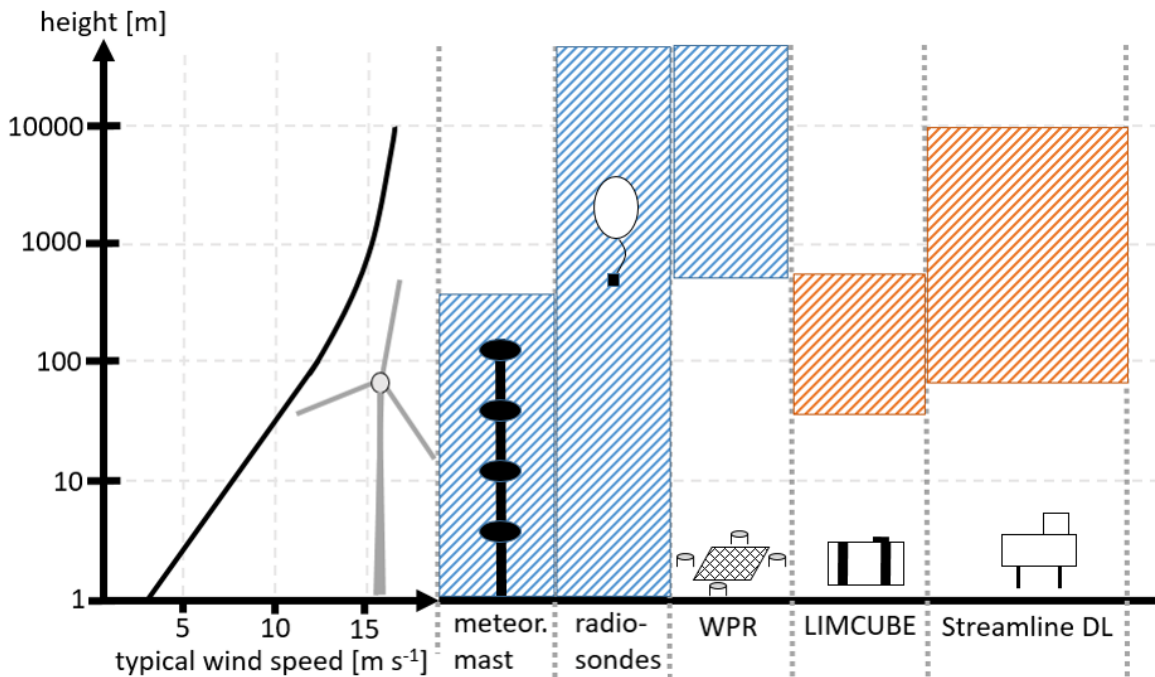
Figure 3.3: Illustration of the different wind velocity measurement methods introduced in Sec. 3.2 in comparison with a typical wind speed profile on a logarithmic height axis (y-axis). WPR refers to wind-profiling radars, LIMCUBE and Streamline DL are Doppler lidar instruments introduced in Sect. 3.2.2. Shaded areas indicate the measurement ranges for the instruments; for orange shading the measurement depends on tracer particles, for blue shading the measurement does not depend on the atmospheric composition.

3.2 Wind measurements in the troposphere

There are three commonly used methods to measure the wind velocity in the lower part of the atmospheric boundary layer, where wind turbines are located (see Fig. 3.3). Meteorological towers (MT) equipped with anemometers (typically sonic anemometers, see Sec. 3.2.3) can reach from the surface up to a few hundred meters. Under certain conditions, active ground-based remote sensing instruments can measure wind profiles from a few ten meters above the ground up to the tropopause and radiosondes provide wind velocity measurements even above the tropopause.

For the lowest layer, from the surface to up to a few hundred meters, meteorological towers, also called meteorological masts, offer a conceptually simple way to measure wind velocities. Once the mast is erected, it can provide continuous wind and other measurements of its site up to its height with high reliability. The highest meteorological masts are the IAP Meteorological Tower in China and the Amazon Tall Tower Observatory in Brazil, both measuring 325 m. Weather masts in Central Europe include the Hamburg Weather Mast at 300 m, the mast of the Royal Dutch Meteorological Institutes located in Cabauw at 213 m and the meteorological tower at the atmospheric boundary layer field site of the DWD in Falkenberg (near Lindenberg) at 99 m.

Alternatively, radiosondes offer measurements from the surface up to altitudes well above the tropopause and are not restrained to a specific location. Their wind velocity measurement is restrained to be calculated from their drift and the resulting wind profile usually is not a vertical profile directly above the launch site. Commercial radiosondes offer wind
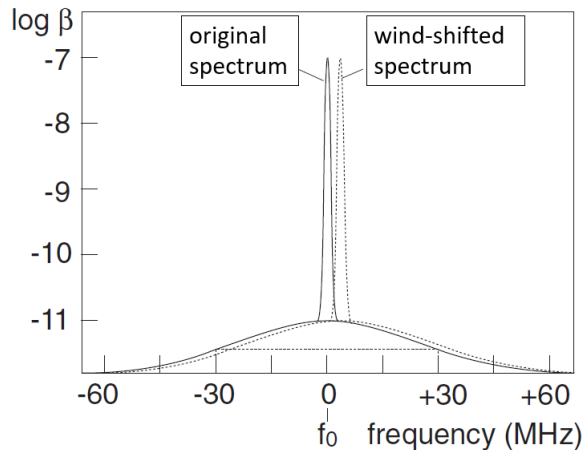
Figure 3.4: Schematic representation of the original (solid) and wind-shifted frequency backscatter distributions relevant in the Doppler lidar technique. The narrow spike represents the aerosol backscatter, the underlying broad peak the molecular backscatter. The indicated shift toward higher frequencies relates to the wind approaching the Doppler lidar (adapted from *Werner*, 2005).

observations with an uncertainty of $0.15\,\mathrm{m\,s^{-1}}$ (Vaisala RS41-SGP). A disadvantage is the coarse temporal resolution in the order of a few hours per day, depending on the site, since the operation is not entirely automatic.

As another approach for wind observations, ground-based remote sensing is able to provide vertical profiles of wind velocity with a high temporal resolution ($< 1\,\mathrm{min}$) from a few tens of meters above the surface up to the tropopause, depending on the instrument. Rapid developments in optical communication technology and their commercial availability during the last decades shaped the near-infrared frequency range for application in atmospheric remote sensing. Lidar (light detection and ranging) is a remote sensing method similar to radar (radio detection and ranging) but operating at nanometer to micrometer wavelengths. There are typically two contributions to the backscattered laser signal; Rayleigh-scattered molecular backscatter and Mie-scattered aerosol backscatter. Figure 3.4 shows a schematic representation of the two backscatter contributions. While the molecular backscatter phase function is Gaussian shaped, the aerosol backscatter phase function produces a narrow peak (*Werner*, 2005).

Sensitive lidars with infrared frequencies are able to extract information about wind velocities from the molecular backscatter (*Werner*, 2005). However, in heights where aerosol particles are present, frequency analysis based on the narrow aerosol backscatter peak is easier and therefore preferred (see Fig. 3.4). By observing the Doppler frequency shift between transmitted and received radiation, Doppler lidars determine the velocity of these aerosol particles in the direction of the instrument. The Doppler frequency shift of the received radiation is then converted to the line-of-sight velocity of the aerosol particles. Since the terminal fall velocity of these small particles is negligible, the calculated particle velocity is assumed to be equal to the wind velocity. Aerosol particles thus serve as tracers for the wind velocity.

To obtain a three-dimensional wind vector, measurements in at least three independent directions are needed. Therefore, the horizontal wind speed measured by the DL is always a spatial average over a certain horizontal area, assuming a horizontally homogeneous wind field, and a temporal average over the time needed to redirect the beam between the different directions. *Frehlich* (2001) conclude that typical wind velocity errors range from $0.1\,\mathrm{m\,s^{-1}}$ to $0.5\,\mathrm{m\,s^{-1}}$. Table 3.2 shows technical parameters for the DL instruments utilised in this

thesis. The lowest measurement range of DLs typically is several tens of meters, the maximum possible height depends on the transmitted power and the composition of the atmosphere. Generally, commercial Doppler lidars are able to derive wind velocities as long as a sufficient amount of tracer particles is available, which includes at least the atmospheric boundary layer as well as optically thin clouds.

For a complete overview, the Wind-Profiling Radar (WPR) technique is introduced briefly. Wind-Profiling radars have been in operation for a few decades and measure wind velocity reliably independent of the synoptic situation. They use wavelengths in the ultra-high frequency (0.3-3 GHz) and very-high frequency (30-300 MHz) regimes to exploit Bragg scattering from refractive index fluctuations due to variations in humidity and temperature, omnipresent in both the clear and cloudy atmosphere (*Röttger and Larsen*, 1990). Although Rayleigh scattering from particles also contributes to the signal, the contribution by Bragg scattering dominates for larger wavelengths (*Bühl et al.*, 2015). WPR emit radar pulses and measure the Doppler shift of the returned signal, which can directly be related to the radial velocity in the beam direction (*Larsen and Röttger*, 1982). Measurements over multiple beam directions yield the three-dimensional wind vector. The vertical wind component is often measured by pointing the beam in zenith direction. A more in-depth description of the measurement principle can be found in Chapter 21 of *Röttger and Larsen* (1990). The DWD operates four WPR in Germany; in Lindenberg, Bayreuth, Ziegendorf and Nordholz. They operate in a low and high mode, giving wind profiles from 500 m to 9.3 km above the surface and 5.5 km to 16 km above the surface, respectively, with vertical resolutions of 150 m and 330 m (*Päschke et al.*, 2014). The possibility to measure in clear and cloudy atmosphere with a WPR is a great advantage compared to every other ground-based remote sensing wind measurement instrument. However, WPRs are immobile, due to their large size of 140 m^2 effective antenna area, and rather susceptible to other transmissions in their UHF or VHF range.

3.2.1 Wind measurements for wind energy purposes

To conclude which instruments are useful in the scope of the project of this thesis, the covered altitude range, temporal resolution and mobility are compared.

Firstly, the different measurement ranges between the instruments are competitive from the wind energy perspective. Most wind turbines have hub heights between 80 m and 150 m above ground and only atmospheric layers that are near the hub height should have an influence on the hub height wind field. This means that most instruments mentioned, except the WPR, do cover this height range. Radiosonde and meteorological mast measurements can have a higher range resolution of only a few meters compared to tens of meters for the DL.

Secondly, the temporal resolution is important. So far, wind measurements at wind farm sites have usually been made before the planning phase, to evaluate the mean wind field over a representative time range. For this project, wind measurements with a temporal resolution higher than few tens of minutes are required in order to use them as additional input for wind power forecasts. These have a typical temporal resolution of 15 min and are updated until up to 5 min before the forecast time. The assimilated measurements need to have a similar resolution to that, which is possible for remote sensing and in-situ instruments. Radiosondes, with typical launch frequencies between 3 and 12 hours, are not suitable, because their temporal resolution is not sufficient to capture the diurnal cycle and rapid changes in the wind field.

Lastly, for wind farm site assessments the instruments only need to be in place for limited time. The sites can be remote, in forested or even off-shore locations. Traditionally, measurements are no longer needed once the WET are installed. However, for the topic of

this thesis, the data assimilation into the power forecast, continuous measurements beyond the wind farms construction are required. Measurements from a meteorological mast and DL observations would both fit well in regards to their spatial and temporal resolution. The aspiration of this work is to evaluate and improve the present power forecasting method and to be able to expand this improved method to other wind farms in the future. Wind masts cannot be relocated easily to other sites and the deployment of new wind masts is costly and requires constructional and bureaucratic effort. Therefore, a mobile solution like a DL, which typically is not bigger than $0.5\,\mathrm{m}^3$ (see Tab. 3.2), seems to be more suitable. In the following, the technical specifications of two commercial DL are compared.

3.2.2 Doppler lidar instruments

Figure 3.5 shows the two DL instruments used on this project, which will be introduced in this section. Subsequently, main differences of both DLs for users of the wind energy sector are summarised.

The HALO Photonics Streamline Doppler lidar XR (Streamline DL, shown in Fig. 3.5b) is a ground-based lidar to measure the three-dimensional wind velocity, derived from observations pointing into different azimuth directions. It emits laser pulses at a center wavelength of $\lambda = 1.5\,\mathrm{\mu m}$, which are scattered by aerosol particles and small cloud particles. The MOLRAO observatory of the DWD has been operating their Streamline DL since 2012. They employ a scan pattern doing velocity-azimuth display (VAD) scans in $n{=}24$ directions, spaced $\alpha{=}15°$ apart, each $\Theta{=}15°$ off zenith (see Fig. 3.6a). It takes $5\,\mathrm{s}$ to emit $75\,000$ laser pulses for each beam direction before the scanner moves to the next azimuth angle. To finish one VAD scan this totals to approximately $3\,\mathrm{min}$. The measurement range of the Streamline DL extends up to a height of $12\,\mathrm{km}$ as long as a sufficient amount of tracer particles are present in the measurement volume. A more detailed description of the scanning pattern can be found in *Päschke et al.* (2014). A list of the technical parameters is compiled in Tab. 3.2.

The measurement principle of the Leosphere Windcube Doppler lidar WLS8 (named LIMCUBE, shown in Fig. 3.5a) is very similar to the one of the Streamline DL. It also uses infrared radiation (at $\lambda = 1.54\,\mathrm{\mu m}$) and measures the Doppler frequency shift of the returning pulses. The pulse length of the LIMCUBE is $75\,\mathrm{ns}$, providing a vertical resolution of $20\,\mathrm{m}$ (see Tab. 3.2). The maximum range covered by the LIMCUBE is $500\,\mathrm{m}$, which results in an incomplete atmospheric boundary layer observation in the middle latitudes, especially in summer during daytime. However, for wind energy applications this does not pose a big disadvantage, since typical heights of wind turbines are well below $500\,\mathrm{m}$.

In contrast to the more complex scanning patterns of the Streamline DL the LIMCUBE measures radial wind velocities $V_{r,1-4}$ in the four main wind directions, each $\Theta = 15°$ off zenith. The horizontal wind vector components u and v are derived from the radial wind velocity of opposing lidar beams and the vertical wind velocity w is calculated by averaging over four consecutive measurements. Figure 3.6 illustrates the LIMCUBEs scanning pattern and the following equations show the calculation of the three wind components:

$$V_{r,1} = u\sin(\Theta) + w\cos(\Theta) \qquad\qquad V_{r,3} = -u\sin(\Theta) + w\cos(\Theta),$$
$$V_{r,2} = v\sin(\Theta) + w\cos(\Theta) \qquad\qquad V_{r,4} = -v\sin(\Theta) + w\cos(\Theta),$$

and therefore:

$$u = \frac{V_{r,1} - V_{r,3}}{2\sin(\Theta)} \qquad v = \frac{V_{r,2} - V_{r,4}}{2\sin(\Theta)} \qquad w = \frac{V_{r,1} + V_{r,2} + V_{r,3} + V_{r,4}}{4\cos(\Theta)}. \tag{3.2}$$
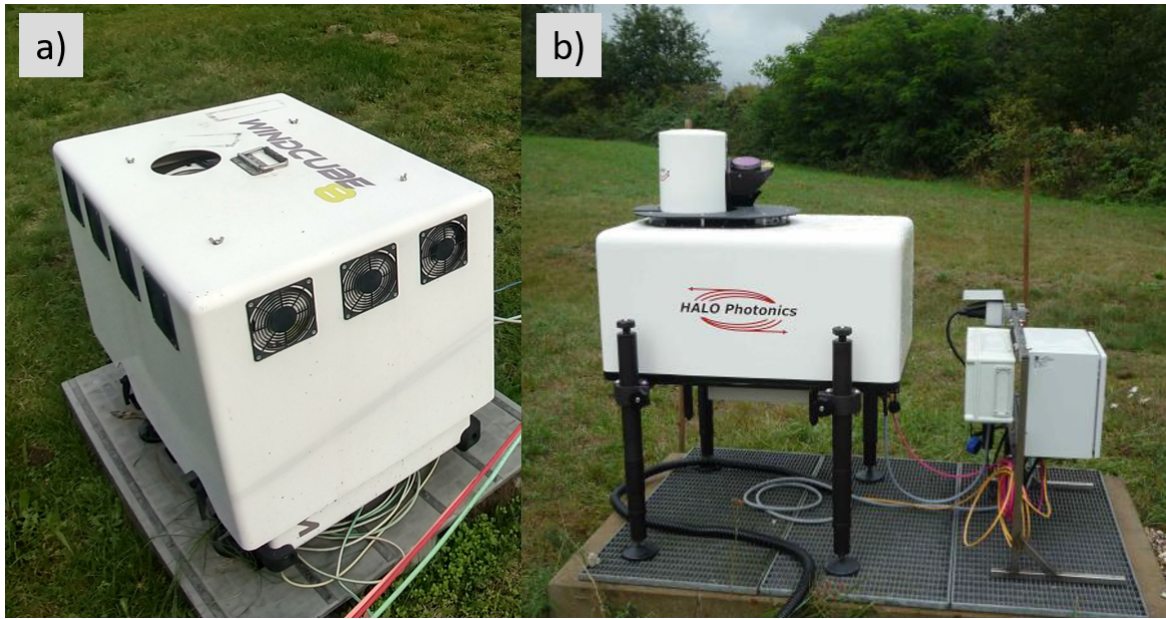
Figure 3.5: Doppler lidar instruments Leosphere Windcube WLS8 named "LIMCUBE" (a) and Halo Photonics Streamline DL (b) operating in Lindenberg. Pictures taken by Pablo Saavedra Garfias (LIMCUBE) and from DWD website (18.10.2021).

Both DLs are user-friendly, also for non-technical personnel. The setup requires only electricity and a data connection, either via ethernet cable to another PC or storage device or wirelessly via a mobile connection. The technical parameters of the Streamline DL are more customisable, with no strict observation patterns specified. LIMCUBE observations are always angled at 15°, pointing into the four main wind directions, averaged over 20 m vertical measurement volume and ranging up to 500 m above the surface. All of these parameters are configurable for the Streamline DL, offering more flexibility for the operators. For the configuration of the Streamline DL detailed in Tab. 3.2, which was also applied for the observational data analysed in this thesis, this results in a lower vertical resolution of the Streamline DL compared to the LIMCUBE. The different maximum ranges for both devices are far above the typical WET height. Regular maintenance for DLs is typically low, barring the complete failure of single parts. Both DL do not have wear parts and do not need regular calibrations. The cost of a new Windcube WLS8 in 2012 was about 150 000 €. The Streamline DL, also acquired in 2012, was about 276 000 €. It was upgraded later by DWD to its XR capabilities and modified in certain parameters. Today, similar devices are more compact and priced lower, e.g., the Leosphere Windcube v2.1 with a range of 300 m is only half the size of the Windcube 8 and available at roughly 100 000 €.
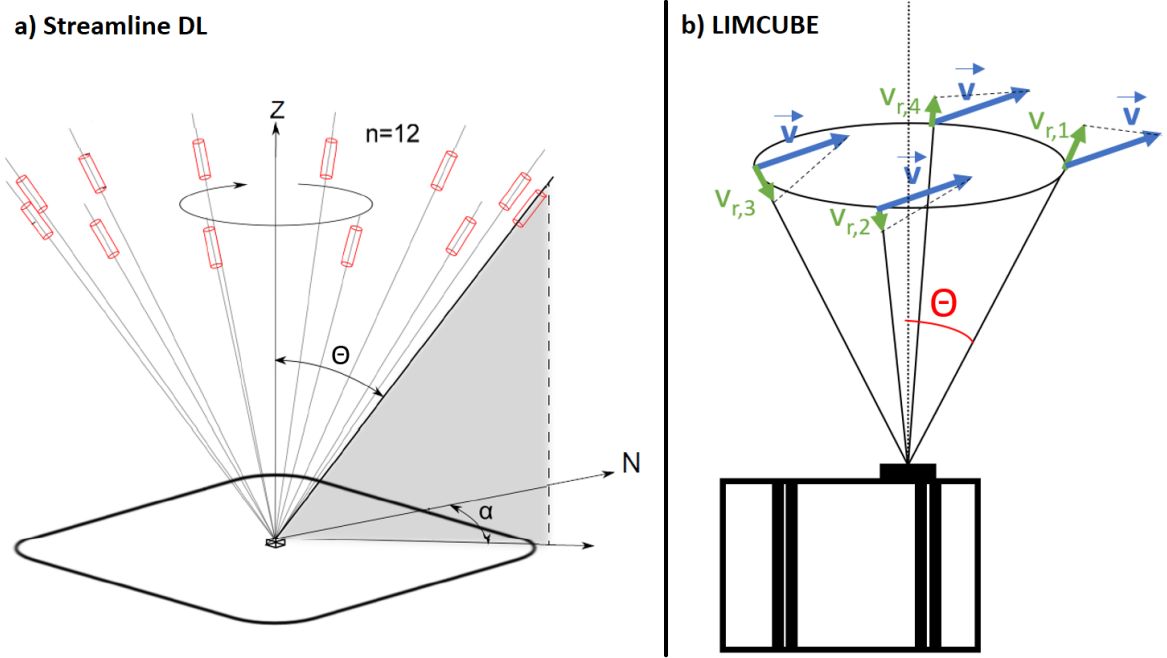
Figure 3.6: a) Illustration of the Streamline DL VAD scanning pattern for n=12 beam directions (adapted from *Päschke et al.*, 2014). b) Illustration of the Leosphere LIMCUBE scanning pattern.

Table 3.2: Technical parameters of the HALO Photonics Streamline and Leosphere Windcube (LIMCUBE) Doppler lidars, as most commonly used during the discussed measurement.

parameters	Streamline DL	LIMCUBE
wavelength [µm]	1.5	1.54
first range gate [m]	69.5	40
range [m]	12 000	500
pulse length [ns]	160	75
range gate length [m]	46	20
number of range gates	250	25
dimensions (L x W x D) [m³]	0.63 x 0.53 x 0.40	0.95 x 0.65 x 0.55
weight [kg]	85	90

3.2.3 Sonic anemometers

Ultrasonic anemometers (usually called "sonic anemometers") are a further possibility to measure wind speed in-situ. They consist of three oppositely installed pairs of sound transmitters and receivers and measure the time taken for a sonic pulse to travel the known distance between transmitter and receiver back and forth. The speed of the pulse depends on the speed of sound superimposed by the wind speed along this direction. Since the speed of sound is equal in both directions it cancels out and only the wind speed remains measurable. Sonic anemometers measurements have a high temporal resolution of 20 Hz or better, which makes them well suited for a wide range of applications (*Nfaoui*, 2012). Typical measurement uncertainties of sonic anemometers are well below $0.1\,\mathrm{m\,s^{-1}}$ for the wind speed and below 1° for the wind direction. In addition, sonic anemometers are virtually free of maintenance and much cheaper than DLs, costing about $5\,000\,€$. Their low cost is only offset by the additional cost of erecting a mast for measurements above the surface in hub height levels.

As previously mentioned, (sonic) anemometers are also installed on WET. However, since they are located at the back of the nacelle, they are always on the lee side of the rotor and, therefore, influenced by the rotor's wake effects. *Allik et al.* (2014) investigated deviations between wind speed data measured with nacelle-mounted anemometers on small wind turbines with hub heights of 16 m and 18 m and, as a reference, anemometers mounted on meteorological towers upwind of the WET. Both mechanical cup anemometers and sonic anemometers on the WET were used to estimate the influence the rotor has on the wind speed measured on its lee side. Additionally, the correlation between the wind speed at the back of the nacelles and on the MT was evaluated. For the field experiment by *Allik et al.* (2014), the mean wind speed measured on the WETs was systematically lower (with values ranging between $1.45\,\mathrm{m\,s^{-1}}$ to $2.25\,\mathrm{m\,s^{-1}}$) than on the reference, revealing a significant difference. However, the correlation between the observations at the nacelle and on the mast in the same height was found to be high with a Pearson correlation coefficient of 0.874. This means that while the absolute wind speed observations at the back of the nacelle are significantly influenced by the rotor blades, they still represent the pattern of the wind speed well. Furthermore, it was found that sonic anemometers were more similar to the mast measurements than observations by the cup anemometers.

Contrarily, *Antoniou and Friis Pedersen* (1997) found that for a considerably larger WET of 55 m hub height with 1 MW rated power the difference between the nacelle measurement relative to an undisturbed mast measurement depends on the wind speed. This difference is minimised for high and low wind speeds and increased for moderate wind speeds. However, this behaviour was found to change with the location of the anemometer on the nacelle, the characteristics of the rotor and the shape of the nacelle, preventing an easy generalisation. To estimate the WETs rotors influence on the nacelle sonic anemometer observations in this study, they will be compared to the observations from the DL and the MT in Sect. 4.2.2.

3.3 Met Office Wind Production Forecast product

The Met Office Wind Production Forecast (WPF) is a product based on NWP model data from the UK-based Met Office. The LoadManager® power prediction tool of the company LEM-Software used in this project (see Sect. 3.4) is currently exclusively based on WPF model data. For Germany, the WPF product is post-processed from the European domain of the renowned Unified Model (UM) by Met Office. This UM by Met Office is a seamless weather and climate forecast model for regional and global applications. The NWP configurations provide grid resolutions from 1.5 km to 10 km. For climate simulations the grid spacing
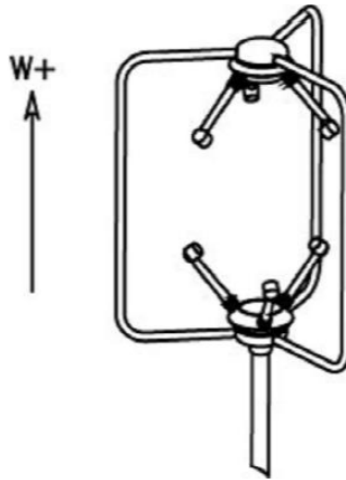
Figure 3.7: Conceptual sketch of a 3D sonic anemometer (adapted from *Ramos et al.*, 2012).

reaches up to 300 km. The temporal scales range from hours to centuries (*Brown et al.*, 2012).

Figure 3.8 shows a conceptual sketch of the creation of the WPF product. The model forecast starts with the Global model. The Global model has a grid spacing of 10 km and a forecast length of six days. Model runs are initiated every six hours. The Global model provides the boundary conditions for the subsequent European domain Euro4 model. The Euro4 model has a grid spacing of 4.4 km and a temporal resolution of 1 hour for the first 54 hours, with model runs performed every 6 hours. Both the Global model and the Euro4 model are based on the Met Office UM. Since the WPR in Lindenberg is part of the European network of operational profilers E-PROFILE, its observations are made available to NWP forecasters and get assimilated into the Global model routinely (e.g., *Hooper et al.*, 2013; *Hisscott*, 2019). The assimilation of WPR data shows a significant improvement of model reliability due to the instrument's high temporal and spatial resolution and consistent measurements, shown for a COSMO (*Necker et al.*, 2018) and a MetOffice Unified Model (*Illingworth et al.*, 2015) configuration.

However, the Euro4 model is a downscaling model; that means it does not perform its own separate observations assimilation cycle. It still benefits from the observations, but via the boundary conditions passed by the Global model as illustrated in Fig. 3.8.

Although the Euro4 grid spacing is 4.4 km, its real terrain resolution is coarser. Numerical models use a flattened orography to avoid numerical instability, smoothed over 2-3 grid lengths. However, this resolution is too coarse for wind energy applications. Orography on sub-kilometer scales can influence the wind field considerably. Therefore, Met Office provides the WPF product. It is a commercial service targeted mostly at the wind energy sector. It includes a series of corrections for weather parameters of interest to account for local orography. The WPF post-processing can be applied to any of Met Offices NWP configurations, for the data in this project it is used on the Euro4 model (personal communication with Justin Krijnen and Ian Pearman, Met Office, 2019 and 2020).

To generate site and hub height specific weather forecasts, two interpolations are done within the WPF. First, a horizontal interpolation based on nearest neighbour and site specifics is performed. For example, a site on land would be characterised and corrected as such, even if the nearest neighbour grid point is in the ocean. Secondly, a vertical interpolation of the weather parameters based on the nearest model level is done.
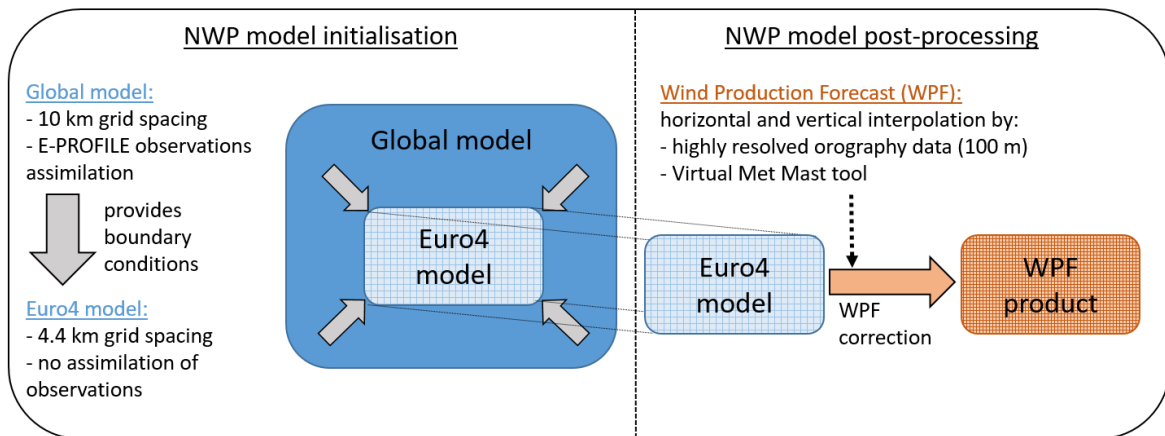
Figure 3.8: Conceptual sketch of the creation of the Wind Production Forecast product (WPF). Left side: The MetOffice Global model with a coarser grid size of 10 km benefits from the assimilation of observational data. The Global model then provides boundary conditions for the Euro4 model (grid size 4.4 km), passing on the benefits of observation assimilation. Right side: Afterwards, the WPF post-processing tool is applied to the Euro4 model data, adjusting for the coarse model resolution with highly resolved orography data and performing the Virtual Met Mast correction (see Sect. 3.3). In the Figure, this is implied by the finer grid pattern for "WPF product" compared to "Euro4 model".

For wind interpolations the MetOffice Virtual MetMast (VMM) tool is used. The VMM is a method to derive site-specific wind information from NWP data as an alternative to physical weather masts (*Standen et al.*, 2017). Introduced to provide an easy way of wind farm site assessments by using it on historical NWP data sets, it is a similarly powerful tool to use on weather forecast data. The problem with NWP data sets concerning site specific wind profiles originates from the roughness parametrization. NWP models need some kind of surface roughness parametrization to deal with sub-grid scale roughness elements, like small troughs and hills. The different UM configurations use the concept of an effective orographic roughness. The grid cells real roughness is increased to account for sub-grid scale orographic elements. This is a valid approach for NWP applications, because sufficiently high above the ground the observed airflow above complex orography appears similar to the airflow over flat terrain with a higher effective roughness length (*Grant and Mason*, 1990). However, one side effect of treating wind like this is a reduction of wind speed in the lower part of the atmospheric boundary layer. For the UK, *Standen et al.* (2017) examined this reduction at heights of 33.3 m, 93.3 m and 513.3 m above ground level. Even in not highly complex terrain, they still noticed a wind speed anomaly at 93.3 m above the ground, which is in the typical height ranges of WET. The VMM tool corrects this bias in two steps: First, by increasing the wind speed to a more consistent value and secondly, by adjusting for the coarse model resolution with highly resolved orography data (100 m gridded dataset) to better reproduce local wind flow. More details of the VMM method can be found in *Standen et al.* (2017). In 2021, after the time period investigated in this thesis, the Euro4 model was discontinued and superseded by an enhanced Global model. Now, the WPF post-processing is applied to the Global model.

For wind energy users, data from the WPF is interpolated to the location of the wind farm. The relevant wind variables in this data set are wind speed and direction in three heights (10 m, 70 m and 355 m). In addition, the WPF includes temperature at 2 m and 355 m, atmospheric surface pressure, relative humidity at the surface, global irradiance, cloud

Table 3.3: List of variables included in WPF product model data.

WPF variable	heights [m]
wind speed	10, 70, 355
wind direction	10, 70, 355
temperature	2, 355
cloud fraction	low, medium, high
relative humidity	2
atmospheric pressure	2
planetary boundary layer height	-

fraction and atmospheric boundary layer height (see tab. 3.3). However, for the considered time period of 1 January 2014 to 30 September 2019 some of these variables have larger gaps. To ensure a sufficiently long data set with consistent variables, those with larger gaps will be excluded for the later data analyses.

3.4 LEM-Software LoadManager® tool for energy forecasts

In this section the power forecast module LoadManager® will be introduced and the basic principle of Artificial Neural Networks (ANN) will be briefly explained. Further, the configuration of the LoadManager® for training and forecast and the modifiable parameters are presented. Subsequently, the integration of observational wind data into the power prediction process is discussed.

The LoadManager® is a modular software used to forecast time series of energy generation, energy consumption and prices. It was developed by the Leipzig based software company LEM-Software. In its original form the LoadManager® uses NWP model data to create power forecasts. Within the LoadManager®, there are multiple different methods to process input data, including simple linear regression models but also more complex ANNs.

The purpose of all these power forecast models is to connect the meteorological information from NWP models to the desired power forecast. The relation between wind speed and power generation for one Wind Energy Turbine (WET) is defined by its power curve (see also Fig. 3.2). For multiple WETs, wake effects within the wind farm add more complexity (see Sec. 4.1.3). Additionally, the input data from the NWP is affected by the model parametrisation (see Sec. 3.3).

Figure 3.9 outlines the LoadManager® method of power predictions. The ANN is used to learn the wind to power relationship along with systematic errors from historic NWP model data during its training (*Zimmer et al.*, 2012). When producing the power forecast, the LoadManager® prediction software uses the weighting function derived from the training on prognostic NWP data to correct for systematic errors.

The structure of an ANN is based on biological neural networks as both are methods, which evaluate input data. In-depth information on the basic structure of ANNs can be found in *Walczak* (2018); the most important details will be introduced in the following.

The ANN in use are multi layer perceptron networks (MLP), a class of feedforward networks. They are organized as a combination of one input layer, one or more hidden layers and one output layer. They work as universal approximators (*Menzies et al.*, 2015). Each layer consists of a number of nodes (neurons). For the input and output layers these typically
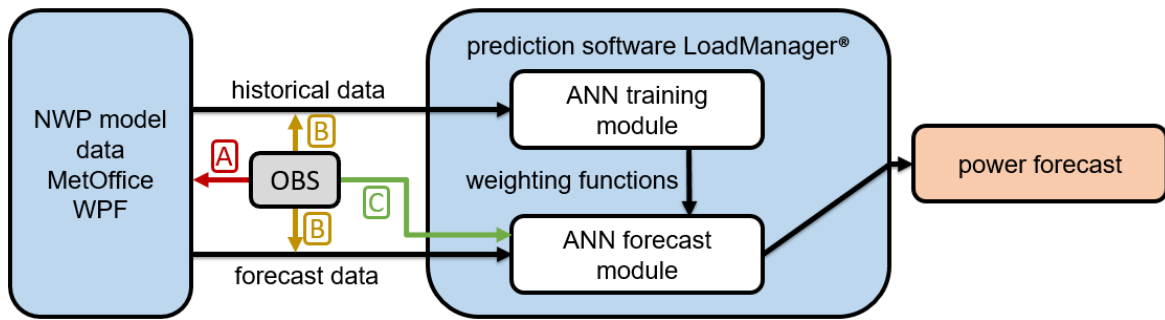
Figure 3.9: LoadManager® method of power predictions. Historical data from the WPF product is used to train a wind power prediction ANN in the LoadManager®. The training yields weighting functions, which are later applied on forecast data to generate the wind power forecast. Observational data (OBS) can be included on various points along the wind power generation chain. They can (A) be assimilated into the NWP model directly, (B) be used to evaluate a bias of historic model data and then correct the bias in the forecast data or (C) be used as additional input during the wind power forecast. More details on the use of observational data in Sect. 3.4.2.
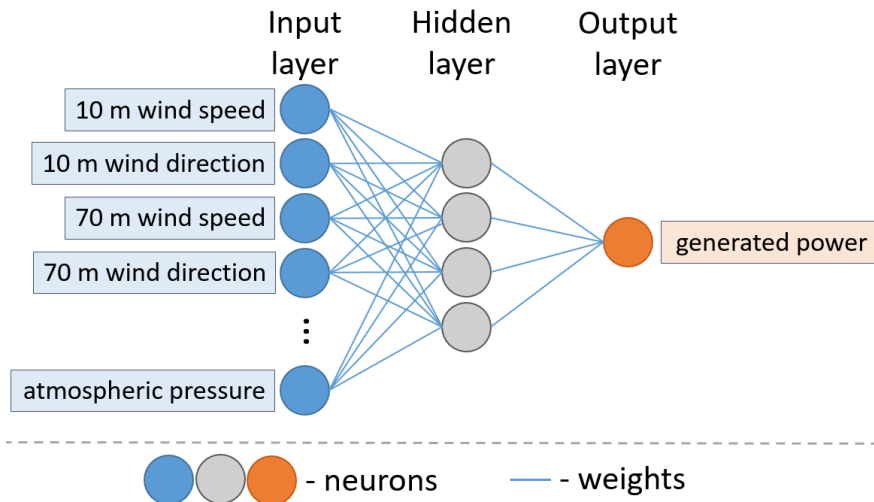

Figure 3.10: Conceptual sketch of an artificial neural network (ANN) with one hidden layer. Input neurons are fed by the chosen input variables; the output neuron represents the generated power.

correlate to the number of input and output values, respectively. For the LoadManager® power prediction ANN, the input neurons are the NWP model parameters. In Fig 3.10 they are displayed as blue circles. The output neuron represents the generated power, displayed as an orange circle in Fig. 3.10. Between input and output layers multiple hidden layers with different numbers of neurons can be located. These hidden layer neurons represent intermediate calculation steps between input and output. Neurons of neighboring layers are linked by weights. These weights are adjusted during the training process. This alteration of the weights represents the learning process of an ANN.

3.4.1 LoadManager® configuration

Firstly, training the ANN requires a training time series (LoadManager®: Learning Range (LR)) and a validation time series (LoadManager®: Recall Range (RR)). Data from the LR is used to adjust the internal weights of the ANN while data from the RR is used to

evaluate these weights. One cycle through the dataset is called epoch. The ANN is trained for multiple epochs, i.e. the same data points are evaluated several times at different stages of the training process. This is to avoid biases leaning towards the last time steps in the dataset. In the LoadManager®, the number of epochs is set implicitly by choosing the total number of samples trained; the number of epochs is then dependent on the length of the input time series. Additionally, the parameter "step size" defines the number of LR data points trained on, before the recall calibration is done on an equal amount of samples out of the RR.

The resulting configuration of the ANN is dependent on the LR data and also, indirectly, dependent on the RR data. Therefore, choosing the LR and RR is a sensitive task, because ideally the model needs to be trained, evaluated and later applied on data sampled in similar atmospheric conditions. A simple way to ensure good representation of occurred synoptic situations in all data sets is to alternately assign consecutive days in the LR and RR. This is preferable to splitting the data set in two halves to avoid biases due to seasonal or interannual variability. There are different power forecast model setups within the LoadManager®. This project will focus on the "Intervall-Optinet" ANN setup of the LoadManager®. Its most significant feature is that it alternately distributes the days of the dataset into LR and RR, instead of using two separate time series for LR and RR like most other LoadManager® setups.

Secondly, the optimal number of neurons and layers for the ANN needs to be determined. There are three possible layers with up to 100 neurons available. The user can choose to either set a fixed number of neurons per layer or to give a range of minimum to maximum neurons. If given ranges for the number of neurons, the LoadManager® repeats the complete training cycle for every possible combination of neurons per layer and in the end keeps the configuration with the smallest error. The measure of error can be set to either percentage deviation, Euclidean distance, absolute deviation or a mix of those. For this projects power predictions, the different configurations (see Sect. 4.4) are trained with eight to twelve neurons on the first hidden layer and zero to one neurons in the second hidden layer. The combination of neurons, which yielded the smallest absolute deviation between predicted power generation and real power generation during training, was saved and later used for power predictions. Therefore, power predictions with different wind velocity sources will not always be done with the same neuron configurations. Instead the most optimal configuration for each set of input variables is utilised.

Thirdly, input and output norms can be set. They allow the model to permit input and output values outside the ranges of the training set. If the training data set only includes wind speeds up to $20\,\mathrm{m\,s^{-1}}$, but in live conditions wind speeds up to $25\,\mathrm{m\,s^{-1}}$ are expected, the input norm can be adjusted to 0.8. Similarly, the output norm can be adjusted to allow the ANN to forecast slightly more power than the training data included. However, this function has to be used with care for wind energy forecasts, since the wind speed to power conversion from WET is not linear (see Sec. 3.1). In this thesis, input norms of 0.1 and 0.9 are used. This allows the input variables to be up to $10\,\%$ lower or higher during operation than they were during training. The output norms are set to 0.2 and 0.8, allowing the predicted power output to be up to $20\,\%$ lower or higher during operation than they were during training. However, the effect of the output norm is minimal, since the output power predicted by the LoadManager® is hard-capped in both directions, by no wind power generation at $0\,\mathrm{MW}$ and the maximum capacity of the analysed wind farm at $44\,\mathrm{MW}$.

The final two parameters of ANN configuration are the momentum and the learning rate. They decide how much the ANN adjusts its weights after each internal calibration and how fast the model converges to a solution. The momentum parameter is in principle based on
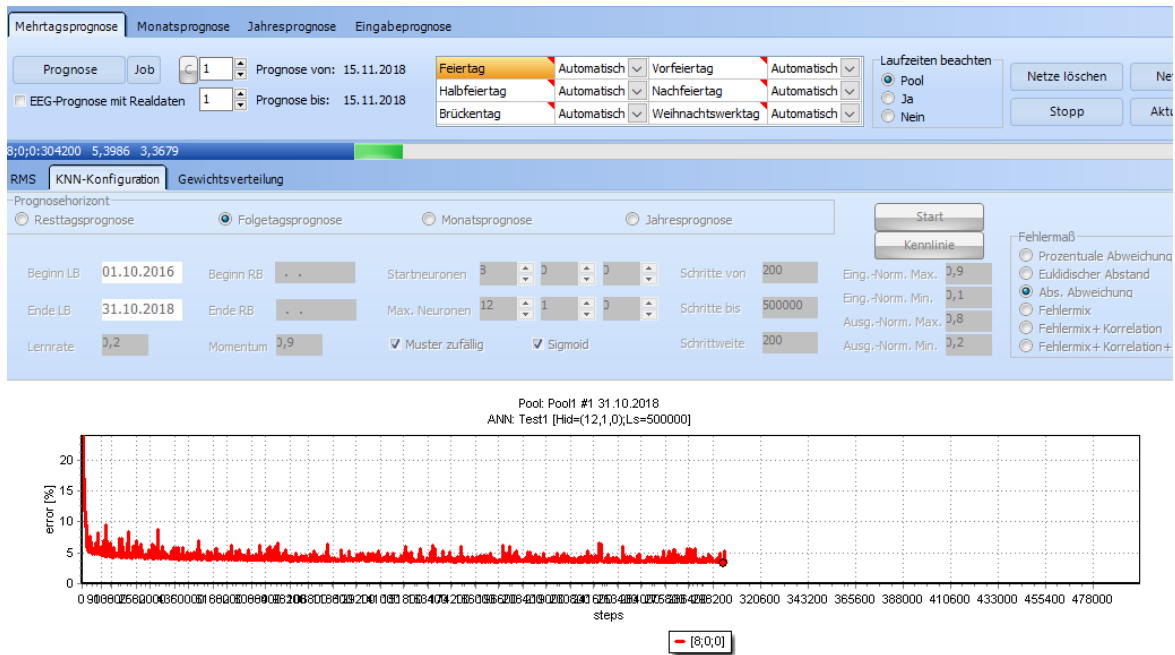
Figure 3.11: Screenshot of LoadManager® user interface by LEM-Software.

the physical *momentum* property. It replaces the gradient, which is used to update the weights at every training step, by a moving average of current and previous gradients (*Qian*, 1999). During the ANN training, the momentum is designed to help overshoot smaller local minima and find the more important global minima. However, if the momentum is too high it may lead to overshooting even global minima and thus destabilising the ANNs solution. A momentum value of 0.9 is typically used (e.g., *Bushaev*, 2017) and was also applied for this project's power predictions. The learning rate is another parameter controlling how much the ANN updates its weights during the training process and has a small positive value, usually between 0 and 1. When choosing a lower learning rate the training process needs more iterations to converge to a solution, as the changes made to the weights at each step are smaller. Larger learning rates speed up the training process, but may lead to the training converging to a suboptimal solution. The developers of the LoadManager® power prediction tool at LEM-Software suggested a learning rate of 0.2, which will be used for the power predictions in this thesis.

3.4.2 Integration of observational wind data into LoadManager® power predictions

There are multiple points in the power prediction process (see Fig. 3.9), where additional remote sensing data might be useful. The first and most obvious point is the NWP base data (version A in Fig. 3.9). Assimilating wind data into NWP models can increase the models reliability (e.g., *Illingworth et al.*, 2015; *Necker et al.*, 2018). However, in the frame of this thesis it is not possible to assimilate observations into the WPF model runs, since they are done externally by Met Office and LEM-Software is only using the final data that was interpolated for the location of the wind farm. Nevertheless, an advantage of the Met Office UM is that wind observation data from E-PROFILE is already assimilated (*Hooper et al.*, 2013). Therefore, the models wind data for the location of Lindenberg should arguably be very good.

Secondly, observational data can be used to correct the systematic bias in WPF model output (Fig. 3.9 version B, see also Sect. 4.2.2). A similar correction would then be used on NWP forecast data during operational power predictions. The ANN training data and wind forecast data would both be corrected and then be closer to the observation, thus providing more reliable power forecasts. However, the reason ANNs are used for power predictions is that they are able to learn systematic errors in the WPF input data during the training already.

Thirdly, observational data can be used directly as input for the LoadManager® power prediction (see Fig. 3.9 version C). Observed wind velocities are not subject to parametrisation like the model data. Additionally, they do not need to be interpolated between model grid points as they are measured at the exact location of the wind farm (nacelle sonic measurements) or at a distance of 6 km (DL and MT). The disadvantage of observational data is that it has no inherent prognostic value. Using the time-shifted observational wind speed in a persistence model (also called "naïve model") can, however, provide power forecasts for the next few time steps. Since changes to the submitted power generation schedule can be submitted until 5 min in advance (see Sect. 2.3), even improved power forecasts for only the next few time steps are valuable.

In this thesis, the third approach will be used. After introducing the study site and the observational data sets in detail in the following chapter, the input data sets for the power prediction based on observational data will be summarised in Sect. 4.4.

4 Methodology

In this chapter, the study site and the observational data sets are introduced in Sect. 4.1. Furthermore, the different available wind velocity data sets are compared to each other in Sect. 4.2. To provide a potential improvement over the persistence wind speed model, the Auto-Regression Integrated Moving Average (ARIMA) model is described and evaluated in its capability to predict short-term wind speed values in Sect. 4.3. Section 4.4 summarises the observational data sets used for wind power predictions. In the end of this Chapter, Sect. 4.5 introduces methods to detect sudden large changes in wind power ramp generation, so called ramp events. These methods are used on the present data set to evaluate the frequency of occurrence of ramp events at the location of the wind farm Beeskow. Main research questions for this chapter are:

- How well are the different wind velocity sources able to represent the wind field at the location of the wind farm?

- Does the ARIMA model improve the wind speed prediction compared to the persistence approach?

- What are the main issues for the current wind power prediction setup?

4.1 Site characteristics

4.1.1 Meteorological Observatory Lindenberg - Richard Aßmann-Observatory

Figure 4.1 shows the study region including the different measurement locations. The village Lindenberg is located about 40 km southwest of Berlin in the German federal state Brandenburg. The surrounding region is mostly characterised by patches of forests and fields and does not provide any complex orography. However, the Meteorological Observatory Lindenberg - Richard-Aßmann Observatory (MOLRAO) is located on a slight ridge in the landscape of approximately 110 m above sea level, which is extending 8 km to the north of the observatory. Its elevated position stems from the historic occupation of launching meteorological weather kites. Weather kites provided vertical profiles of the atmosphere in Lindenberg as early as 1905, when Richard Aßmann founded the observatory to systematically measure the vertical distribution of wind, temperature and humidity.

Today, the MOLRAO supersite comprises extensive measurements including boundary layer profiling, atmospheric remote sensing and the DWD radiation facility. It also is one of the four German Wind-Profiling Radar (WPR) locations. Although the lowest measurement range gate of the WPR at 500 m height is above the typical WET height, the assimilation of its observations into NWP models improves the WPF model's performance at this location (see Sect. 3.3). Additionally, the atmospheric remote sensing suite of MOLRAO includes the Doppler Lidar (DL) introduced in Sect. 3.2.2, a 35 GHz Doppler cloud radar, multiple microwave radiometers, the Raman lidar RAMSES and a ceilometer. MOLRAO is also part of the European CloudNet network, a project created for structured measurements of clouds with the goal to improve their representation in weather and climate models (*Illingworth*
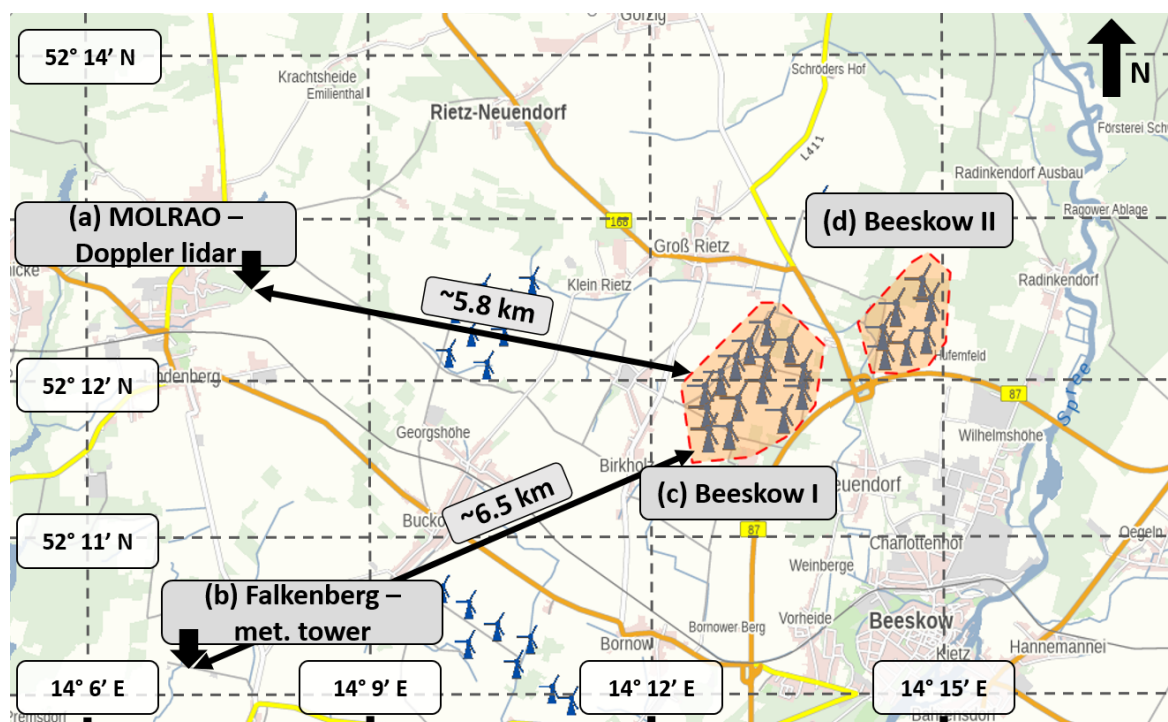
Figure 4.1: Map of the study site: (a) marks the MOLRAO observatory of DWD where the Doppler lidar measurements are recorded (52°12'35"N, 14°7'43"E, 104m a.s.l.), (b) marks the location of the meteorological tower at the DWD atmospheric boundary layer research site (52°10'N ,14°07'20"E, 74 m a.s.l.), (c) indicates the location of the first part of the investigated wind farm Beeskow I (52°11'48"N, 14°13'E, 55-65m a.s.l.), (d) marks the location of Beeskow II (52°12'10"N, 14°14.5'E, 50m a.s.l.). Figure adapted from Energie- und Klimaschutzatlas Brandenburg, provided by the Department of Economics, Labour and Energy.

et al., 2007). CloudNet itself is part of ACTRIS, the European Aerosol, Clouds and Trace gases Research InfraStructure, aimed at further innovating measurement techniques and observations methods of short-lived atmospheric components.

Complementing the observatory, the DWD boundary layer field site Falkenberg is situated 5 km south of MOLRAO. It provides a 99 m high meteorological measurement tower with sonic anemometers, among other instruments, and one more DL. While the DL in Falkenberg is not always operated in a way to measure the horizontal wind field, a limited data set of wind velocity observations from the meteorological tower will also be compared to the DL measurements at MOLRAO, the ultrasonic anemometer measurements at the back of the nacelles and the WPF model performance in Sect. 4.2.

4.1.2 Wind farm Beeskow

The wind farm consists of 22 wind energy turbines of type Vestas V80-2.0MW (described in Sect. 3.1 and is split into two parts, Beeskow Neuendorf (Beeskow I, 52°11'48"N, 14°13'E) and Beeskow Hufenfeld (Beeskow II, 52°12'10"N, 14°14.5'E). However, the power data are only available at the feeding point, where the generated power from the whole wind farm is fed into the electric grid. Therefore, no separation between single WETs or even parts of the wind farm can be made. Beeskow I and II are located about 6 km and 7.5 km east of MOLRAO, respectively. Since MOLRAO is located on higher terrain, the height difference between the observatory and the wind farm ranges approximately between 40 m and 55 m.
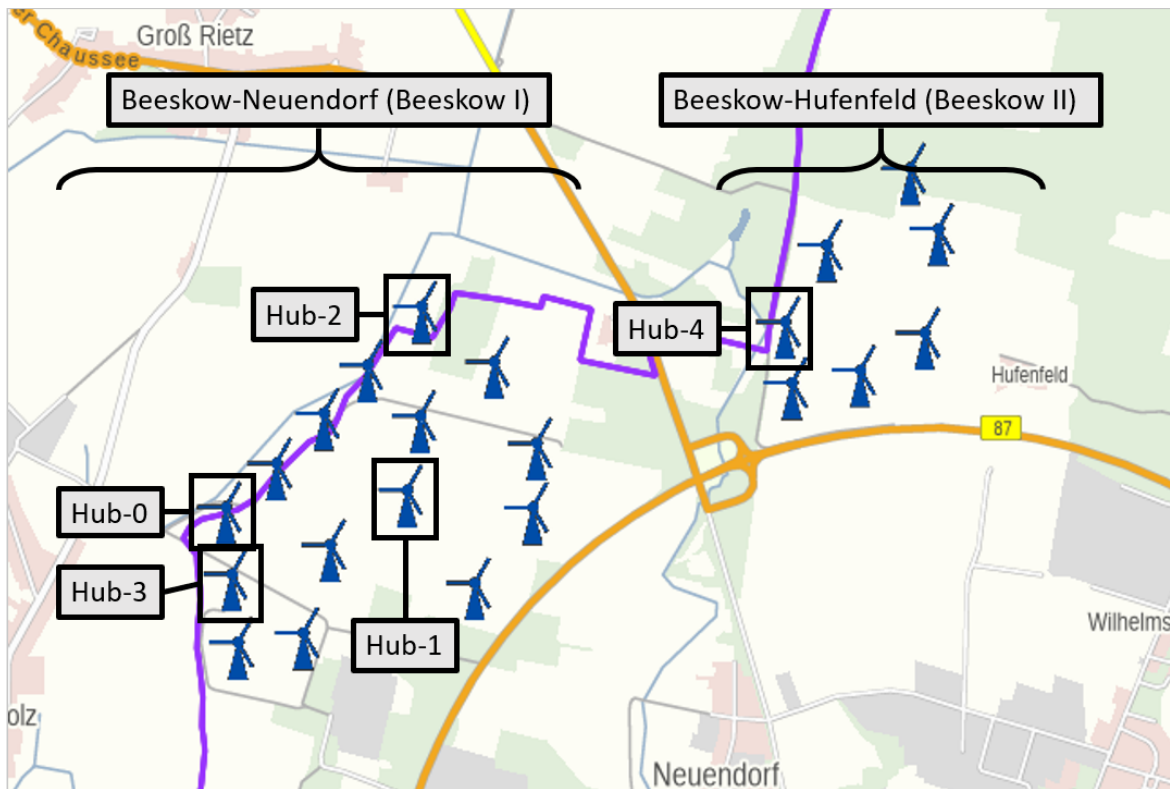
Figure 4.2: Map of the wind energy turbines at wind farm Beeskow. Nacelle wind velocity data are available for the labeled wind turbines. Figure adapted from Energie- und Klimaschutzatlas Brandenburg, provided by the Department of Economics, Labour and Energy

Technical data to Vestas V80-2.0MW turbines are summarised in Tab. 3.1 and their power curve is shown in Fig. 3.2.

Time series of wind speed and wind direction for five of the 22 wind turbines, four from Beeskow I and the fifth from Beeskow II (see Fig. 4.2), are available from 1 January 2014 to 30 September 2019. Sonic anemometers are mounted at the back side on top of every nacelle to measure wind velocities. The measurements are averaged over 10 min time periods. However, only the horizontal wind speed is used for further evaluation, since the wind direction is not available for more than 70 % of the time steps.

4.1.3 Wind farm internal wake effects

Wake effects describe effects of the wind farm on the downwind wind field. Most prominently these include a decrease of downwind wind speed and an increase in turbulence intensity, reducing the power output of downwind turbines considerably. Wake effects can be minimised by planning the layout of the wind farm according to the most prominent wind directions they are exposed to (*González-Longatt et al.*, 2012). The extent of wake effects depends on the atmospheric condition. Generally, for stably stratified conditions wake effects extend further, while a turbulent atmosphere compensates for them more quickly. In stable conditions, wake effects can extend up to 50 km downwind (*Lundquist et al.*, 2019). In this section, wind speed wake effects within the wind farm are briefly evaluated. Within the relatively small confines of the wind farm (longest dimension 3 km), wind energy turbines at the downwind

end experience a wind field influenced by the rest of the wind farm, independent of the atmospheric condition.

For the present data set, wind speed data are available for the five wind energy turbines marked in Fig. 4.2. To evaluate the nature of wake effects, wind speed measurements from the nacelles of these turbines are compared to each other, depending on the wind direction. Since the wind direction observations are severely limited for the nacelle sonic measurements, wind direction information of the meteorological mast in Falkenberg is used. However, wind direction observations on the mast in the narrow wind direction sector between 60° and 85°are affected by flow of the wind around the mast itself, causing artefacts in this sector. Therefore, this wind direction range is disregarded in the following analysis and marked grey in the plots of this section.

First, a histogram of the wind speed difference between sonic anemometer measurements on two wind energy turbines is shown in Fig. 4.3a. The radial sectors represent the wind direction divided into 9°-bins. The radius represents the wind speed difference between both locations in bins of $0.5\,\mathrm{m\,s^{-1}}$. Wind speed difference values between $\pm 0.1\,\mathrm{m\,s^{-1}}$ are disregarded and marked grey, because this is in the typical uncertainty range of sonic anemometers (see Sect. 3.2.3). In addition, wind speed values lower than the cut-in speed of $4\,\mathrm{m\,s^{-1}}$ have been omitted. The color of the bins shows the absolute count per bin divided by the number of data points in each wind direction sector to account for the difference in frequency of wind directions.

Figure 4.3a shows the difference between the turbines marked as Hub-0 and Hub-1 in Fig. 4.2. Outside the grey ring are data points where Hub-0 observed the higher wind speed values. Inside the grey ring are data points where Hub-1 observed the higher wind speed values. Figure **??**b shows a map of the Western part of the wind farm, Beeskow I. Hub-0 is located at the Western edge of the wind farm and Hub-1 is located at a distance of $700\,\mathrm{m}$ in the center of the wind farm.

To emphasize the nature of wake effects, three areas are highlighted on both sides of the figure. Firstly, the orange area shows Westerly winds. For Westerly winds, Hub-0 is the first turbine of the wind farm and, therefore, affected by the undisturbed wind. Although there is no other wind turbine directly in front of it, Hub-1 in the center of the wind farm presumably experiences a decelerated stream. This assumption is evidenced by the histogram, showing that for Westerly flows the wind speed at Hub-0 is often higher than at Hub-1, most frequently by 0.5-$1.0\,\mathrm{m\,s^{-1}}$. Secondly, for South-Westerly winds highlighted by the yellow area, the wind speed difference (Fig. 4.3a) is even higher than for the orange area. Hub-0 is again affected by the mostly undisturbed flow. However, for this wind direction sector, another wind turbine is in front of Hub-1 at close distance (see Fig. 4.3b). Therefore, there is a smaller distance to compensate for upwind wake effects, causing an even larger difference toward higher wind speed values observed at Hub-0. Thirdly, the red area shows a Southern sector, which indicates on average higher wind speed values for the central turbine Hub-1. As seen on the map, for Southerly winds Hub-0 is downwind of two other closely located wind energy turbines, causing a decrease of the wind speed before reaching Hub-0. Contrarily, in this wind direction sector Hub-1 is not affected by other turbines upwind and, thus, experiences on average higher wind speed values than Hub-0.

Comparisons between wind speed observations at the locations of any two wind energy turbines show similar effects, depending on their location relative to each other. No single nacelle anemometer can provide wind speed observations representative of the undisturbed wind for all wind directions. Therefore, the mean wind speed measured between the five available locations, which presumably represents most of the wind farm, will be used for further analyses. To compare the observations to a representative data set, which is not
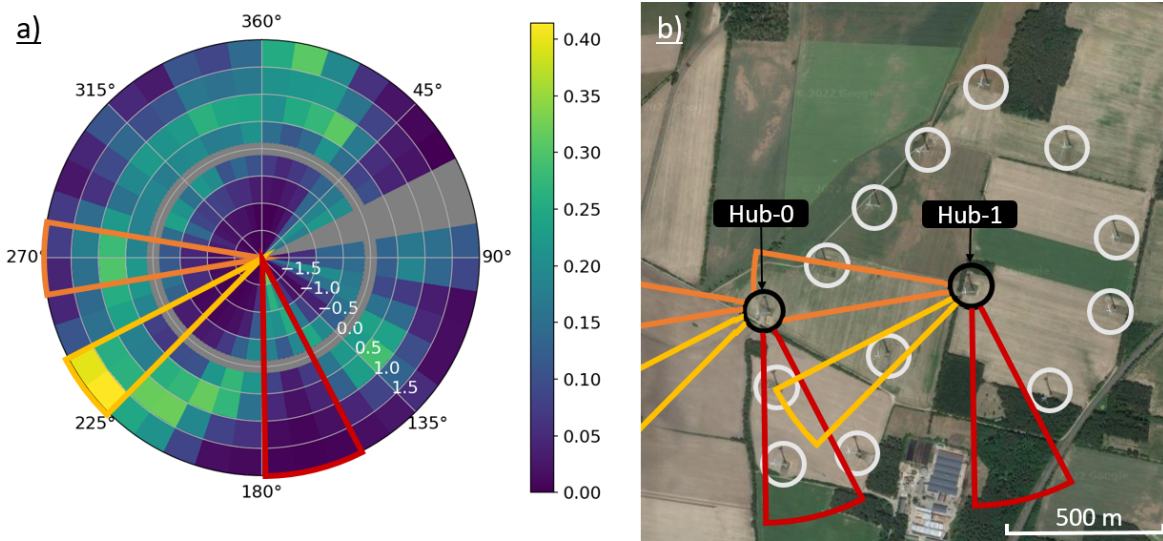
Figure 4.3: Illustration of wake effects between nacelle anemometer wind speed measurements on two wind energy turbines in the wind farm Beeskow. The colored sectors are highlighted for the comparison in Sect. 4.1.3. a) Frequency of occurrence of wind speed difference (radial axis, in $\mathrm{m\,s^{-1}}$) for 40 equally-sized wind direction sectors (angular axis). Each wind direction sector has been normalised by the number of data points in this sector and, therefore, the bins add up to 1. Areas marked as grey are omitted from the comparison. b) Satellite image of the wind farm. Black circles mark the two wind energy turbines being compared, white circles mark the remaining wind energy turbines of Beeskow I. Satellite image from Google Maps.

immediately affected by the wind farm, the wind speed observed at the meteorological tower Falkenberg is used, under the assumption that the wind field at both locations is comparable. Figure 4.4a shows the wind speed difference between the mean of the five available nacelle anemometers and the MT. In addition to the wind speed difference bins, the red dotted line represents the mean wind speed difference per wind direction sector. Ideally, the mean wind speed difference would be zero for all wind direction bins, i.e., the measurements of the meteorological tower agree perfectly with the mean of the nacelle observations. For most Westerly wind directions, ranging from 180° up to 330°, the red line is near $0\,\mathrm{m\,s^{-1}}$, albeit slightly negative. That means while the wind speed at both locations agrees well, the mean wind speed of the wind farm is usually lower than in the undisturbed flow. This is caused by some of the nacelles being downwind of others and, therefore, observing a decreased wind speed, which slightly reduces the mean value as described above. However, for these Westerly wind directions, the mean value of the sonic anemometers represents the undisturbed wind speed measured at the MT well.

Contrarily, the wind speed difference in Fig. 4.4a for Easterly winds is below -1.0 $\mathrm{m\,s^{-1}}$ for wind directions between 30° and 170°, significantly deviating from the ideal value of $0\,\mathrm{m\,s^{-1}}$. For these wind directions, the mean value of the nacelle wind speed observations is severely below the wind speed measured at the location of the meteorological tower. Thus, the mean value of the sonic anemometers does not represent the undisturbed flow well for most Easterly winds.

This bisection is caused by the asymmetrical spacing of the wind energy turbines, for which wind speed data are available (see Fig. 4.2). Ideally, wind speed observations from turbines on all edges of the wind farm would be available to ensure that for each wind direction at least one anemometer measures the (mostly) undisturbed wind. However, within the available data set,

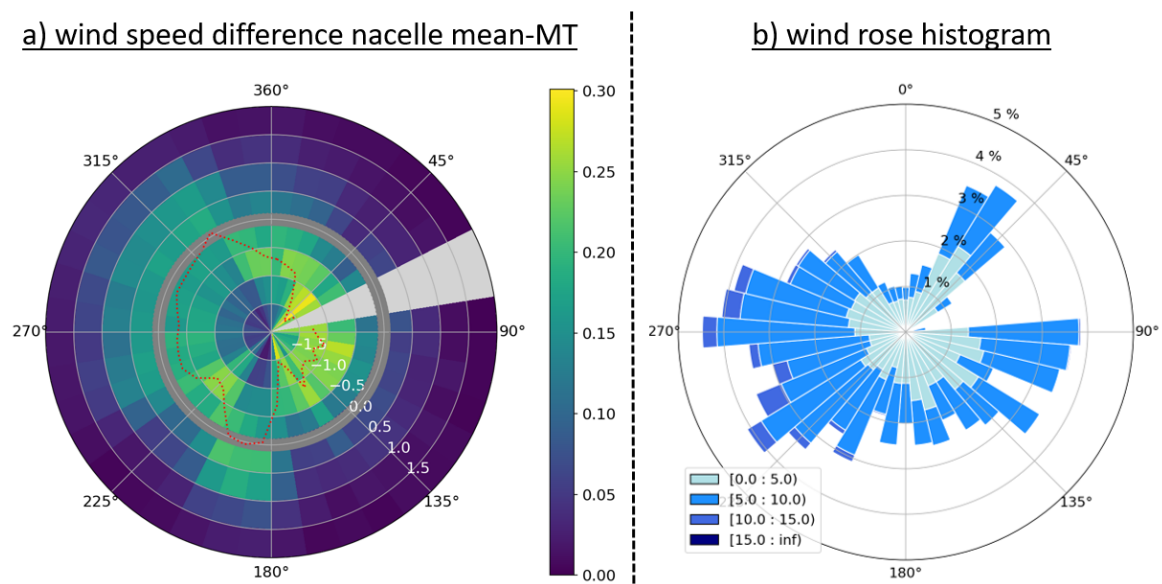## a) wind speed difference nacelle mean-MT

b) wind rose histogram

Figure 4.4: a) Same analysis as in Fig. 4.3, but for the wind speed difference between the mean wind speed of the five available nacelle sonic measurements at the wind energy turbines and the measurements at the meteorological tower (MT). The red dotted line represents the mean wind speed difference for each wind direction sector. b) Wind rose histogram for the same wind direction bins as left. The shade of the blue bars represents wind speed bins in $\mathrm{m\,s^{-1}}$, their length represents the frequency of occurrence.

the Southeastern and Eastern edges of the wind farm are not covered. Instead, the Western edge is represented at multiple locations; Hub-0, Hub-2 and Hub-3 at the westernmost edge of Beeskow I and Hub-4 at the Western edge of Beeskow II, albeit the latter is downwind of Beeskow I for Westerly winds. Therefore, Westerly winds are evidently represented well by the nacelle mean value. On the contrary, the Eastern edge of the wind farm is not covered at all. Only Hub-4 is located in Beeskow II, but not on the easternmost edge. Thus, Easterly winds are underrepresented by the nacelle mean. Even if observations from certain locations are excluded from the mean (not shown), the agreement with the meteorological tower for Easterly winds remains weak.

Figure 4.4b shows the wind rose histogram for the mean nacelle wind speed with MT wind direction data. Westerly and Southwesterly winds are more frequent and more frequently feature higher absolute wind speed values. As explained in Sect. 3.1, wind power generation considerably increases at higher wind speed values, thus, increasing the relevance of a good representation of Westerly winds. Since Easterly winds are slightly less frequent and on average slower, the worse representation of Easterly winds is acceptable in the scope of this thesis. For the future operational application of this approach, the responsible wind farm operators presumably have nacelle wind speed data of all their wind energy turbines available, eliminating this limitation.

4.2 Comparison of wind speed sources

Different wind velocity data sets are used in this thesis, which are summarized in Tab. 4.1. Their availability throughout the investigated period is illustrated in Fig. 4.5, where each cell represents three month of data.
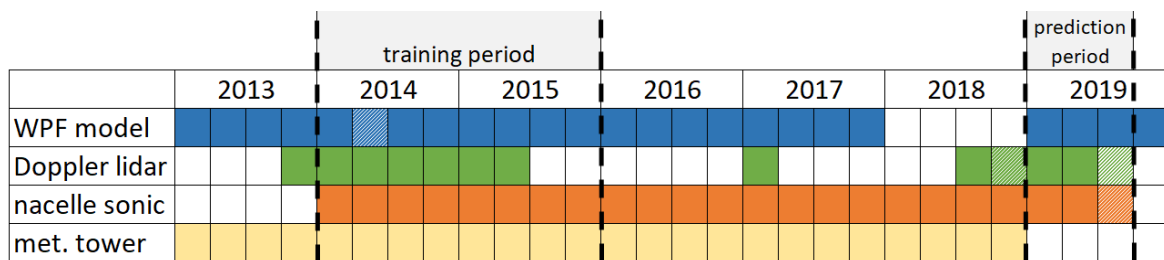
Figure 4.5: Periods of available wind velocity data for the WPF model and the instruments evaluated in this Chapter. Colored cells represent times with data fully available, hatched cells represent periods with data gaps and white cells mean no data are available. Highlighted are the training period used for the ANN training of the different wind power prediction configurations (see also Sect. 4.4) and the prediction period, used to evaluate differences in performance of these wind power predictions in Chapter 5.

Table 4.1: Comparison of temporal resolution Δt, height above ground z_{ag} and height above sea level z_{asl} of the different wind velocity sources. Heights for observations from the meteorological tower and Doppler lidar include the two closest levels to the hub height above ground (100 m)

data source	Δt [min]	z_{ag} [m]	z_{asl} [m]
WPF model	60	70	118
Doppler lidar	30	69 / 115	173 / 219
nacelle sonic anemometers	10	100	150-165
met. tower sonic anemometers	10	80 / 99	154 / 173

To compare the varying wind velocity sources, their different temporal resolutions and observation heights have to be taken into account, as listed in Tab. 4.1. The WPF product provides NWP model data. The model wind velocity is available at 10 m, 70 m and 355 m above the model ground level of 50 m on a temporal resolution Δt of 1 h for the time period from 1 January 2014 to 31 December 2017 and 1 January 2019 to 31 December 2019. The Doppler lidar measurement levels closest to hub height are at 69 m and 115 m above the ground, putting them at 173 m and 219 m above sea level, respectively. Doppler lidar data are available on a temporal resolution Δt=30 min from 1 January 2014 to 30 May 2015 and 1 January 2019 to 4 July 2019. Nacelle sonic anemometers measure at the hub height of 100 m and their data are averaged over Δt=10 min. Sonic measurements are available from 1 January 2014 to 22 September 2019. The meteorological mast's highest measurement levels are located at 80 m and 99 m above the ground and the data are averaged over 10 min for the time period of 1 January 2014 to 31 December 2018.

Each data set has distinct advantages and disadvantages. Firstly, the DL data set is able to provide reliable 3D wind information in a vertical profile above the ground with a temporal resolution of Δt=30 min. However, the DLs measurement is not exactly at hub height and the instrument is not located directly at the wind farm (see Sect. 4.1.2). Secondly, the nacelle sonic anemometers provide measurements with a high temporal resolution (Δt=10 min) directly at the location of the wind farm and exactly at hub height. However, they are likely influenced by the wake of the rotor blades (see Sect. 3.2.3), do not provide a vertical profile and lack wind direction data for long periods of time. Lastly, the meteorological tower (Δt=10 min) observes the wind velocity at approximately the hub height and can provide a vertical profile, albeit only between ground and hub height. The location of the tower is about the same distance from the wind farm as the DL. While the model data are calculated
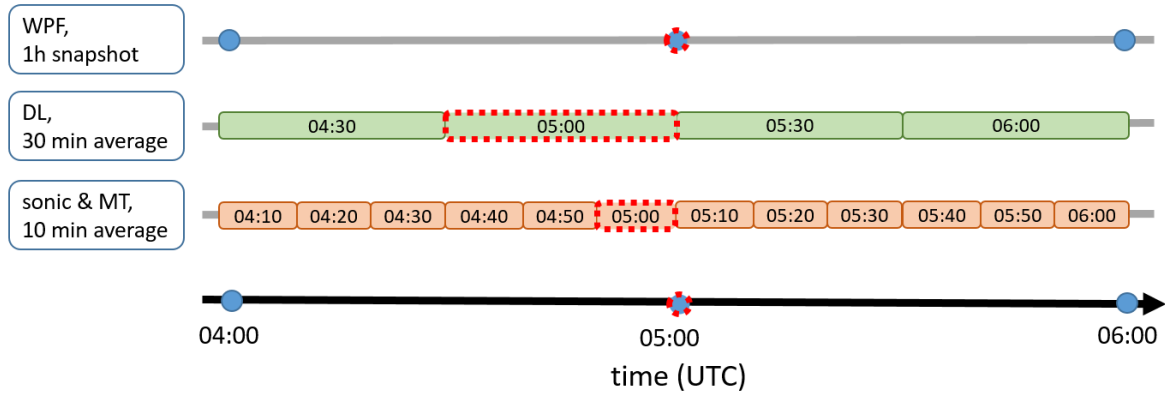
Figure 4.6: Illustration of the difference in temporal resolution Δt between the different wind velocity data sets. This difference may influence the comparison of wind velocity in Sect. 4.2.2.

for a snapshot moment, the measurements are average values over their respective time period. Therefore, the most recent time step of the observational data set is used in comparison with the WPF time step. Figure 4.6 illustrates an example for the wind data comparison at the time step 05:00 UTC. The WPF model output is a snapshot for each hour, short-lived wind speed changes in between the time steps are not represented. In contrast, the instrument's temporal averages are influenced by short-lived wind speed changes in between their time steps, as long as the instrument's observation frequency is sufficiently high. The colored rounded boxes in Fig. 4.6 represent the temporal averages of the observational data sets. For example, the green box (Δt=30 min) labeled "05:00" represents the temporal average of the measured wind velocity of the DL from 04:30 to 05:00 UTC. The orange box (Δt=10 min) labeled "05:00" represents the temporal average of the measured wind velocity from sonic anemometer or MT from 04:50 to 05:00 UTC. For the example time step 05:00 UTC, the indicated (red dotted border in Fig. 4.6) temporal average windows are used to compare them to the WPF snapshot time step. This figure is meant to illustrate the difference in wind data set composition. Though the data sets cannot be perfectly aligned in this thesis, the shown temporal differences are underlying influence factors for the following intercomparisons.

4.2.1 Comparison of observational data sets

This sections investigates differences in performance between DL, nacelle sonic anemometers and MT measurements in the study region. In general, good agreements between the instruments are expected (e.g., *Kelley et al.*, 2007; *Lane et al.*, 2013). *Smith et al.* (2006) compared wind data of a DL instrument to measurements of cup- and sonic anemometers on a MT. The comparison of wind speed values of the DL with MT anemometers at 100 m above the ground yielded a correlation with R^2=0.99. Furthermore, *Päschke et al.* (2015) compared DL data to measurements of radiosondes and a WPR during a one year period at MOLRAO. They also concluded very good agreements between the DL and the other systems (root mean square errors between $0.7\,\mathrm{m\,s^{-1}}$ and $0.9\,\mathrm{m\,s^{-1}}$). Differences between DL and WPR were mainly attributed to their slightly different fields of view. However, since the elevation angle of the WPR (Θ=74.8°) is similar to the Streamline DL (Θ=75°), the difference in measurement volumes is mainly caused by the coarser vertical resolution of the WPR. Most of these comparisons have been done for instruments at the same location. Still, since the orography in the study region is not complex, only minor differences between them are expected.
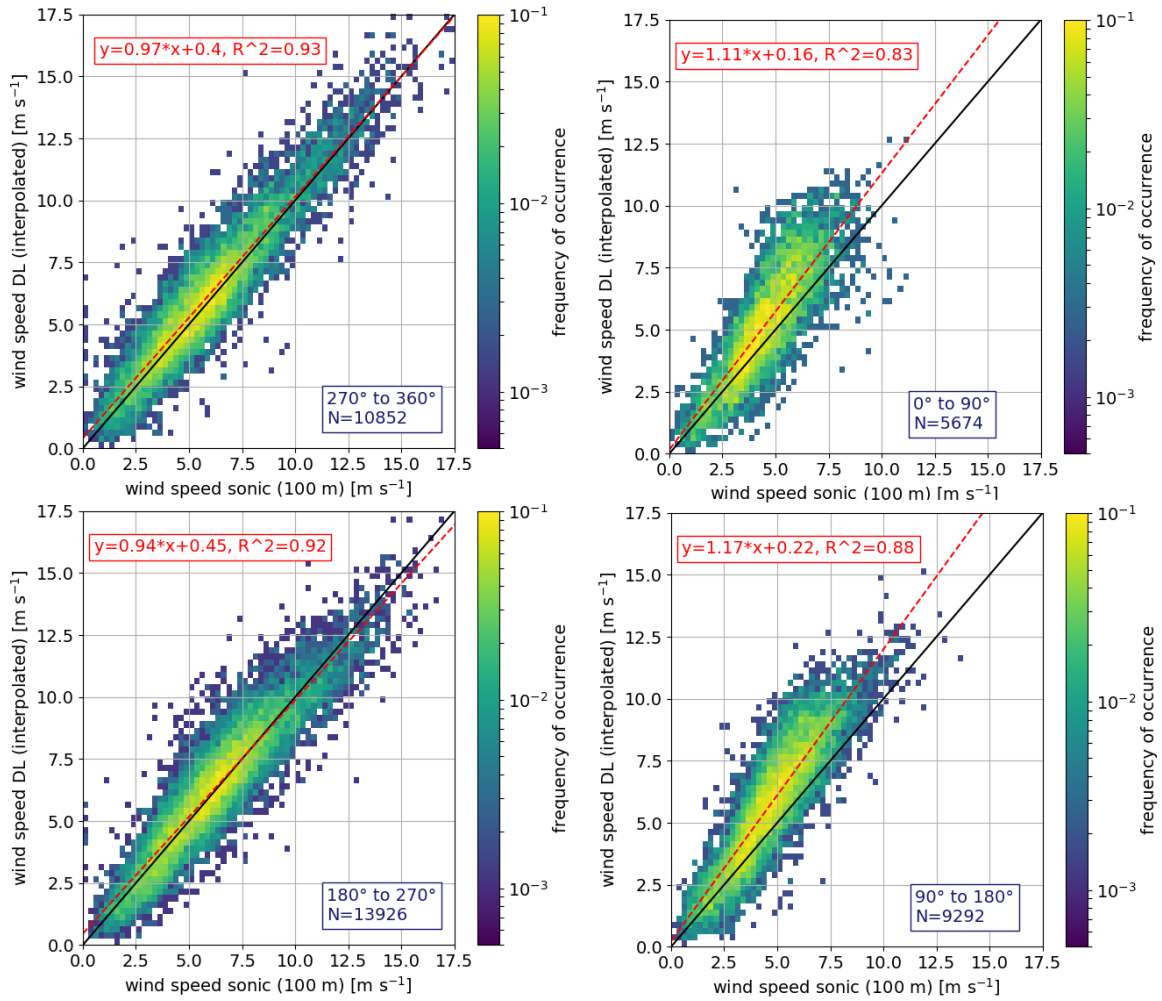
Figure 4.7: Wind speed comparison between nacelle sonic anemometer measurements and Doppler lidar for four wind direction sectors (described in the blue box). Two-dimensional (2D) histograms of both wind speed sources are shown. The solid black line marks the 1:1 line and the dashed red line represents a linear regression with its parameters displayed in the red box. The amount of data points N in this comparison is shown in the blue box. The bin width of the histograms is $0.25 \, \mathrm{m \, s^{-1}}$.

Figure 4.7 shows the comparison between nacelle sonic anemometer and DL observations for four wind direction sectors as two-dimensional (2D) histograms of both wind speed sources, colored by the frequency of occurrence of each bin. The bin width is $0.25\,\mathrm{m\,s^{-1}}$. The solid black line marks the 1:1 line and the dashed red line represents a linear regression with its parameters displayed in the red box. If the relationship between the values of the compared wind speed data sets is a function of the wind speed, the slope deviates from one. A bias between the data sets would be indicated by a non-zero offset. The correlation coefficient R^2 is a measure of the spread of the data cloud around the regression line. The higher R^2, the lower the random error in the correlation. The wind direction sectors are defined as 90° intervals between the four cardinal points and will in the following be described by their edges (e.g., N-E referring to North-to-East: 0°-90°). Over two years of simultaneous measurements have been evaluated. Doppler lidar data were averaged over 30 min, nacelle sonic data were averaged over 10 min (see also Fig. 4.6). While some differences between the instruments caused by the different averaging times, especially for periods with changing wind velocity, are expected, the agreement between them is still very acceptable.

The right panels of Fig. 4.7, the histograms for both N-E and E-S, show a clear deviation from the 1:1 line and a lower wind speed in general, usually below $10\,\mathrm{m\,s^{-1}}$. In addition, the wind speed difference between both instruments increases for higher wind speed values. For N-E the amount of samples N is smallest and the correlation coefficient R^2=0.83 is the lowest, representing the larger spread. On the left panels, for the Western sectors S-W and W-N, R^2 is much higher indicating a smaller spread around the linear regression fit. Measurements for the Western sectors agree better with each other and do not exhibit an increased wind speed difference for higher wind speed values. Additionally, Westerly winds are more frequent and wind speed values range up to $15\,\mathrm{m\,s^{-1}}$.

Figure 4.8 shows the comparison between nacelle sonic anemometer and MT observations for the same four wind sectors for almost three years of 10 min averaged data. Although the time period of three years is similar to the DL comparison, this amounts to approximately three times as many data pairs due to the equal averaging times. For all sectors, the correlation coefficient is $R^2 \geq$ 0.9, indicating a very good agreement between observations at both measurement locations. For the sectors N-E and E-S, the wind speed difference increases for higher wind speed values. The best correlation R^2=0.93 occurs for S-W, which is the most frequent wind direction. South-Westerly winds stream first across the MT before reaching the wind farm.

The reason for the increased wind speed difference at higher wind speeds in N-E and E-S are likely wake effects of the wind farm, as explained in Sect. 4.1.3. The available nacelle sonic measurements are predominately from the Southwestern end of the wind farm (see Fig. 4.2). For Easterly winds, most of the nacelle sonic anemometers measure the wind speed only after it crossed the majority of the wind energy turbines already. Therefore, the wind speed is lower than it would have been in an undisturbed flow upwind of the wind farm. This underestimation of the wind speed increases for higher wind speed values, because the amount of wind energy transformed to electrical power increases. However, since Eastern wind directions are less frequent and usually characterised by lower wind speed values, this downside of the nacelle sonic measurements is expected to be less influential on the wind power predictions. Therefore, the nacelle sonic data set remains a promising input data set for the later wind power predictions.

Even though the DL wind speed at hub height correlates slightly less with the sonic observations than the MT, the advantage of the DL data set for the wind power predictions is that it does not only include the hub height wind speed. In addition, it will feature the wind velocity in three heights, covering also layers directly above the rotor, and the DL data
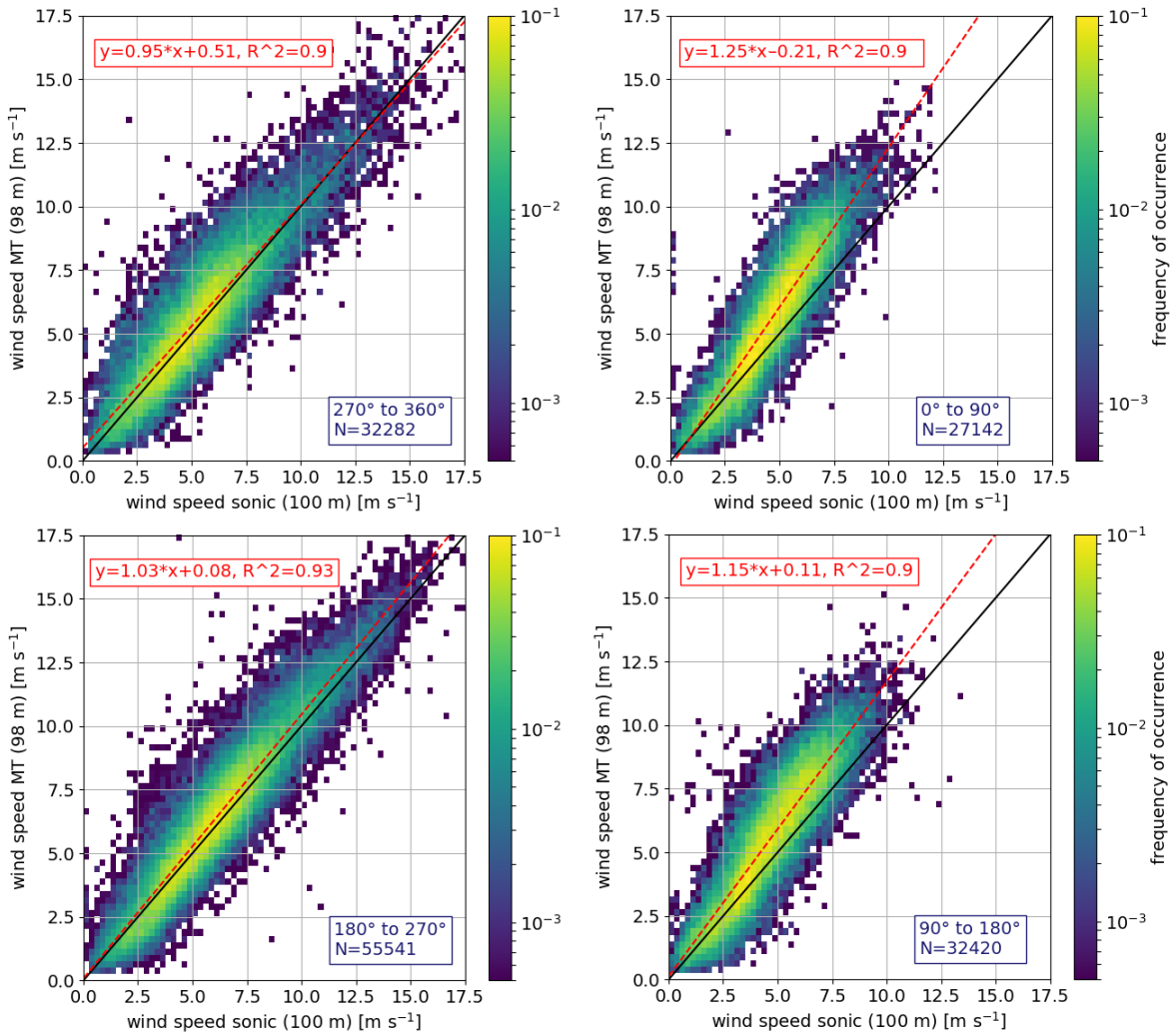
Figure 4.8: Same analysis as Fig. 4.7, but for the comparison between nacelle sonic anemometers and meteorological tower anemometers.

set also includes wind direction values. This additional wind velocity information may compensate the slightly higher wind speed difference in the power prediction. Therefore, the DL data set continues to be promising for wind power predictions done subsequently as well.

Apart from specific detriments explained above, the agreement between nacelle sonic, DL and MT data is very good. Therefore, it can be concluded that

(i) the nacelle sonic anemometer data set is able to represent the wind field at the wind farm, apart from the more rare Easterly wind directions (that usually have lower wind speed values).

(ii) the distance of approximately 6 km between the wind farm and the measurement locations of the the DL and MT does not hinder very good agreements between the observations.

(iii) wind farm wake effects are visible in these comparison plots. However, in an operational setting it is assumed that the wind farm operator has access to real-time observations of *all* of their wind energy turbines, nullifying (or, at least, reducing) that effect.

4.2.2 Comparison of observational data and WPF product

The wind speed observations by the nacelle sonic anemometers, Doppler lidar and the meteorological tower anemometers are compared to the WPF model data. Figure 4.9 shows a comparison between the WPF model wind data and the observational wind speed data sets. On the left side are histograms of the wind speed difference between measured and modeled wind speed. The bins have a width of $0.25\,\mathrm{m\,s^{-1}}$ and are centered around $0\,\mathrm{m\,s^{-1}}$ (vertical black line). The vertical red line marks the bin with the maximum probability density, additionally noted in the red box. Positive wind speed differences means the model wind speed was higher than the observed wind speed. On the right side, the previously introduced two-dimensional histogram plots of both wind speed sources including the linear regression parameters are displayed. Systematical errors as characterised by the slope and offset of the linear regression are not critical for power predictions with artificial neural networks, as the ANN learns and, therefore, neutralises these systematical errors. However, for shutdown criteria based on the absolute wind speed as, e.g., for the protection of bats, a systematical bias matters.

The nacelle wind speed with sonic anemometers is observed at $100\,\mathrm{m}$ height, closely above the nacelle, while the WPF model wind speed is interpolated for $70\,\mathrm{m}$. In the wind speed difference histogram (Fig. 4.9a) a systematic difference between the WPF model and the sonic is found. The modeled wind speed is on average $1.25\,\mathrm{m\,s^{-1}}$ higher than the observed wind speed. The height difference between both data sources does not account for this, as the on average higher wind speed would be expected at larger height (see Sect. 2.2). Furthermore, the roughness parametrisation effect also cannot account for this, as it rather leads to a slow bias. Since this parametrisation effect is corrected by the WPF post-processing (see Sect. 3.3), it is possible that it was overcorrected, leading to a fast bias. Lastly, the influence of the rotors on the wind speed observation might cause this bias. However, the comparisons between the model and the other observational data sets (DL: Fig. 4.9c and MT: Fig. 4.9e) show approximately the same bias, which most likely originates from the model. The 2D histogram (Fig. 4.9b) of wind speed observations from the sonic and the WPF model shows the linear regression and its parameters. Despite the regressions deviation from the 1:1 line, the correlation between wind speed observed by the nacelle sonic anemometers and the WPF model is in good agreement with R^2 values of 0.86. In addition, the 2D histogram shows that the wind speed difference is a function of the wind speed with a slope smaller than 1. Thus, the higher the wind speed, the higher the difference between model and observation is. The slope of this function depends on the wind direction. For Easterly wind directions, which occur less frequently, the slope is smaller than for Westerly directions (see Appendix Sect. A.1). This is caused by wake effects within the wind farm and the asymmetrical spacing of the available nacelle sonic anemometers (see Sect. 4.1.3).

Figures 4.9c and 4.9d show the comparison of WPF model and DL observations. The DL observations are, contrary to the nacelle sonic anemometers, not directly influenced by the wind energy turbines. However, the distance and orography between MOLRAO and the wind farm can affect the wind profile between both locations. For the DL, the nearest range gate to the $70\,\mathrm{m}$ model height is the second range gate above the ground at $69.5\,\mathrm{m}$. Signals from the DL range gates near the ground typically have a worse signal-to-noise ratio associated with them. Noise that originated in the instrument is less disruptive the farther away the DL's return signal comes from. Therefore, the first range gate signal is usually disregarded. Since in this data set, data with too high uncertainty very rarely extends to the second range gate at $69.5\,\mathrm{m}$, data from this height is used in this thesis. For a six month period, the DL measured with another configuration with higher vertical resolution, which shifted the
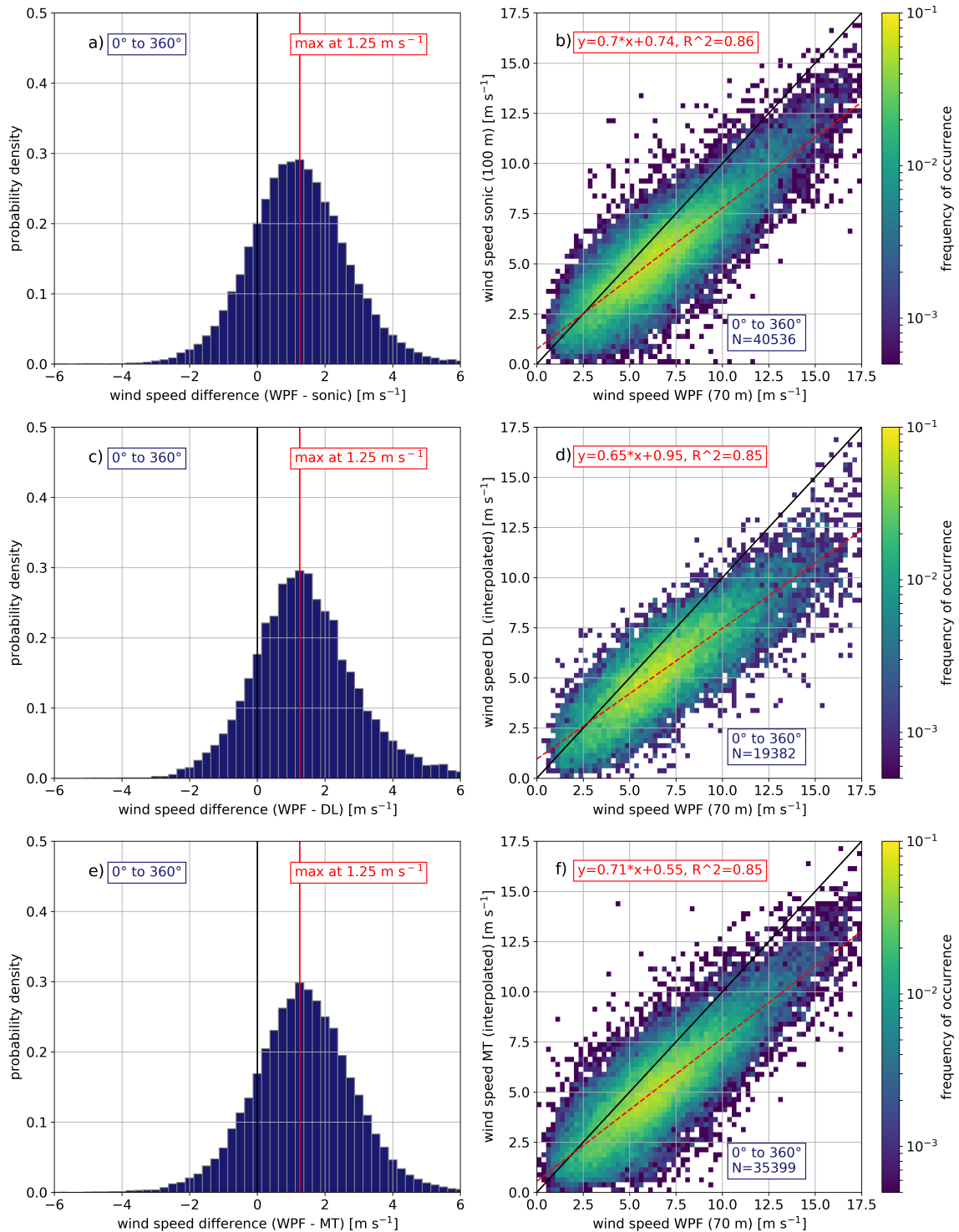
Figure 4.9: Wind speed comparison between observational data and WPF model data for all wind directions (blue box); first row: nacelle sonic anemometer, second row: Doppler lidar, third row: meteorological tower. Left side: Histogram of wind speed differences. The bin with the maximum probability is indicated in red. Right side: two-dimensional (2D) histogram of both wind speed data sources. The solid black line marks the 1:1 line and the dashed red line represents a linear regression with its parameters displayed in the red box. The amount N of data pairs in this comparison is shown in the blue box. The bin width for the 1D and 2D histograms is $0.25 \, \mathrm{m \, s^{-1}}$.

resulting range gates. This period is included in this comparison, but the DL wind speed is linearly interpolated between 60.9 m and 78.2 m, the DL range gates closest to the model height.

In most aspects, the DL-WPF comparison histogram (Fig. 4.9c) is similar to the comparison of WPF data with nacelle sonic anemometer data. A fast bias of the WPF model is shown in the wind speed difference plot with a maximum of the probability density at $1.25\,\mathrm{m\,s^{-1}}$ as for the WPF-nacelle sonic comparison. In the 2D histogram (Fig. 4.9d) the slope of the linear regression between WPF and DL is slightly lower than for the nacelle sonic comparison, but the correlation coefficient R^2 is similar.

The comparisons between the WPF model and MT data in Fig. 4.9e and Fig. 4.9f show similar patterns as the comparisons of WPF to sonic and DL data. The average wind speed difference is also $1.25\,\mathrm{m\,s^{-1}}$ compared to the model. As for the previous instruments, the mean wind speed difference increases for higher wind speeds. The MT data set is used to enable additional comparisons to evaluate the performance of the nacelle sonic anemometers and the DL. However, it will not be used for wind power predictions later on, since it does not fulfil the idea of mobility to be deployed at wind farms of user choice that is the background for this project (see Sect. 3.2.1).

Concluding the comparison with WPF, all observational data sets show a similar deviation from the NWP model wind speed. This suggests a constant bias of the model towards a higher wind speed on average resulting in an increasing difference for increasing wind speed. The data set of sonic anemometer measurements is the data set recorded closest to the wind farm and at exactly hub height of the wind energy turbines. As discussed in Sect. 3.2.3, the observations by the sonic anemometers at the back of the nacelle are influenced by the turbine itself. However, there is no obvious difference between the nacelle sonic comparison and the other two comparisons with DL and MT. Therefore, to conclude, it is expected that all observational data sets represent the real wind speed at the wind farm well.

4.2.3 Comparison of two different Doppler lidars

The two different DL instruments introduced in Sect. 3.2.2 were deployed in a comparison field experiment at the DWD observatory MOLRAO in Lindenberg, which was conducted between 14 July 2020 and 20 October 2020. The aim of this field experiment was to compare the instruments of the Leipzig Institute for Meteorology (LIM) to similar systems of DWD. In this section, wind velocity measurements of the two DL instruments are compared to explore which device is more suitable to provide reliable wind speed information for wind power predictions. As a reminder, the main practical difference between both instruments is the different vertical resolution (LIMCUBE: 20 m, Streamline: 45 m) in their respective operational measurement configurations. The LIMCUBE's higher vertical resolution results in five available measurement levels (70 m, 80 m, 100 m, 120 m and 140 m) over the wind turbine rotor area, while the Streamline DL has only two (69 m and 115 m). Therefore, assuming the same quality of measurements between both instruments, LIMCUBE observations are expected to more accurately display the wind field in the rotor level.

Due to a overheating problems of the LIMCUBE in the beginning of the field experiment, only one month (24 September 2020 to 20 October 2020) of the three months observation period can be used for this comparison. However, one month of data is not sufficient to train a reliable ANN model configuration with the LoadManager®. That means a direct comparison of two power predictions based on the different instrument's observations is not possible. Therefore, only the wind velocity measurements of both instruments will be com-
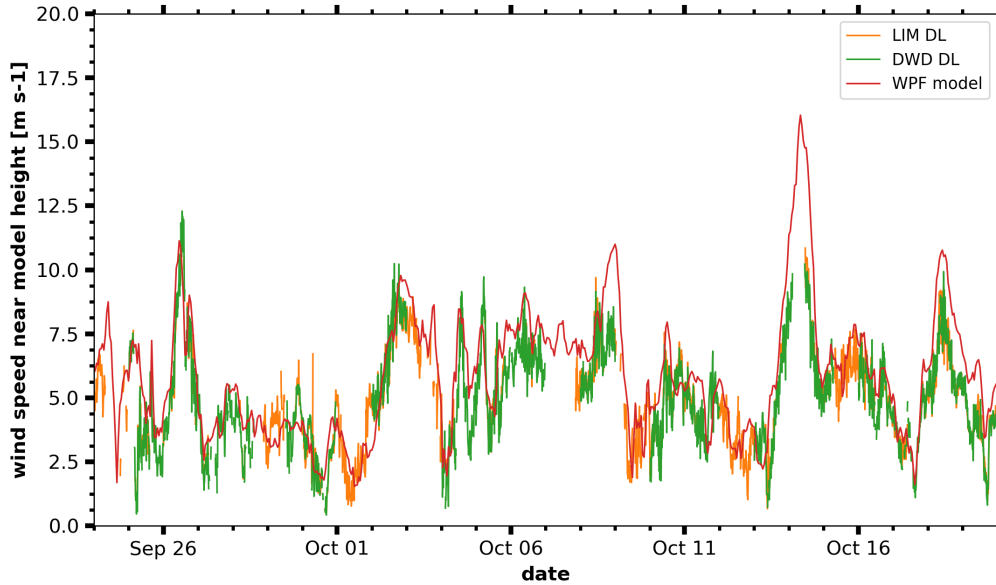
Figure 4.10: Time series of wind velocity measurements at 70 m observed during the field experiment at MOLRAO and WPF model data between 24 September 2020 and 20 October 2020. Periods with missing data for both DL are usually rain events.

pared, assessing similarities and differences in the observations to estimate how these could translate to better power forecasts.

Figure 4.10 provides an overview of the 70 m horizontal wind speed in the considered time period. This measurement level has been chosen because the WPF model data is interpolated to 70 m and both DL instruments have a nearby range gate (LIMCUBE: 70 m, Streamline: 69.5 m). Agreements between both DL measurements and the model data are generally good, however, the WPF data is smoother and tends to slightly overestimate the DL observations (as discussed in Sect. 4.2.2).

Since the other range gates of both instruments do not coincide, the measurements need to either be interpolated or compared to its nearest neighbour from the other instrument. However, some of the observation volume is outside of the typical extent of the logarithmic wind profile (see Sect. 2.2), so a general logarithmic interpolation is not suitable. Alternatives could be linear or spline interpolations. For this comparison, the Streamline DL observations were interpolated to the finer LIMCUBE grid. The adjustments to the wind speed values observed by the Streamline DL by both interpolation methods are usually in the order of $< 1\,\mathrm{m\,s^{-1}}$ compared to the nearest neighbour approach, depending on the vertical distance to the nearest range gate. Since differences between both interpolation approaches for this data set proved to be negligible, the linear interpolation has been chosen for the sake of simplicity.

Figure 4.11 shows 2D histograms of the horizontal wind speed comparison between both DL instruments for four range gates. Generally, both instruments agree very well at all investigated range gates. The red boxes show the regression parameters for a linear fit and the corresponding correlation coefficient R^2. The slope of the linear regression is very close to 1 for all shown range gates and $R^2 > 0.95$ further emphasizes the good fit.
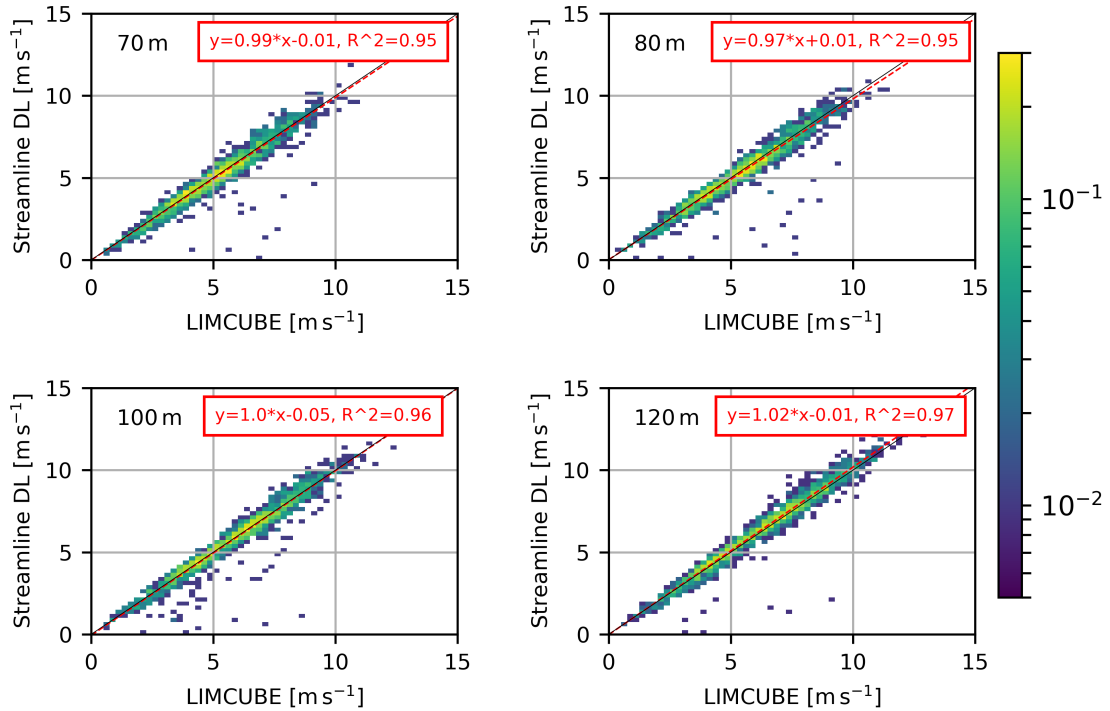
Figure 4.11: Two-dimensional histograms of the horizontal wind speed of LIMCUBE and Streamline DL during the comparison field experiment in Lindenberg between 24 September 2020 and 20 October 2020. The height in the top left corner refers to the range gate of the LIMCUBE observations, to which the Streamline DL observations have been linearly interpolated. The red boxes show the linear regression parameters for the linear fit (red line).

Therefore, it can be concluded that within this study both devices provide very similar observations in the height range of the investigated wind turbines and none of the two instrument's measurements can be preferred over the other. However, for the acquisition of a DL instrument, the Windcube can be preferred as it is the cheaper of the two instruments.

4.3 Short term wind speed prediction with the ARIMA model

While the simplicity of the persistence approach for short-term forecasting (see Sect. 3.4.2) is appealing, a more sophisticated approach should yield better results. An Auto-Regression Integrated Moving Average (ARIMA) model is a type of model containing different standard techniques to analyse and forecast time series data. In the following section, the three-part structure of an ARIMA model is introduced briefly, more details on it can be found in *Hyndman and Athanasopoulos* (2021). Subsequently, the process of determining the parameters needed for the ARIMA prediction is explained and the performance of this approach for the wind speed prediction is estimated.

4.3.1 Setup of the ARIMA model

The ARIMA model consists of three parts. The first of the three ARIMA parts is the auto-regression (AR) model. The auto-regression model is a kind of linear regression model. Such

models perform best if the time series is not correlated to itself and the variable at every time step is independent of previous time steps. Auto-regression uses the relationship between the observation and its previous ("lagged") time steps to predict the forecast variable and is characterised by the parameter p. An AR model of order p can be written as

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \tag{4.1}$$

calculating y_t, the forecast of the variable at time step t, using a constant c, the variable at previous time steps y_{t-n} with parameters ϕ_n and an error term ε_t, which is white noise. Here, the prediction of y_t depends only on its own lags.

The second ARIMA part, integration (I), uses differencing between the observation and the observation at a previous time step to eliminate trends and seasonality in the time series and make it stationary and is characterised by the parameter d. Differencing is the easiest way to achieve stationarity of the time series, which is desirable for good performance of the included AR model. The first difference (d=1) of y_t can be written as

$$y'_t = y_t - y_{t-1}. \tag{4.2}$$

and displays simply the change of variable y between time steps $t-1$ and t. The order of differencing determines the parameter d.

Lastly, the moving average (MA) part contains a model that uses the relationship between the observation and residual errors from a moving average function applied to previous observations, characterised by parameter q. The MA model can be written as

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \tag{4.3}$$

with the parameters θ and white noise error terms ε_{t-n} of previous time steps.

Combining the three parts yields the ARIMA model, written as

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \tag{4.4}$$

where the differenced series y'_t is a function of both its lagged observations and the lagged errors. The parameters p, d and q characterise the ARIMA model and have to be specified with suitable values. Differencing is needed, when the time series is non-stationary, exhibiting either a trend or a seasonal pattern. Although wind speed data do not show prominent seasonality as, e.g., the temperature, it can show trends, depending on the time frame of the evaluated series. Because of the more random nature of wind speed, longer wind speed time series tend to be more stationary. Whether differencing is needed for a certain time series can be evaluated with the Augmented Dickey–Fuller test. This test is a commonly used unit-root test, which examines the null hypothesis of an ARIMA model (*Cheung and Lai*, 1995). Non-stationary wind speed time series usually become stationary after first order differencing.

The ARIMA model can be expanded further for Seasonality (SARIMA) and for use with eXogenous factors (SARIMAX). A time series is seasonal if it exhibits a regular pattern, e.g., a daily, monthly or yearly cycle. Exogenous variables are secondary observational time series with the same time steps as the primary time series. These variables are used as parallel input next to the primary variable but not modeled after the ARIMA scheme. For example, when forecasting temperature, the variable itself (temperature) usually exhibits a strong seasonality on both a daily and a yearly cycle, suggesting the use of the SARIMA model. To further improve this forecast, a secondary observational variable, which influences the temperature, can be added as an exogenous factor. Variables like precipitation or cloud fraction are not

cyclical themselves and, therefore, difficult to predict with the ARIMA approach. However, they have a reasonably consistent effect on the temperature and would, presumably, improve the performance of a SARIMAX temperature model if added as exogenous factor.

Since the wind speed is not subject to seasonal variations as strongly as, e.g., the temperature, the seasonality module is not needed in this case. Similarly, the module to process exogenous input is not used, since there is no variable affecting the wind speed measured in parallel.

4.3.2 Determining the parameters of the ARIMA model

Figure 4.12 serves to illustrate the ARIMA parameters. If either p (AR part) or q (MA part) are 0, the other parameter can be found by evaluating the autocorrelation function (ACF) and the partial autocorrelation function (PACF). However, if both p and q are positive, those functions do not help in finding suitable values (*Hyndman and Athanasopoulos*, 2021). One way to determine these is to assign different values for p and q and compare the resulting wind speed simulations to the observational data set, as proposed by *Grigonytė and Butkevičiūtė* (2016). This was tested here for multiple possible combinations of p and q but did not consistently result in improvements compared to the persistence wind speeds.

The popular python package pmdarima (*Smith et al.*, 2017) includes the *auto_arima* function. This function is used to fit the best ARIMA model to any given time series, according to a provided information criterion (details can be found in the packages documentation). In that way, the model for each subset of the time series is based on the optimal set of parameters p, d and q. To ensure computational efficiency, possible p and q values were capped at 4, d was capped at 2.

Furthermore, the length of the subset of the time series used for the ARIMA model is investigated. The backlog b describes the amount of time steps before the prediction time step used to optimise the model (see Fig. 4.12). The longer the data subset (= the higher b), the more information the model has to determine its parameters θ (MA part) and ϕ (AR part). However, information on the variables pattern a few days before may not be relevant to its prediction, especially for wind speed. In this thesis, six values for b between 48 (referring to twelve hours) and 960 (10 days) are considered. The notation "ARIMA96" refers to an ARIMA model with a b=96.

The wind speed prediction period starts on 1 January 2019 and ends on 23 September 2019, containing approximately 25 000 time steps. The *auto_arima* function chooses a different parameter combination for every prediction time step, based on the backlog length b. Some ARIMA parameter combinations result in special cases for the model, notably, the ARIMA(0,1,0) model is equal to the persistence model. To evaluate the automatically chosen parameter combinations, they are counted and analysed. Figure 4.13 shows the number of occurrences of combinations (p,d,q), which are evaluated as optimal by the *auto_arima* function for more than 3 % of the time steps on at least one of the investigated ARIMA configurations. Combinations (p,d,q) which did not exceed this threshold are summarised into the last bin "others (48)", indicating 48 remaining combinations in that category.

The most prominent difference between ARIMA runs with different b values is the high amount of (0,1,0) (=persistence) occurrences for lower b values. The ARIMA96 configuration, calculating the optimal ARIMA model on a subset of 96 time steps (1 day), yields the (0,1,0) combination as best possible combination for approximately half of the time steps. For ARIMA48 and ARIMA192 the persistence model performs best for over 30 % of the time steps. Therefore, for this relatively large number of time steps, the ARIMA approach does not lead to an improvement over the persistence approach.
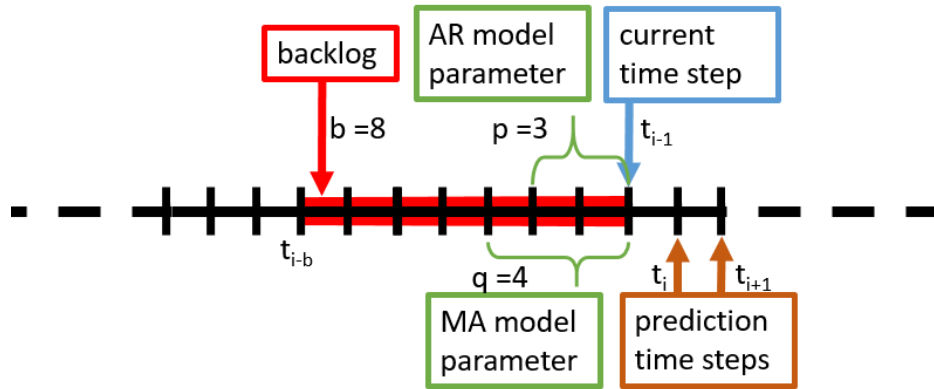
Figure 4.12: Illustration of the different parameters characterising the ARIMA model configuration with example values. The ARIMA models internal parameters are evaluated by the *auto_arima* function based on a subset of the time series of length b, beginning at time step t_{i-b} until the current time step t_{i-1} (red area). The *auto_arima* function yields the combination of the parameters p, d and q (green brackets), for which the prediction error is minimal. To predict the wind speed at time steps t_i and t_{i+1}, the ARIMA model (see Eq. 4.4) is applied with parameters p, d and q. There, p determines the amount of lagged time steps used for the auto-regression (AR) model, q the amount of lagged residual errors for the moving average (MA) model and d whether the time series needs to be differenced to reach stationarity.

However, for higher b values the number of occurrences is more evenly distributed between the shown combinations (Fig. 4.13). For a considerable amount of time steps the persistence combination (0,1,0) still performs well, but higher order combinations are increasingly represented. In general, most model combinations use parameters $p \leq 2$ and $q \leq 2$, indicating that usually only the previous 1-2 time steps are needed for the prediction of the next time step. Therefore, limiting $p \leq 4$ and $q \leq 4$ proves to be no issue for most configurations. For b=48, ..., 480 the share of time steps optimally predicted by combinations with p=4 or q=4 is less than 10 %. For ARIMA960, *auto_arima* determines for almost 25 % of the model steps a combination with p=4 or q=4 as optimal setup. Furthermore, combinations with $p \geq 3$ or $q \geq 3$ amount to approximately half of the ARIMA960 model setups. For this specific configuration with a large backlog b a higher limit of p and q could be investigated to increase its prediction accuracy.

In conclusion, the higher the length b of the subset of the wind speed time series, which is used to determine the optimal ARIMA model parameters, the higher frequency of higher orders (p,d,q) of the optimal ARIMA model. Shorter subsets favor more simple model configurations, with a large share of time steps resorting to the persistence configuration (0,1,0).
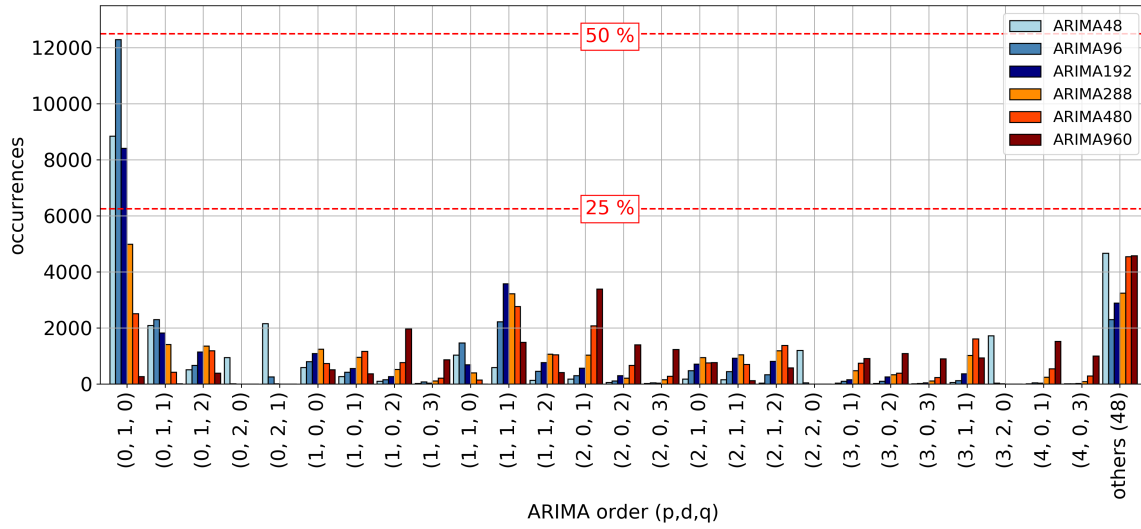
Figure 4.13: Order (p,d,q) of the different ARIMA configurations (colors) as returned by the *auto_arima* function. Each ARIMA configuration has been evaluated for the prediction of approximately 25 000 time steps. The number of times a specific ARIMA order (x-axis) has been determined as optimal is indicated at the y-axis. The horizontal dashed red lines mark an amount of $25/50\,\%$ of all time steps. ARIMA configurations with higher backlog b (reddish hues) are more evenly distributed over the shown ARIMA order triples than configurations with lower b (blue hues), which tend to favor lower order ARIMA models. The last bin "others (48)" summarises the remaining 48 combinations of (p,d,q).

4.3.3 Performance of the ARIMA model for wind speed prediction

Table 4.2 shows the root-mean-square error of the ARIMA modeled wind speed for six different values of backlog b and for the persistence model. For b values much lower than 48 the ARIMA model becomes unreliable and yields larger errors. The RMSE for b=48 is still higher than for the persistence shifted wind speed, but for the other configurations the RMSE of the ARIMA modeled wind speed is lower than persistence, indicating a better wind speed prediction through ARIMA. For the 15 min prediction ARIMA192 shows the lowest RMSE, roughly 5 % better than persistence. For the 30 min prediction, ARIMA288 shows the lowest RMSE, still about 2 % better than persistence. However, although the ARIMA approach reduces the RMSE of the wind speed prediction compared to the persistence approach, the improvement is rather modest. Consequently, the different wind speed time series have been subjected to a two-sample Student's t-test. Even though these data support the thesis that the ARIMA wind speed prediction is better than the persistence method, the effect is not statistically significant.

This improvement compared to the persistence approach is not evenly distributed over the whole data set. Times with fluctuating wind speed are situations where the ARIMA approach excels. Figure 4.14 shows an example of 24 h in June 2019 including the wind speed observed by the sonic anemometer (solid, grey), the persistence-shifted wind speed (solid, orange) and the prediction by the ARIMA192 model (dotted, black). From 19 UTC to 2 UTC the observed wind speed is fluctuating between consecutive 15 min time steps. The dotted ARIMA192 wind speed prediction follows the observed wind speed considerably better than the persistence predicted wind speed. For the 12 h period shown in this plot, the root-mean-square error of the persistence approach is at RMSE_{pers}=1.47 m s^{-1} twice as large as for the ARIMA192 model RMSE_{ARIMA}=0.74 m s^{-1}.

Table 4.2: Root-mean-square error in $[\mathrm{m\,s^{-1}}]$ for different ARIMA configurations in comparison to the persistence approach. Only time steps when the observed wind speed was higher than $4\,\mathrm{m\,s^{-1}}$ (cut-in wind speed of Vestas V80-2.0MW) were considered, since situations with lower wind speed do not contribute to wind power generation. However, relative errors between the different configurations and persistence for all wind speed values are similar to this.

prediction time	15 min	30 min
persistence	0.675	0.861
ARIMA48	0.709	1.006
ARIMA96	0.657	0.884
ARIMA192	0.645	0.854
ARIMA288	0.652	0.845
ARIMA480	0.647	0.846
ARIMA960	0.651	0.857

While the difference of the ARIMA approach compared to the persistence approach is relatively small, as indicated in Tab. 4.2, it still provides an improvement in wind speed prediction. One reason for that relatively small change is the chaotic nature of the wind speed. Furthermore it is found that more complex ARIMA configurations with high b values do not necessarily provide better wind speed predictions. For the prediction times evaluated here ($+15\,\mathrm{min}$ and $+30\,\mathrm{min}$), the optimal backlog values are $b=192$ (corresponding to 2 days) and $b=288$ (3 days), respectively. Sect. 5.2 explores, whether this slight improvement in wind speed prediction leads to an improvement in wind power prediction.

4.4 Observational data sets for wind power predictions

LoadManager® power predictions in this thesis (see Sect. 3.4) were typically performed for two years of training data, split evenly into a learning range and a validation range. Table 4.3 gives an overview of the input data sets, which will be used for wind power predictions. Since the wind power forecast requires a temporal resolution of 15 min, the input data are converted to that resolution via spline interpolation within the LoadManager®.

The WPF reference data set includes all model variables currently used for power predictions. For the power prediction process using observational data, the WPF wind velocity is replaced with the observed data sets of DL or nacelle sonic anemometers. In addition, the WPF model temperature on both available heights ($2\,\mathrm{m}$ and $355\,\mathrm{m}$), the atmospheric surface pressure and the surface relative humidity are used, since these variables proved to be important during the ANN training. The high weight of the temperature values is presumably caused by the dependence of the power generated through wind on the atmospheric stability. *Wharton and Lundquist* (2012) conclude that the power generated under stable conditions is up to $15\,\%$ higher than under strongly convective conditions for the same wind speed.

For the DL-based power predictions, three different heights ($116\,\mathrm{m}$, $208\,\mathrm{m}$ and $487\,\mathrm{m}$) are used as input variables for the ANN. The altitude of $116\,\mathrm{m}$ is chosen as the lowest input range gate, because it is closest to the wind turbines hub height at $100\,\mathrm{m}$. Two additional range gates are chosen to provide information about the atmospheric layers directly above the rotor layer. Doppler lidar measurements provide wind speed and direction information.

For the sonic-based power predictions, wind speed data are available at the hub height $100\,\mathrm{m}$. However, only the horizontal wind speed is used as ANN input variable, since the wind direction is not available for over $70\,\%$ of the time, as stated previously. Both the per-
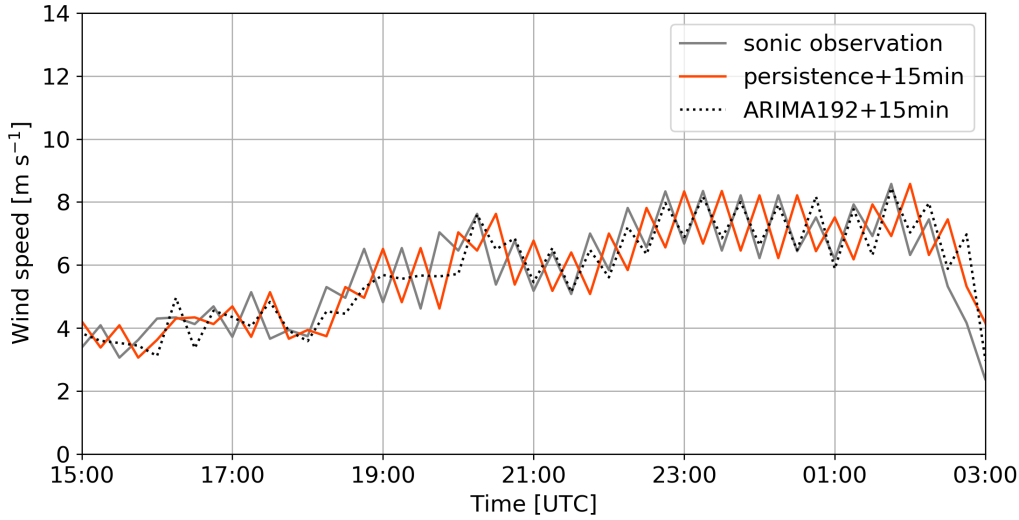
Figure 4.14: Time series of sonic anemometer wind speed observation for 24 h on 14 June 2019. In addition to the observation (gray), the corresponding persistence (orange) and ARIMA modeled (dotted) wind speed are displayed. Especially during periods with fluctuating wind speed, as seen here between 19 UTC to 2 UTC, the ARIMA approach outperforms the persistence approach.

sistence and the ARIMA approach are based on the nacelle sonic anemometer data set. Prior evaluation suggests to use the mean wind speed of the five turbines instead of measurements from a single turbine, because it resulted in the lowest statistical errors. This is supported by the analysis of wake effects within the wind farm in Sect. 4.1.3, showing that the difference between MT observations and the nacelle observations is least dependent on the wind direction if the wind speed data are averaged over the five turbines. For the ARIMA modeled wind speed detailed in Sect. 4.3, six configurations with different backlog values b are used as input for the LoadManager® wind power prediction.

Technically, the DL-based power prediction using the time shift of 30 min is also a persistence model. However, since it is the only DL-based configuration analysed in this thesis, it is not further specified in subsequent figures and discussions as "DL persistence" but instead referred to as "DL" configuration. The distinction between the "persistence" approach and the "ARIMA" approach is in the following only emphasised for the sonic-based power predictions.

In conclusion, there are 16 power prediction runs discussed in this thesis, one WPF reference run and 15 runs including observational wind speed data:

- 1 run based on WPF data (reference)

- 1 run based on DL data (+30 min)

- 2 runs based on persistence ultrasonic anemometer data (+15 min and +30 min)

- 12 runs based on ARIMA models ultrasonic anemometer data (six different backlog values, each for +15 min and +30 min)

Table 4.3: Input data sets for wind power predictions with wind velocity sources and additional variables from the WPF product (T - temperature, RH - relative humidity, p - atmospheric pressure)

simulation	forecasting horizon (min)	wind speed source	height (m)	additional variables (WPF)
WPF (reference)	-	WPF model	10, 70, 355	T (2 m, 355 m), RH, p, cloud fraction (low, medium, high), solar radiation
DL	30	Doppler Lidar	116, 208, 487	T (2 m, 355 m), RH, p
persistence	15, 30	sonic anemometer	100	T (2 m, 355 m), RH, p
ARIMA	15, 30	sonic anemometer	100	T (2 m, 355 m), RH, p

4.5 Wind power ramp events

Numerical Weather Prediction (NWP) models perform well in predicting the general development of the wind speed up to a few days in advance. However, typical challenges of wind power forecasts are features not resolved by the underlying meteorological model and generally rapidly changing wind speeds, so-called ramp events. The term wind ramps refers to large, rapid increases (up-ramps) or decreases (down-ramps) in wind speed. One cause of unexpected and sudden changes in wind speed can be low-level jets. However, as NWP models improve continuously, they are even able to predict local phenomena like low-level jets to some extent (*Tay et al.*, 2021). Other reasons for ramp events include, e.g., the passage of frontal systems or local circulation patterns.

This section discusses the detection of wind power ramp events and their occurrence in the data set investigated in this thesis. Wind power prediction results for wind ramp events in comparison to the entire study period will be analysed in Chap. 5.

4.5.1 Ramp detection methods

The influence of wind ramp events on wind power generation is a widely discussed topic (e.g., *Bossavy et al.*, 2013; *Bianco et al.*, 2016). A ramp event can be described by the following parameters (*Gallego-Castillo et al.*, 2015):

- **magnitude** (ΔP_r): the relative change in power generation during the ramp event,

- **window length** (WL): the time window during which the ramp event takes place,

- **ramp rate** ($\Delta P_r/WL$): the intensity of the ramp,

- **time** (t_0): the time of occurrence of the ramp, usually either the starting time or the center time,

- **direction**: whether it is an up-ramp or a down-ramp.

So far, there is no unified method for the identification of wind power ramps. *Gallego et al.* (2011) and *Pichault et al.* (2021a) outline methods to use continuous wavelet transformation to characterise ramp events. Other studies (e.g., *Kamath*, 2010; *Bianco et al.*, 2016) use binary approaches with similar definitions, which usually depend on the absolute or relative power change ΔP_r during a certain time window with a window length WL. Here, the latter approach is applied.

The values of these parameters then depend on the specific wind farm and user. *Gallego-Castillo et al.* (2015) analysed ramp definitions in the literature and found thresholds for ΔP_r between $10\,\%$ and $75\,\%$ and for WL between $30\,\text{min}$ and $6\,\text{h}$. Most common are ΔP_r between $30\,\%$ and $50\,\%$ for WL between $1\,\text{h}$ and $3\,\text{h}$. The magnitude ΔP_r usually refers to a percentage of the rated power of the wind farm. However, *Kamath* (2010) argues that the installed capacity of the wind farm might change over time, due to upgrades or temporary shut downs. On the one hand, a certain ramp event classified based on the rated power might not be classified as a ramp event on a later day, when the installed capacity increased in the meantime. The accompanying challenge for the grid operator, with a sudden increase or decrease in power generation, is however still the same. On the other hand, an absolute power change threshold would lead to an increase in the frequency of ramp events, if the installed capacity of the wind farm increases (*Gallego-Castillo et al.*, 2015).

Since the capacity of the studied wind farm Beeskow did not change during the investigated period, either a relative or absolute power change can be utilised, yielding the same results. However, for the sake of clarity, the relative power change will be used in this study. *Bianco et al.* (2016) propose three methods of finding wind power ramps in a data set; the *fixed-time interval* method, the *minimum-maximum* method and the *explicit-derivative* method, briefly explained in the following. In-depth information about the three methods can be found in *Bianco et al.* (2016).

First, the fixed-time interval method, using a sliding time window of the specified window length WL is explained. For every time step of the data set, the difference in power between the start and end point of this window is calculated

$$\Delta P_r = P_s - P_e. \tag{4.5}$$

If the difference in power exceeds the threshold, all time steps in this time window are considered a ramp event. The sign of ΔP_r determines the direction of the ramp (see Fig. 4.15a). While the approach is simple, *Bianco et al.* (2016) identify two possible disadvantages: larger values of ΔP_r may occur within the time window than defined by its edges and two ramp events with opposite direction may be overlapping in time.

Secondly, the minimum-maximum method applies a similar sliding time window of length WL. However, instead of the power difference between the start and end time step, the difference between the observed maximum and minimum power during that window is calculated:

$$\Delta P_r = P_{max} - P_{min}. \tag{4.6}$$

If ΔP_r is above the threshold, this time window is considered as a ramp event. If $t(P_{max}) > t(P_{min})$, it is considered an up-ramp, otherwise a down-ramp event (see Fig. 4.15b). This method eliminates the two disadvantages of the fixed-time interval method; ramp events only last from one extreme value to the other and, while one time step may simultaneously be end- and start point for oppositely directed ramps, up- and down-ramps never overlap.

Lastly, the explicit derivative method calculates a smoothed derivative of the power generation over all time steps within the time window. If $|\partial P/\partial t| \geq \Delta P_r/WL$, the power change within that time window is considered as a ramp event. The algorithm then searches the beginning of the up-ramp (down-ramp) event by identifying the minimum (maximum) in a time interval $WL/2$ before the ramp time step. The end is found by detecting the maximum (minimum) in $WL/2$ after the initial ramp time step (see Fig. 4.15c, d).

For all methods, if continuous time steps longer than WL are marked as a ramp event with the same direction, it is concatenated into one ramp event. Therefore, single ramp events may be longer than WL.
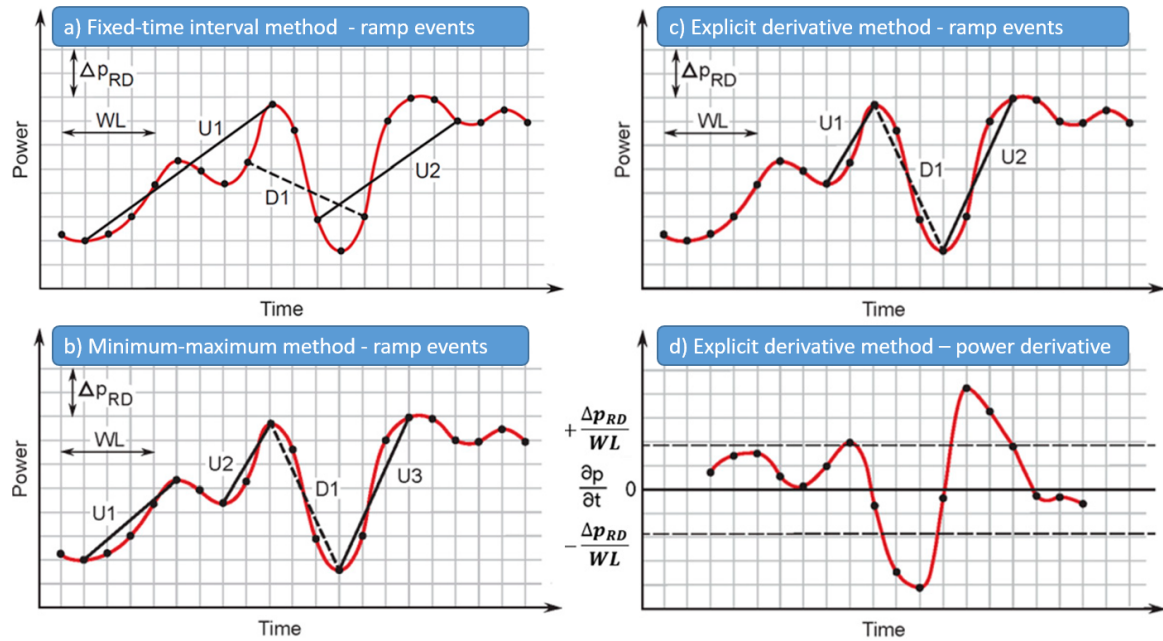
Figure 4.15: Different ramp detection methods evaluated in this thesis. Solid lines describe up-ramp events, dashed lines down-ramp events. WL is the window length of the ramp window, ΔP_{RD} the defined ramp power threshold. a)-c) show ramp events found for a theoretical power time series, d) shows the smoothed power derivative corresponding to the power curve from the other panels. Figure adapted from *Bianco et al.* (2016).

For the wind farm Beeskow, power data from 1 January 2016 until 30 September 2019 have been analysed for the three different ramp detection methods proposed by *Bianco et al.* (2016). In the following, the absolute difference in number of wind ramp events will be discussed, before analysing the seasonal and diurnal cycle of ramp events at the wind farm Beeskow.

Table 4.4 shows the absolute number of ramp events for the three different ramp detection methods for five ramp power thresholds ΔP_r (30 %, 40 %, 50 %, 60 % and 70 %) and three sets of window lengths WL (30 min, 60 min and 120 min). Since both the fixed-time interval method and the explicit derivative method are based on fixed window lengths, their absolute ramp numbers are similar. If they differ, the latter counts slightly more ramp events, because there the ramp events are only defined from the minimum to the maximum during WL and, therefore, two consecutive ramp events with only one or two non-ramp time steps in between are counted as two separate events, while the fixed-time interval method likely concatenates them into one event. This becomes clear when considering the average ramp time length for the fixed-time interval method, which is 15-25 % longer than for the explicit derivative method (not shown). The minimum-maximum method counts on average more ramp events than the other two methods. Since it evaluates the power change within the time window from minimum to maximum, and not from its start to the end, this method also finds ramps shorter than WL and, therefore, detects more ramp events. However, seasonal and daily patterns are displayed independently of the method. Since the minimum-maximum method avoids the possible drawbacks of the fixed-time interval method and can reliably register ramp events shorter than WL, the minimum-maximum method will be used predominately for the next analyses.

Table 4.4: Absolute number of up-ramp (U) and down-ramp (D) events from 1 January 2016 until 30 September 2019 in Beeskow. Ramp events are classified by the power change ΔP_r over a time window with window length WL with the fixed-time interval method, minimum maximum method and explicit derivative method by *Bianco et al.* (2016) for three ramp power thresholds ΔP_r (30 %, 40 % and 50 %). Using higher ΔP_r thresholds (60 % and 70 %) leads to a further decrease of the number of identified ramp events.

WL	ΔP_r ramp direction	30 % U	D	40 % U	D	50 % U	D
30 min	fixed-time interval	311	303	149	122	72	54
	minimum-maximum	344	332	162	128	79	56
	explicit derivative	311	305	149	124	72	54
60 min	fixed-time interval	408	406	198	181	95	96
	minimum-maximum	528	521	250	231	116	113
	explicit derivative	415	412	199	181	95	96
120 min	fixed-time interval	485	489	227	261	130	134
	minimum-maximum	679	682	338	347	172	180
	explicit derivative	496	505	230	262	130	134

4.5.2 Ramp occurrences found in previous studies

In this section, the occurrence and frequency of ramp events will be discussed. The understanding of the synoptic reasons for the occurrence of ramp events is of great importance for the development of ramp adaptation tools. However, since the definitions of ramps are varying, the synoptic reasons for differently defined ramps vary as well.

The grid area on which the ramps are evaluated determines the scale of the underlying synoptic processes. If the analysis is based on a larger grid region with multiple wind farms, the causes for ramps are expected to be large scale meteorological processes, like frontal zones, because they are expected to affect most of the wind farms equally. If the ramp events are analysed for a single wind farm or turbine, micro- and mesoscale processes are likely becoming more relevant, because local effects can still affect a single wind farm's output considerably (e.g., *Gallego-Castillo et al.*, 2015; *Cutler et al.*, 2007).

Zack (2007) divides ramp events into horizontally and vertically induced events. Horizontal processes include large-scale events, like frontal passages, as well as mesoscale events, like regional circulation patterns. Both are reasonably well represented by NWP models. In contrast, vertical processes, e.g., driven by convection, are generally not that well represented by NWP models. Therefore, their prediction is considerably more difficult. Still, even if the specific meteorological processes are not captured in exact timing and location by the model, it can provide useful information hinting at their occurrence.

Gallego-Castillo et al. (2015) summarised studies on ramp event occurrences for multiple sites in North America and revealed that up-ramps were found more frequently than down-ramps. For the relatively flat Canadian Alberta region *AWS-Truewind* (2008) found that up-ramp events take place mainly in May to July, down-ramp events rather from August to January. *Kamath* (2010) found for the Columbia Basin region, North-Western USA, that (i) moderate to large down-ramps rarely occur in the morning, (ii) up-ramps also occur more frequently in the afternoon than in the morning and (iii) there are more days with up-ramps than with down-ramps. Still, even though the results of these studies mostly agree, *Gallego-Castillo et al.* (2015) note that due to the complex nature of ramp events and the limited

sample of study regions, the reported experiences cannot easily be generalized to case studies at other sites. For Germany, *Steiner et al.* (2017) investigate meteorological causes for the occurrence of wind power ramps, focusing on cyclone activity. Germany-wide power forecast errors were analysed and a connection with cyclones or troughs in the North Sea, Baltic Sea or over Germany identified. However, studies on the diurnal cycle and the seasonality of ramp events in Germany are sparse.

Since the reasons and the frequency of ramp events is largely dependent on case specifics, the above-mentioned findings will be used as a base for discussion of ramp events at the wind farm analysed in this thesis in Sect. 5.1.

5 Results and discussion

Though the quality of wind power predictions has been consistently increased during the last decade, persistent reasons for remaining uncertainties are ramp events (introduced in Sect. 4.5). The occurrence of ramp events at the studied wind farm Beeskow is discussed in Sect. 5.1 and evaluated for the diurnal and seasonal cycle for almost four years of power generation data. Afterwards, the quality of the performed wind power predictions is analysed. For that purpose, two different aspects are investigated. Section 5.2 discusses the quality of the power predictions regarding their absolute power difference to the observed power generation. In this *volume*-oriented approach, a power prediction configuration based on one wind speed source is better than a power prediction configuration based on another wind speed source, if the difference to the actually generated wind power is lower and, therefore, the necessary balancing services need to be less extensive. Contrarily, Sect. 5.3 details a *cost*-oriented approach. Every balancing service is associated with a certain cost. Since the balancing price is not constant and can be either positive or negative in sign, a higher absolute power difference at a certain time step may not necessarily lead to higher balancing costs for the wind farm operator. Therefore, this section aims to identify the most profitable power prediction configuration. In Sect. 5.4 both approaches are compared and the results are summarised briefly. Consequently, the research questions for this chapter are:

- When and how often do ramp events occur at the study site Beeskow?

- Do ramp events exhibit a seasonal or diurnal cycle and do they agree to previous studies?

- Do observational data as additional power prediction input improve the quality of wind power predictions?

- Is the *volume*-oriented or the *cost*-oriented approach more meaningful?

5.1 Occurrence of ramp events

The occurrence of ramp events at the wind farm Beeskow has been evaluated for the three different ramp detection methods introduced in Sect. 4.5. Five sets of ramp power thresholds ΔP_r (30 %, 40 %, 50 %, 60 % and 70 %) and three sets of window lengths WL (30 min, 60 min and 120 min) are evaluated in this study. For the wind power generation profile of the studied wind farm, only the most extreme ramps exceed a ΔP_r of 50 %. All figures provided in the following show results for WL=120 min (8 time steps) and ΔP_r=30 %, if not indicated otherwise. This is the least strict combination and, therefore, shows the highest absolute number of ramp events. However, all ramp threshold combinations have been evaluated and differences to the shown combination setup are described in the text. In App. A.2, Fig. A.3 and Fig. A.4 show results for six additional ramp threshold combinations (ΔP_r=30 %, 40 %, 50 % and WL=30 min, 60 min).

Several of the results in this section are already included in *Lochmann et al.* (2022), of which a revised version is currently under review. However, the analysis in this thesis covers additional aspects compared to the manuscript.
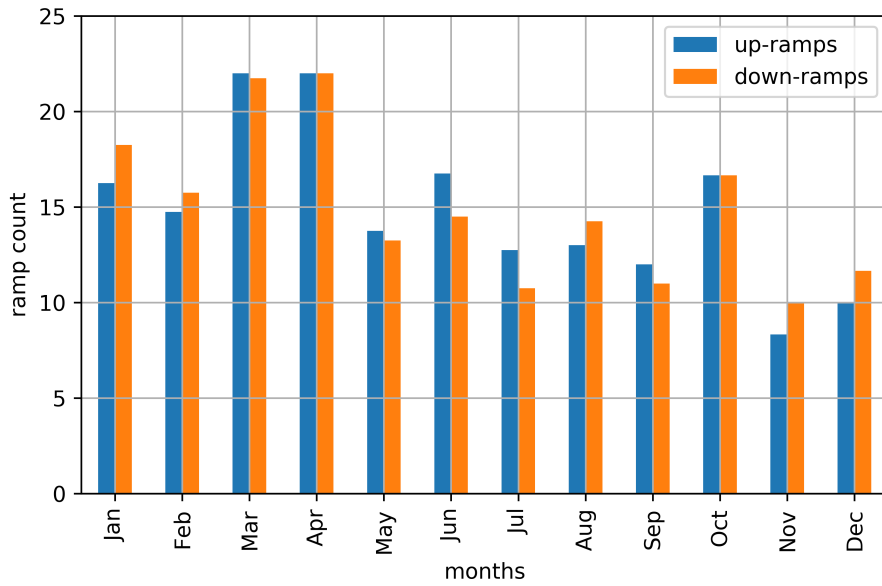
Figure 5.1: Annual occurrence of wind power ramp events per month, based on 45 months of data (obtained between 1 January 2016 and 30 September 2019), exemplified for window length WL=120 min and ramp power change threshold ΔP_r=30%. More ramp threshold combinations are shown in Fig. A.4.

5.1.1 Seasonal occurrence of ramp events

Figure 5.1 shows the average number of ramp events per month, based on 45 months of data (obtained between 1 January 2016 and 30 September 2019). While for January to September four years of data are available, for October to December only three years of power generation data are available. Therefore, the monthly ramp count per year has been calculated by dividing the monthly sum of ramp events between January and September by four and for October to December by three. A seasonal pattern is obvious independently of the ramp definition. Up- and down-ramp events are most frequent in March and April and least frequent in November and December. During the rest of the year, their number is relatively constant. The observed maximum in spring almost coincides with the maximum of ramp activity found by *AWS-Truewind* (2008) for Alberta, albeit it is observed two months earlier in Germany. The study region in Alberta is situated close to the Rocky Mountains, which is characterised by subarctic climate (Köppen-Geiger: Dfc), while the region around Beeskow, Germany, is classified as warm humid continental climate (Köppen-Geiger: Dfb). Compared to Germany, the on average colder subarctic climate in Alberta with a later spring onset is a possible explanation for a later onset of increased ramp activity.

For most combinations of WL and ΔP_r the number of up-ramps in Beeskow throughout the year is marginally higher than the number of down-ramps, similar to the previous studies (*AWS-Truewind*, 2008; *Kamath*, 2010). Additionally, in Beeskow there are slightly more up-ramp than down-ramp events between March and July and slightly more down-ramp then up-ramp events during November to February, however, their difference in frequency of occurrence is not very pronounced.

Table 5.1 shows seasonal occurrences and the diurnal cycle of observed wind power ramps for all the investigated power ramp thresholds and window lengths. The resulting patterns are similar for the different thresholds.

The upper half of Table 5.1 lists the seasonal results of detected ramp events. The total number of ramp events show a slightly higher number of up-ramps than down-ramps for most combinations of WL and ΔP_r. This agrees with previous studies and is explained by the passage of frontal systems, usually causing quick wind speed increases followed by a wind speed decrease over a longer time period (*Gallego-Castillo et al.*, 2015). This effect seems to be more relevant for shorter WL, as for the longest analysed WL=120 min, all except one ΔP_r register more down- than up-ramp events.

Furthermore, the number of total observed ramp events shows an approximate reduction by a factor of two when increasing ΔP_r by 10 %, independently of the chosen WL. Looking at the seasons separately, this reduction of the number of ramp events by half is applicable predominately in spring and summer. For all WL, categories with high ΔP_r only consist of a small number of samples and, therefore, do not always fit this pattern. While this pattern is not necessarily true for all seasons and combinations of thresholds, it is a helpful rule of thumb for wind farm operators and their choice of ΔP_r in the definition of ramp events.

All combinations of WL and ΔP_r show the highest number of ramps in spring, followed by summer and the lowest number in autumn and winter, agreeing well with *AWS-Truewind* (2008). However, there is no significant difference in seasonality between up- and down-ramp events.

Table 5.1: Upper half: Number of monthly observed wind power ramp events per year classified by three different window lengths WL and five different ramp power thresholds ΔP_r and divided by season. Wind power generation data from 01 January 2016 until 30 September 2019 (adding up to 45 months), the seasonal occurrences have been weighted per year. Lower half: Number of observed wind power ramp events per time of day for three sections: night (21-05 UTC), morning (05-13 UTC) and afternoon (13-21 UTC).

WL		30 min					60 min					120 min				
seasonal	ΔP_r	30%	40%	50%	60%	70%	30%	40%	50%	60%	70%	30%	40%	50%	60%	70%
DJF	up	15.3	5.8	2.8	1.5	0.6	28.5	10.7	4.2	1.8	0.8	41.0	20.1	7.5	2.8	0.8
(winter)	down	16.8	3.3	0.7	0.7	0.3	32.3	10.0	4.3	1.8	0.67	45.7	19.8	10.0	4.0	0.9
MAM	up	33.3	16.5	7.8	4.0	1.5	48.5	24.5	11.8	6.0	2.8	57.8	32.0	17.5	10.5	5.0
(spring)	down	34.0	15.0	7.0	2.5	0.8	45.8	24.5	12.3	5.8	2.8	57.0	31.3	17.3	9.5	6.5
JJA	up	26.3	13.8	8.0	6.0	3.0	34.5	18.8	11.3	7.0	4.3	42.5	21.0	13.8	8.0	5.5
(summer)	down	22.3	10.3	5.5	3.0	1.5	32.0	15.3	9.0	5.3	3.8	39.5	22.3	11.3	6.0	4.8
SON	up	14.0	5.5	1.42	0.7	0.3	26.5	10.8	2.3	0.9	0.9	37.0	15.1	5.7	2.6	0.9
(autumn)	down	13.6	4.0	0.8	0.8	0.6	27.0	10.42	3.6	1.1	0.6	37.7	17.6	8.6	2.3	1.17
total	**up**	**88.8**	**41.5**	**20.0**	**12.2**	**5.4**	**137.7**	**64.8**	**29.5**	**15.7**	**8.8**	**178.3**	**88.2**	**44.4**	**23.9**	**12.3**
	down	**86.6**	**32.6**	**14.0**	**7.0**	**2.9**	**137.1**	**60.1**	**28.8**	**13.8**	**7.8**	**179.8**	**90.8**	**47.1**	**21.8**	**13.3**
diurnal	ΔP_r	30%	40%	50%	60%	70%	30%	40%	50%	60%	70%	30%	40%	50%	60%	70%
21-05 UTC	up	36	13	5	3	1	75	27	12	6	2	106	51	21	12	2
(night)	down	43	11	4	3	1	65	23	8	6	4	102	46	18	9	6
05-13 UTC	up	109	45	18	10	3	180	78	30	14	8	271	120	56	25	14
(morning)	down	94	33	12	4	2	165	67	29	10	6	224	100	56	21	10
13-21 UTC	up	198	83	55	34	17	272	144	73	41	24	301	166	91	56	32
(afternoon)	down	194	103	39	20	8	290	140	75	38	20	355	200	105	54	36
total	**up**	**343**	**141**	**78**	**47**	**21**	**527**	**249**	**115**	**61**	**34**	**678**	**337**	**168**	**93**	**48**
	down	**331**	**147**	**55**	**27**	**11**	**520**	**230**	**112**	**54**	**30**	**681**	**346**	**179**	**84**	**52**

5.1.2 Diurnal occurrence of ramp events

The dynamic of the wind field in the atmospheric boundary layer exhibits a diurnal cycle, especially on days with a high amount of incoming solar radiation at ground (see Sect. 2.2). On a typical day, the wind speed in the lower layers increases after sunrise when the heating of the surface leads to thermal convection. During the day, a well mixed convective boundary layer develops as explained in Sect. 2.1. Towards sunset, the momentum transport from the surface ceases and a stable boundary layer is formed at night, typically reducing the mean wind speed in the lowest layers.

The diurnal cycle of ramp events observed at the wind farm Beeskow is shown in Fig. 5.2. Ramps are sorted by the starting time of the individual event. The Figure shows that both up- and down-ramp events mostly start during daytime. Up-ramp events are significantly more frequent than down-ramps in hours 6 to 11. This coincides with the start of thermal convection during the first hours of the day, leading to momentum transport through the atmospheric boundary layer, which can be one reason for local ramp events.

From hours 12 to 15, down-ramps are found to be more frequent. Toward the end of the day, the wind speed typically decreases again. However, since there are seasons with strong thermal activity throughout the day, the number of up-ramp events is consistently high during daytime until hour 18. During the night, ramp activity is lower and down-ramps are slightly more frequent.

Ramp events caused by large-scale synoptic phenomena (e.g., fronts) are less likely connected to a diurnal cycle and, therefore, presumably occur during all hours of the day. The findings for Beeskow agree well with *Kamath* (2010) in (i), observing less down-ramps in the morning, and (ii), more up-ramp events in the afternoon than in the morning.

The lower half of Table 5.1, examining the diurnal cycle of ramp events, roughly shows a reduction of ramps by half when increasing ΔP_r by 10 % as found for the whole data set, although again not necessarily for all combinations of WL and ΔP_r. Furthermore, most ramp events occur in the afternoon (13-21 UTC). For shorter WL, the share of ramps found in the afternoon is larger: While for WL=30 min and ΔP_r=30 % afternoon ramps make up roughly 60 % of the total ramps, for WL=120 min and ΔP_r=30 % afternoon ramps represent less than 50 % of total ramps. In addition, for higher ΔP_r, the frequency of afternoon ramps for all WL increases, e.g., to 80 % for WL=30 min and 70 % for WL=120 min. Summarising, this means that the occurrence of steep power ramps, either due to short WL, large power changes or both, is most likely in the afternoon.
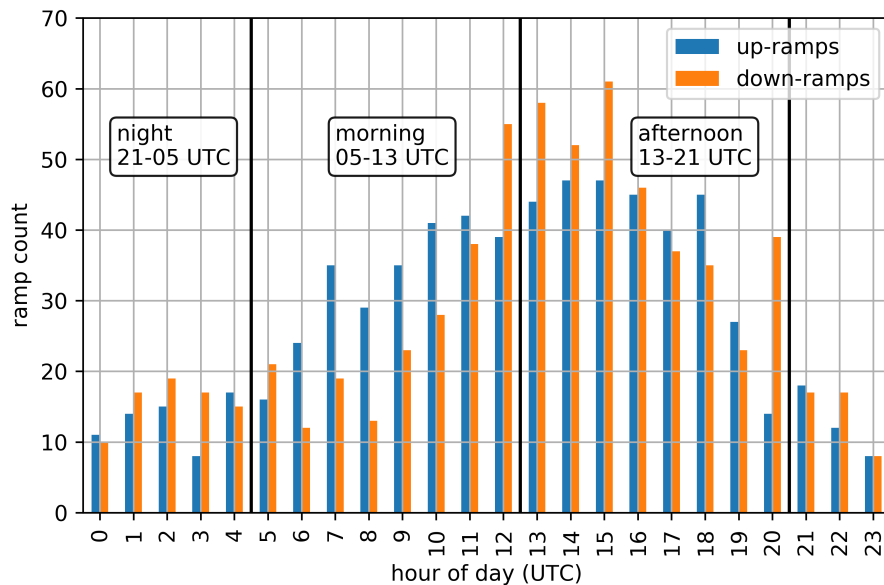
Figure 5.2: Total number of ramp events per hour of day (UTC), classified by starting hour of the ramp event. Based on 45 months of data, exemplified for window length WL=120 min and ramp power change threshold ΔP_r=30%. More ramp threshold combinations are shown in Fig. A.3.

5.1.3 Frequency of occurrence of ramp events in power predictions

The occurrence of wind power ramps in power predictions is investigated for the different wind speed input configurations introduced in Sect. 4.4. They are analysed for the number of wind power ramps predicted and whether these predicted wind power ramps match with observed wind power ramps. The number of "detected" ramp events refers to the total number of ramp events predicted. If predicted ramp events align in at least one time step with observed ramp events, they count as "matches". To keep results concise they are shown for only one ramp threshold configuration (WL=120 min, ΔP_r= 30 %). While other ramp threshold configurations with stricter thresholds show a smaller total number of ramp events, their results proportionally agree with the results shown here.

For the power prediction period of about nine months (see Fig. 4.5), the number of observed down-ramps is slightly higher than the number of up-ramp events (see Fig. 5.3). In general, power predictions based on observational data predict three to ten times more ramp events than the WPF prediction, depending on the used observational data set.

The WPF prediction predicts only less than 10 % of the observed ramp events. The reason for this is the smooth input wind speed data, generated by the WPF model. While the power prediction with this base configuration of WPF without additional input is on average competitive with the other prediction setups (see Sect. 5.2), it cannot predict ramp events well. Most of the few ramp events that are predicted by the WPF setups do match with observed ramp events. Predictions based on DL data capture more ramp events than WPF predictions, around 20 % of the observed number of ramps. All of the ramp events detected with the DL-based power predictions can be matched with observed ramp events.

Predictions based on nacelle sonic wind speed observation capture by far the highest percentage of ramp events, 35 % to 70 % of the observed ramps, based on the chosen ramp threshold configuration (see Fig. 5.3). This holds true for ARIMA configurations and persistence configurations, as they are both based on sonic anemometer wind speed data. ARIMA48, al-
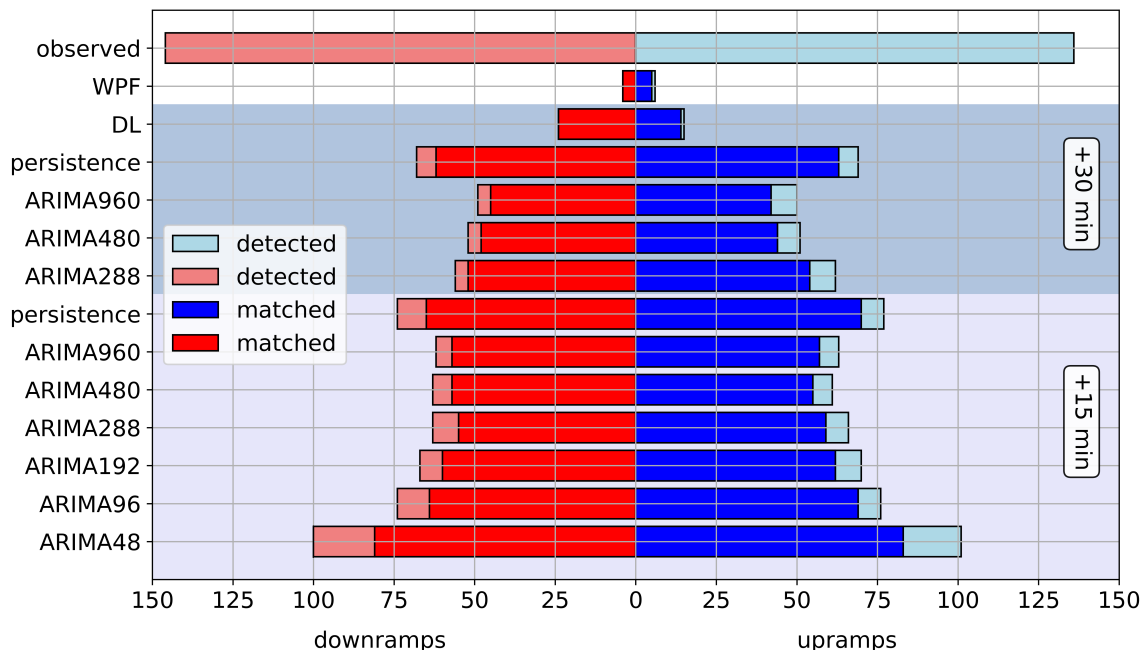
Figure 5.3: Number of ramp events for different power prediction configurations. Light red (blue) bars indicate the total number of detected down-ramps (up-ramps); strong red (blue) bars show the number of ramp events matched to observed ramp events.

though having the highest RMSE for its wind speed predictions (see Sect 4.3.3), captures the highest number of ramp events. The higher b for the ARIMA configurations, the smaller is the number of ramp events captured. This is due to the *auto_ arima* function more frequently yielding higher order AR (p value, see Sect. 4.3) and MA (q value) models for the wind speed prediction. This results in more lagged observations and their residual errors used to predict the next time steps wind speed value, therefore, exhibiting a slight smoothing effect on the prediction. ARIMA configurations with lower b feature lower order p and q combinations more frequently (see Sect. 4.3), resulting in less smoothing effects. Generally, power prediction setups for a forecasting horizon of 15 min detect a slightly higher total number of ramp events than for a forecasting horizon of 30 min, but otherwise not much difference is observed.

5.2 Volume-oriented analysis

The quality of a power prediction configuration is evaluated by the error between the predicted and the generated power of the selected wind farm. In the following, this is called power deviation P_{dev}. It is evaluated for wind power predictions based on the wind speed input data sets summarised in Sect. 4.4. A power deviation $P_{dev}>0$ means that the power prediction underestimated the actual wind power generation and more power is generated than previously predicted. Contrarily, $P_{dev}<0$ means that the power prediction overestimated the actual wind power generation and less power is generated than previously predicted.

As for the previous section, several of the results in this section are already included in *Lochmann et al.* (2022), of which a revised version is currently in revision. However, the analysis in this thesis covers additional aspects compared to the manuscript.
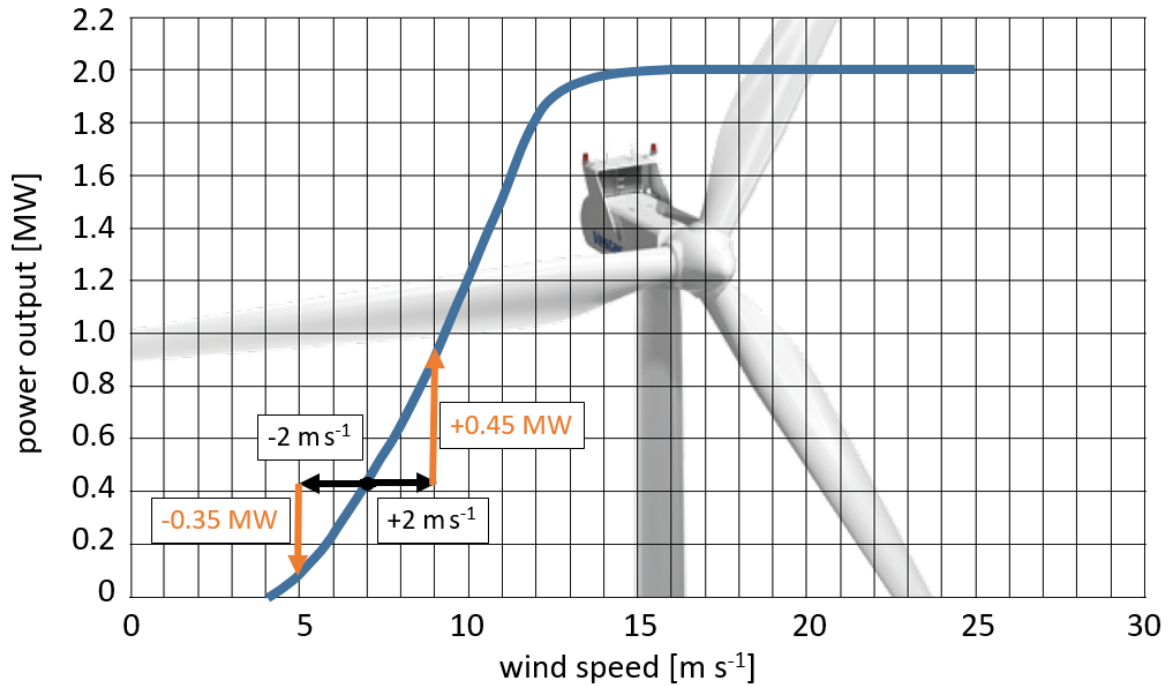
Figure 5.4: Vestas V80 power curve with wind speed errors (Figure adapted from *Vestas*, 2011). The same absolute error in wind speed (black) results in different absolute changes in power generation (orange). Wind speed underestimations lead to larger absolute deviations between power prediction and power generation than wind speed overestimations.

5.2.1 Bias in LoadManager® power predictions

Power generation by wind energy turbines depends mostly on the wind speed. Furthermore, the characteristic power curve of the wind energy turbine is not a linear function (see Sect. 3.1). If the real wind speed deviates from the wind speed prediction, the corresponding wind power prediction will be erroneous. However, because of the nonlinear relation, the deviation in generated power is not equal for over- and underestimations of the wind speed.

Figure 5.4 shows an example for a wind speed deviation of $2\,\mathrm{m\,s^{-1}}$ in both directions from the initial prediction of $7\,\mathrm{m\,s^{-1}}$. If the wind speed is overestimated, i.e., the real wind speed is $2\,\mathrm{m\,s^{-1}}$ lower than predicted, the power generation of one wind energy turbine is approximately $0.35\,\mathrm{MW}$ lower than expected. If the wind speed is underestimated, i.e., the real wind speed is $2\,\mathrm{m\,s^{-1}}$ higher than predicted, the power generation of one wind energy turbine is approximately $0.45\,\mathrm{MW}$ higher than expected. That means, for the predominant medium load situations with wind speed values between $4\,\mathrm{m\,s^{-1}}$ and $8\,\mathrm{m\,s^{-1}}$, the same absolute difference between the observed wind speed and the predicted wind speed with different signs leads to higher values of P_{dev} while underestimating wind speeds (Fig. 5.4). Although this effect reverses when the power curve flattens for wind speed values above $12\,\mathrm{m\,s^{-1}}$, since these are less frequent (see Fig. 5.5a) it does not compensate this effect completely. For wind speed values between $8\,\mathrm{m\,s^{-1}}$ and $12\,\mathrm{m\,s^{-1}}$ the power curve is almost linear, causing moderate wind speed errors with either sign to produce power deviations of similar absolute values. Therefore, if the wind speed prediction errors are evenly distributed around zero, wind power prediction errors (P_{dev}) are *not* evenly distributed but instead exhibit a bias towards underestimations of generated power.

Figure 5.5 shows statistics for wind speed values observed by the nacelle sonic anemometer to support these arguments. Figure 5.5a shows the probability density of wind speed values, highlighting that the majority of wind speed values are below $8\,\mathrm{m\,s^{-1}}$. Figure 5.5b represents a 2D histogram of the wind speed values in comparison to the generated power per wind energy turbine and the characteristic power curve (black line). This figure shows that the highest frequency of occurrence of these data pairs agrees well with the theoretical power curve. To estimate the power generation per wind energy turbine, the power generation of the whole wind farm has been divided by the number of turbines, because only the power generation of the whole wind farm combined is available. Since the power generation is a mean value over the whole wind farm, the theoretical maximum power generation of $2\,\mathrm{MW}$ can only be reached when every wind energy turbine is operating at maximum capacity. However, this is unrealistic in the local wind climate and the maximum of $2\,\mathrm{MW}$ per wind energy turbine is not reached. Figures 5.5c and 5.5d show that the wind speed prediction error for the nacelle sonic anemometer and DL observations is evenly distributed around zero. The distribution of the sonic wind speed prediction error in Fig. 5.5c is slightly narrower and has a more distinct peak at $0\,\mathrm{m\,s^{-1}}$ than Fig. 5.5d, but is otherwise similar. For better comparability, the wind speed prediction error is only shown for the forecasting horizon of 30 min for both the sonic persistence and the DL configurations. For the shorter forecasting horizon of 15 min, the distributions are even narrower and exhibit a higher peak at $0\,\mathrm{m\,s^{-1}}$. Furthermore, the wind speed prediction error of the ARIMA setups looks similar to the one of the sonic persistence configuration (not shown).

Therefore, since medium wind speed values below $8\,\mathrm{m\,s^{-1}}$ are more frequent than values above $12\,\mathrm{m\,s^{-1}}$ and the wind speed prediction error is uniformly distributed in over- and underestimations, a bias towards underestimations of the generated wind power by the LoadManager® is caused.

Figure 5.6 shows histograms of the power deviation P_{dev} for four wind power prediction setups. In the top-right corner the mean of the power deviation $\overline{P_{dev}}$ is displayed. For all wind power prediction setups $\overline{P_{dev}}$ is negative, showing the previously explained underestimation of the wind power prediction. The reference WPF setup and the DL setup show the highest absolute mean deviations $\overline{P_{dev}}$ and the histograms are slightly skewed towards negative values. The mean power deviation of the DL setup is even lower than $\overline{P_{dev}}$ of the WPF setup, which is mostly due to more very low P_{dev} values ($P_{dev}<3\,\mathrm{MWh}$) in the DL power prediction than in the WPF power prediction. This becomes apparent when comparing the median values of the WPF and DL configurations, which are similar. For configurations using nacelle sonic anemometer data, namely the persistence and the ARIMA configurations, this mean underestimation is about a factor of 3-4 smaller and the histogram is more symmetrical around $P_{dev}=0\,\mathrm{MWh}$. The narrower, more symmetrical distribution of P_{dev} is another advantage of the sonic-based wind power predictions, which additionally feature median values close to $0\,\mathrm{MWh}$.
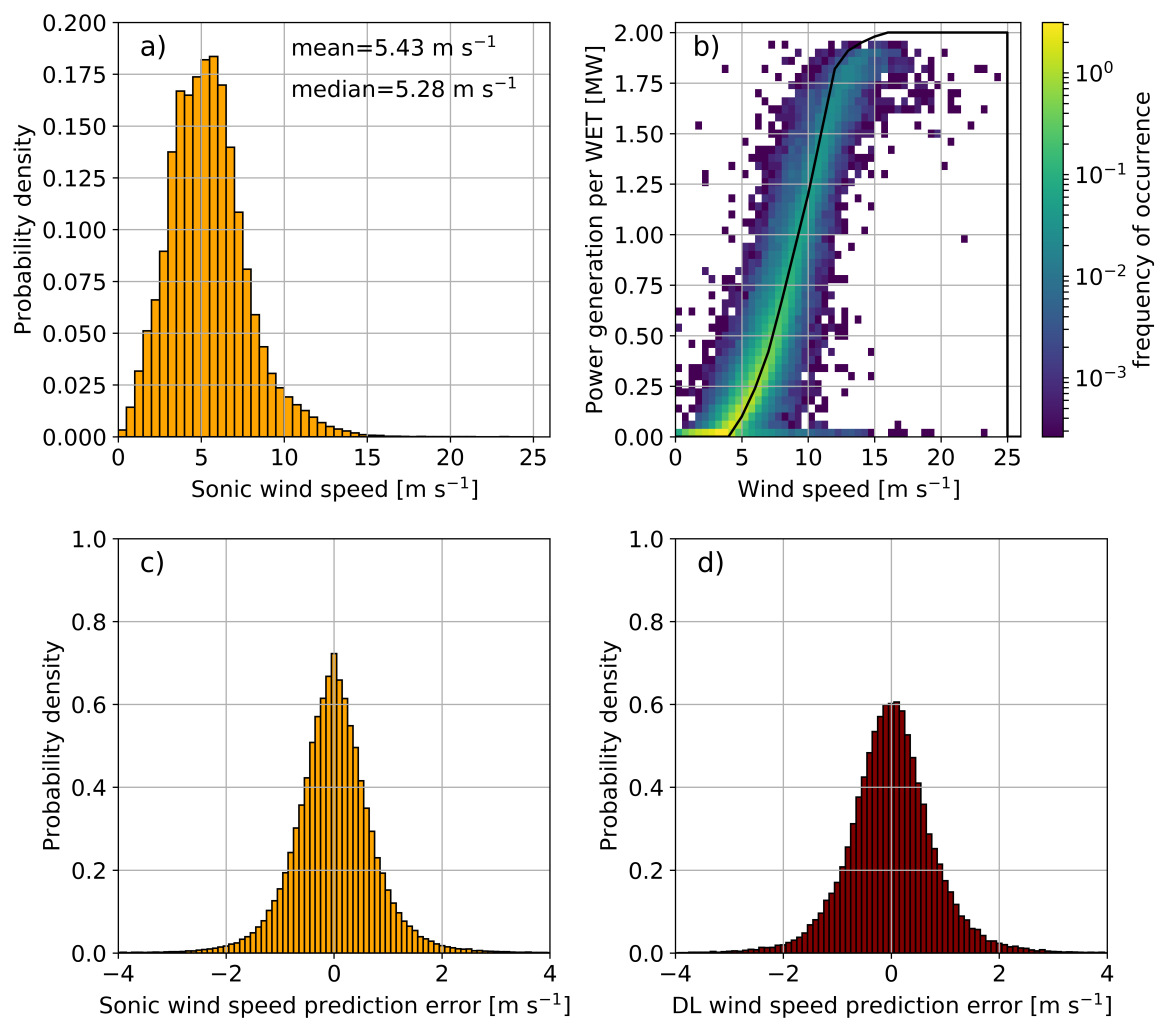
Figure 5.5: a) Histogram of wind speed measurements of nacelle ultrasonic anemometers, averaged over all five available locations. b) 2D histogram of wind speed measurements as a function of the power generation per wind energy turbine (WET). The solid black line represents the WETs power curve. c) Wind speed prediction error for the persistence prediction at 30 min in advance for nacelle ultrasonic measurements. d) Wind speed prediction error for the persistence prediction at 30 min in advance for Doppler lidar measurements.
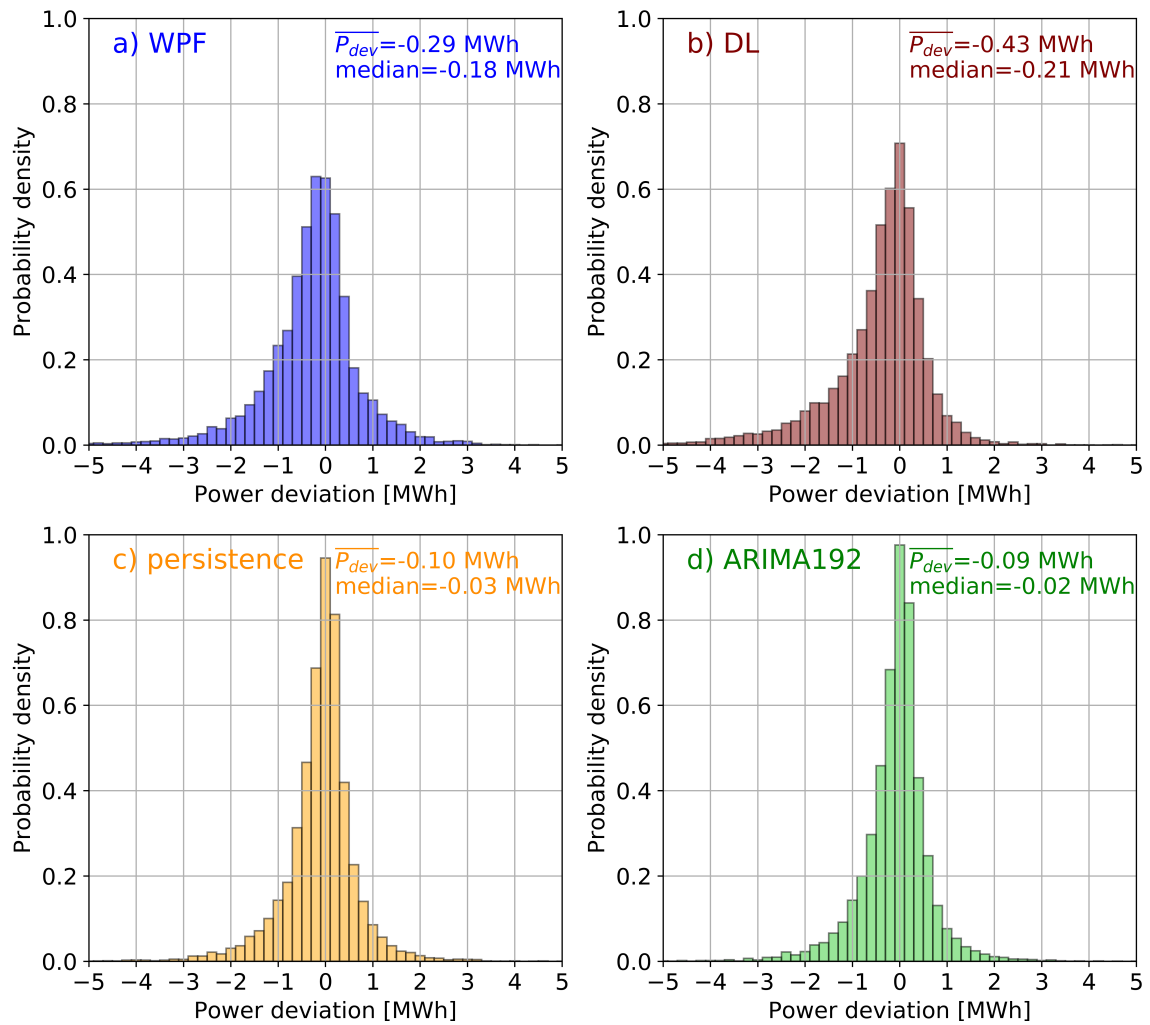
Figure 5.6: Histograms of power deviation P_{dev} for four wind power prediction configurations of the nine month prediction period (1 January 2019 to 30 September 2019) based on different wind velocity data: a) reference WPF, b) Doppler lidar data (30 min ahead), c) nacelle ultrasonic anemometer observations with persistence approach (15 min ahead), d) nacelle ultrasonic anemometer observations with ARIMA192 approach (15 min ahead). Mean and median values are indicated in the Figure.

5.2.2 Power deviations for the entire data set

Figure 5.7 shows the absolute power deviations $|P_{dev}|$ for the entire 45 months data set, divided into overestimations and underestimations. This analysis is shown for $+15\,\text{min}$ and $+30\,\text{min}$ prediction steps and for different wind power prediction configurations. The currently operational WPF configuration (blue boxes) is shown as a reference value. The wind power prediction based on DL data (brown) is only shown for $+30\,\text{min}$ as the data were averaged on that temporal resolution. Power deviations of the sonic-based persistence power predictions (orange) are shown for both forecasting horizons of $15\,\text{min}$ and $30\,\text{min}$. For the sonic-based ARIMA power prediction (green) at $15\,\text{min}$ ahead $|P_{dev}|$ is shown for all six ARIMA setups. For the sonic-based ARIMA power prediction at $30\,\text{min}$ ahead (also green) $|P_{dev}|$ is shown for the three most promising ARIMA setups ARIMA288, ARIMA480 and ARIMA960. Subsequent plots in this chapter comparing different wind power predictions will show the same combination of configurations.

The mean value of the absolute power deviations, i.e., the mean absolute error, serves as a proxy to compare different wind power prediction configurations. To evaluate the power prediction improvement by replacing NWP wind velocity data, the difference between the observation-based prediction and the reference WPF prediction is important (see white labels in Fig. 5.7). The white labels show the difference to the reference prediction first for the overestimations and then for the underestimations. In the case of power predictions, absolute improvements are more meaningful than relative values, because a deviation of, e.g., $\pm\,2\,\text{MWh}$ always has the same real world implications for balancing efforts, irrespective of whether the total power generation is, e.g., $5\,\text{MWh}$ or $25\,\text{MWh}$.

In addition, the difference in frequency of occurrence between under- and overestimations has to be taken into account. As stated above and shown in Fig. 5.6, for the reference WPF-based power prediction and for the DL-based power prediction underestimations are much more frequent than overestimations. For the WPF and the DL power predictions, $65\,\%$ and $67\,\%$ of the time steps show underestimations of $P_{dev}<$-$0.1\,\text{MWh}$, respectively. Additionally, these distributions are slightly skewed toward negative values (see Fig. 5.6). This shape of the WPF power prediction P_{dev} distribution causes its mean overestimation to be absolutely lower than the mean underestimation. For the sonic-based ARIMA and persistence power predictions, P_{dev} is more symmetrically distributed around $P_{dev}=0\,\text{MWh}$. Underestimations are only slightly more frequent than overestimations, with a share of $57\,\%$ (persistence) and $56\,\%$ (ARIMA192) of the time steps having underestimations of $P_{dev}<$-$0.1\,\text{MWh}$. Therefore, mean absolute underestimations and overestimations are expected to be more similar for the sonic-based power predictions. However, since the mean overestimation of the WPF power prediction is absolutely lower than the mean underestimation (Fig. 5.6), the improvement by the observation-based power prediction approaches is less for overestimations than for underestimations.

Figure 5.7 shows that generally, power underestimations (plain boxes) have a wider distribution than overestimations (hatched boxes). In addition, underestimations feature on average higher $|P_{dev}|$ values than overestimations. The improvement (white labels) during power underestimations compared to the reference run is higher than for power overestimations for all power predictions setups. This agrees well with the abovementioned expectations based on Fig. 5.6. Since the ARIMA modeled wind speeds are very close to the persistence shifted wind speeds, it is expected that the wind power predictions based on both of these configurations are also similar. For the short prediction times examined here, all predictions based on nacelle sonic observational data result in more accurate power forecasts than those performed with the WPF model input. The wind power predictions based on DL data
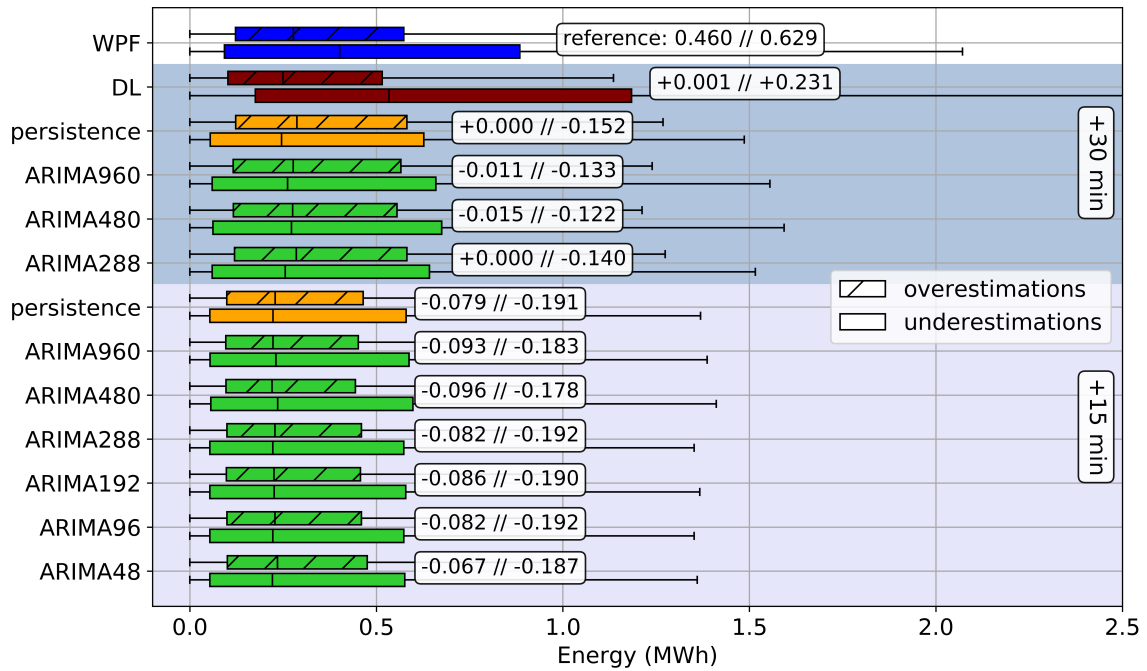
Figure 5.7: Differences between wind power generation and differently configured predictions for all time steps, divided into overestimations (hatched) and underestimations (plain). The boxes represent data from the first to third quartile, the whiskers data from the 10th to the 90th percentile. The black vertical line is the median. The blue bar refers to the reference WPF configuration, brown the DL configuration, orange bars the persistence configurations and green bars depict ARIMA configurations (details in Sect. 4.4). The white labels show the mean power deviation value of the reference WPF prediction in MWh and, subsequently, the difference in mean power deviation to the reference WPF for each power prediction model configuration, first for the overestimations and then for the underestimations.

achieve approximately the same mean power deviation for power overestimation scenarios as the WPF-based reference while during underestimation scenarios DL-based predictions are less good.

As expected, power deviations for configurations, for which both prediction times are presented, are generally lower for the shorter forecasting horizon of 15 min than for 30 min for both under- and overestimations. For 30 min in advance, all power predictions show no or only very small improvements compared to the WPF prediction during overestimations. Contrarily, the improvements during underestimations (except for the DL-based power prediction), ranging from 0.12 MWh to 0.15 MWh, are significant compared to the mean power deviation of the reference run of 0.629 MWh. The sonic-based persistence configuration is slightly better than the one based on ARIMA, reducing the mean power deviation during underestimations by about 0.15 MWh.

For the forecasting horizon of 15 min, most ARIMA configurations perform slightly better than the persistence prediction. All of them show improvements compared to the reference power prediction, reducing the mean power deviation per time step by 0.07 MWh to 0.1 MWh for overestimations and by 0.18 MWh to 0.19 MWh for underestimations. In contrast to the 30 min ahead predictions, for 15 min in advance there is also an improvement during overestimations, though still smaller than during underestimations. Even though the ARIMA wind power predictions result in slightly larger absolute deviations than the sonic persistence

power prediction, this difference is not significant. However, the improvement compared to the WPF-based reference run is clearly significant for all wind power prediction configurations based on observational data, except for the DL prediction.

Although the prediction based on DL data provides also information on wind velocity profiles (see Sect. 4.4), it yields larger deviations than the WPF-based reference run. Consequently, data from a vertically pointing DL located 6 km to the WNW of the wind farm do not provide enough additional information on the wind field to result in an improvement of the wind power prediction in this study, which is in contrast to studies investigating horizontally pointing Doppler lidars and radars (*Pichault et al.*, 2021a,b; *Valldecabres et al.*, 2020), which were exclusively installed for wind power predictions. In addition, the larger distance between the observatory, where the DL is installed, and the wind farm compared to the above-mentioned studies, likely contributes to higher deviations. This result suggests, that the vicinity of the DL to the wind farm is crucial. However, including the DL observations from MOLRAO was done as a test to study the usefulness of DL observations performed routinely there for other purposes. Moving the vertically-pointing DL into the wind farm was thus not a possibility but could be investigated in a future study.

In an additional analysis, the DL power deviations were evaluated only for the wind sector where the DL is strictly upwind of the wind farm. In theory, the persistence approach should be most valuable for this wind sector. The distance between the DL and the wind farm area ranges approximately from 6 km to 8 km. Thus, for an exemplified wind speed of $7\,\mathrm{m\,s^{-1}}$ it would take around 15 min to 20 min from the observatory to the wind farm. In this sector of Western to Northwestern wind directions, the power deviations of the DL power prediction are comparable to the power deviations of the sonic-based approaches. The narrower this wind direction sector is chosen, the more the power deviation of the DL power prediction decreases. However, eliminating other wind directions is not an option for the potential operational application of this approach, because wind power predictions are needed for all wind directions. Therefore, the comparison of the different wind power predictions shows the DL power deviations for all wind directions. Summarising, as shown in Fig. 5.7, the DL power deviations do not show an improvement compared to the reference.

5.2.3 Power deviations for ramp time steps

Figure 5.8 shows the histograms of $|P_{dev}|$ for ramp time steps only, including two time steps before and after the ramp event. Ramp events have been identified by the Minimum-maximum method with power threshold $\Delta P_r{=}30\,\%$ and window length $WL{=}120\,\mathrm{min}$ as explained in Sect. 4.5. Compared to the histograms of P_{dev} for the entire data set (Fig. 5.6), the distributions are wider, because the frequency of occurrence of larger absolute power deviations is higher during ramp events. In addition, for most of the shown examples $\overline{P_{dev}}$ deviates further from 0 MWh during ramp time steps than during all time steps, creating a larger bias towards underestimations. This bias is larger for the WPF and DL configurations than for the persistence or ARIMA power predictions. For down-ramps, $\overline{P_{dev}}$ of the reference WPF prediction is about half of the $\overline{P_{dev}}$ value of the DL configuration. Compared to the sonic-based power predictions, the reference $\overline{P_{dev}}$ is 70 % higher than for the persistence prediction and approximately 160 % higher than the ARIMA192 prediction. During up-ramp events, $\overline{P_{dev}}$ is closer to 0 MWh than during down-ramp events for all configurations. The reference WPF setup yields $\overline{P_{dev}}{=}$ -0.45 MWh during up-ramps, which is only slightly higher than for the entire data set ($\overline{P_{dev}}{=}$ -0.29 MWh). The mean power deviation for the DL setup is about twice as high than for the reference. The sonic-based power predictions result in positive $\overline{P_{dev}}$ values, which are in their absolute values approximately 75-80 % lower than the reference.
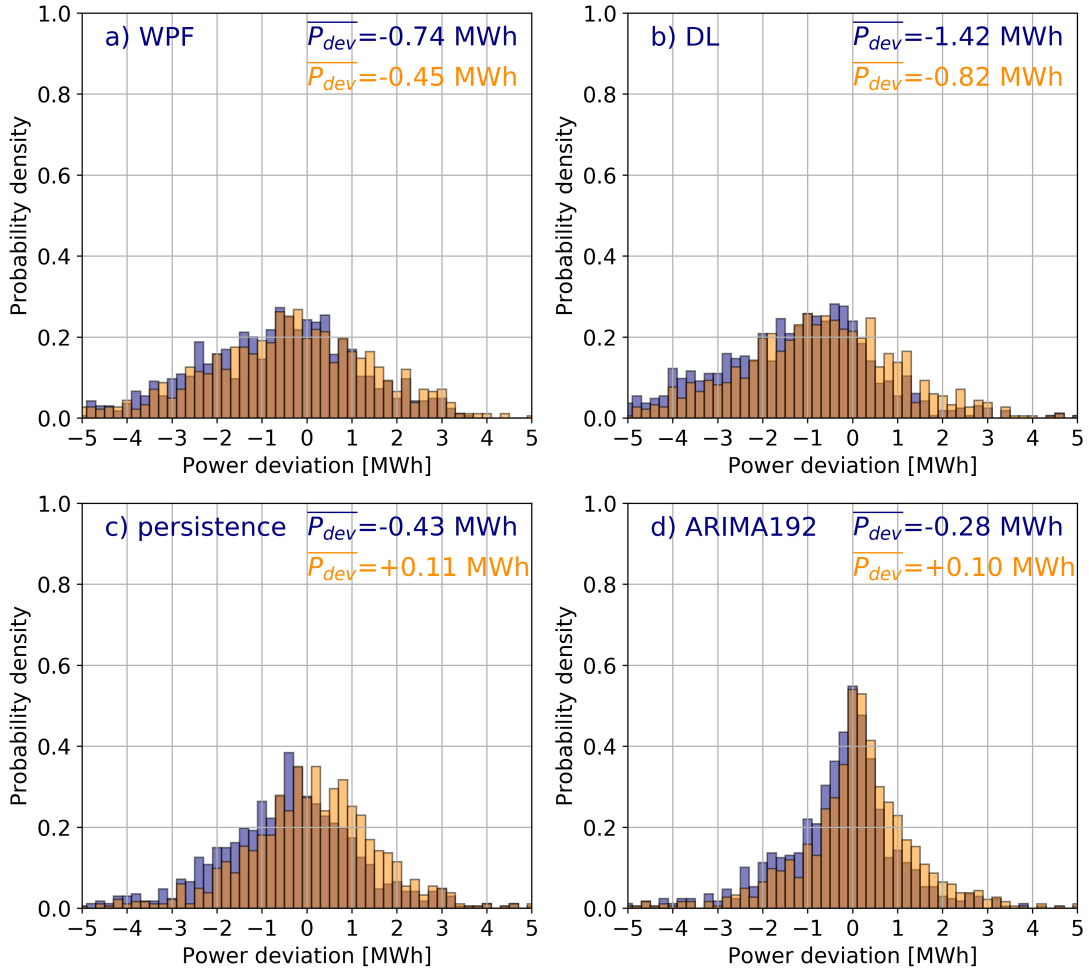
Figure 5.8: Histograms of power deviation P_{dev} during down-ramp (blue) and up-ramp events (orange) for four wind power prediction configurations based on different wind velocity data: a) reference WPF, b) Doppler lidar data (30 min ahead), c) nacelle ultrasonic anemometer observations with persistence approach (15 min ahead), d) nacelle ultrasonic anemometer observations with ARIMA192 approach (15 min ahead).

In addition, $\overline{P_{dev}}$ for the persistence and ARIMA configurations are in their amount similar to their $\overline{P_{dev}}$ values for the entire data set, albeit with the opposite sign and a less narrow distribution around 0 MWh.

Since the number of underestimations is significantly higher than the number of overestimations for most configurations, the direct comparison of their mean absolute errors can be misleading. In this section, the performance of the different power predictions setups depending on the sign of the ramp event is evaluated by the mean absolute deviation of over- and underestimations combined.

Figure 5.9 shows the absolute deviations per ramp time step divided into up-ramps (hatched) and down-ramps (plain) as introduced in Fig. 5.7 for the entire data set. In general, $|P_{dev}|$ is approximately three times higher during ramp events. The reference WPF power prediction is the only configuration, which shows a higher mean absolute deviation for down-ramp than for up-ramp events. Contrarily, all power predictions based on observational wind speed data show slightly lower P_{dev} for down-ramps than for up-ramps. Therefore, the improvement

Table 5.2: Mean Doppler Lidar (DL) power deviations during ramp events for different wind direction sectors. The DL is located at an angle of 285° relative to the wind farm, the wind direction sectors in this Table are centered around this angle.

power prediction	wind direction sector	$\overline{P_{dev}}$ up-ramps [MWh]	$\overline{P_{dev}}$ down-ramps [MWh]
WPF	0°-360°	1.492	1.523
DL	0°-360°	1.642	1.299
DL	195°-15°	1.745	1.358
DL	240°-330°	1.665	1.385
DL	270°-300°	1.569	1.287
DL	280°-290°	1.499	1.241

during down-ramp events is for all configurations larger than during up-ramps. Similar to considering all time steps, the mean absolute power deviation for ramp events using the ARIMA configurations is pretty close to the mean absolute power deviation of the persistence wind power prediction. However, during ramp events of both directions, the persistence predictions outperform all ARIMA configurations slightly.

For a forecasting horizon of 30 min, the sonic-based power predictions show only a minor improvement of 0.01 MWh-0.09 MWh compared to the reference. The DL power prediction during up-ramps, however, performs worse than the reference run. For a forecasting horizon of 15 min, the improvements range from 0.16 MWh to 0.21 MWh during up-ramps and from 0.25 MWh to 0.30 MWh during down-ramps. That means, for 15 min in advance the improvement during ramp events is up to twice as high as during all time steps. Therefore, the value of observational data in the wind power prediction increases during ramp events. For the DL-based power prediction, the mean absolute deviation is better than the WPF-based reference value during down-ramp events, but less good during up-ramp events. Evaluating DL absolute power deviations only for Western to Northwestern wind directions decreases them to values comparable to the sonic-based approaches with the same forecasting horizon, similar to the evaluation for the entire data set. Table 5.2 shows the mean power deviation for the DL configuration for different wind direction sectors centered around the line of sight between the DL and the wind farm. Only when analysing the 10° and 30° sectors around the line of sight (280°-290° and 270°-300°, respectively), an improvement compared to the DL power prediction for all wind directions is shown. During down-ramp events, for all chosen wind direction sectors the reduction of $\overline{P_{dev}}$ compared to the WPF reference is evident. However, for up-ramp events, no wind direction sector results in a lower $\overline{P_{dev}}$ of the DL setup compared to the reference, albeit for the narrowest sector of 280°-290° $\overline{P_{dev}}$ of both configurations is close.

The results in this section are shown for one ramp threshold definition (WL=120 min, ΔP_r=30 %). For other ramp threshold definitions, the improvement during ramp time steps is smaller than shown here. For WL=30 min, none of the power predictions configurations is able to provide better power predictions than the reference. This is caused by the nature of the persistence approach. For very steep ramp events the assumption that the wind speed remains the same is not very plausible. Although the ARIMA approach uses lagged observations as predictors, it is evidently also not able to accurately predict large changes for a short time range of only two time steps. For WL=60 min, the sonic-based power predictions do provide an improvement over the reference run for ΔP_r=30 % and ΔP_r=40 %, albeit of smaller absolute values than for WL=120 min and ΔP_r=30 %. Power predictions for ramp events with ΔP_r>40 % are not improved by any power prediction configuration.
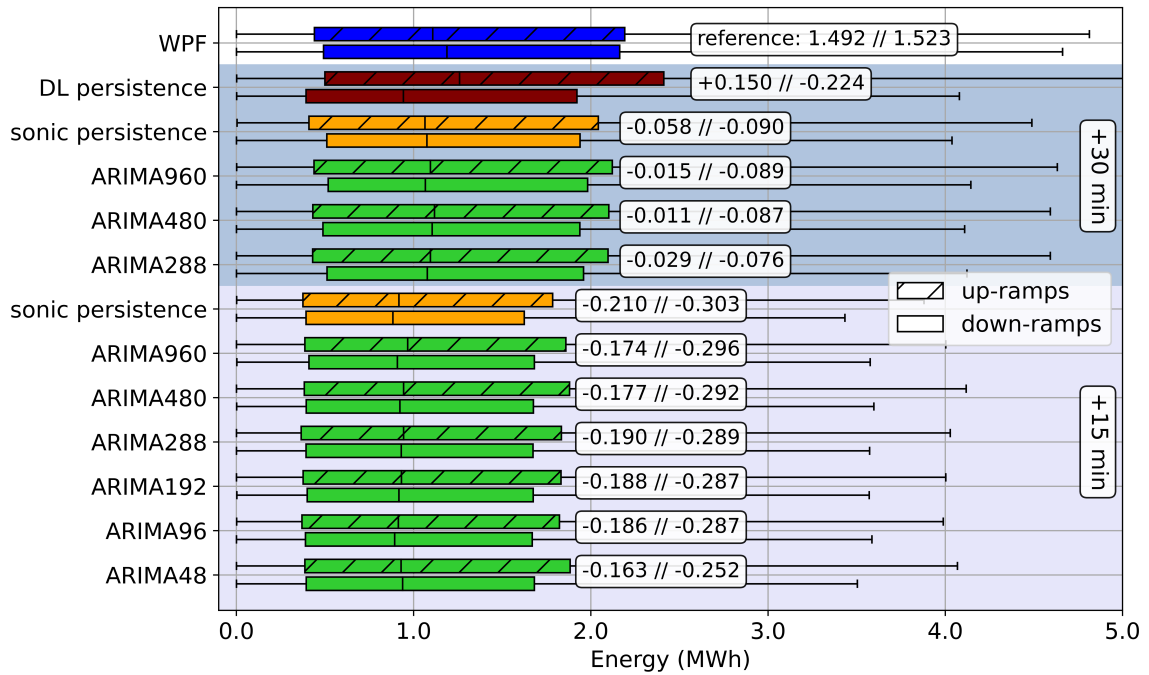
Figure 5.9: Differences between wind power generation and differently configured predictions P_{dev} for ramp time steps and two time steps before and afterwards. The boxes represent data from the first to third quartile, the whiskers data from the 10th to the 90th percentile. The black vertical line is the median. Green bars are ARIMA configurations, orange bars are persistence configurations, red the DL configuration and blue the reference WPF configuration (details in Sect. 4.4). The data are divided into P_{dev} during up-ramp events (hatched) and down-ramp events (plain). The white labels show the difference in mean deviation $\overline{P_{dev}}$ between each power prediction model configuration and the reference WPF prediction in MWh for up-ramps and for down-ramps, respectively.

However, power predictions during moderately steep ramp events are improved by the observation-based power predictions as discussed in this section. Very steep ramp events are rare at the wind farm Beeskow (see Tab. 5.1) and the lacking improvement for these time steps does not affect the overall performance of the observations based power predictions, as shown in Sect. 5.2.2.

5.2.4 Summary

For the whole data set (Sect. 5.2.2) the mean absolute error between wind power prediction and observed power generation for the reference WPF-based wind power prediction amounts to 0.46 MWh during time steps with overestimations and 0.63 MWh during time steps with underestimations. The higher value for underestimations is caused by the bias of the LoadManager®'s power predictions towards underestimations, which is illustrated in Fig. 5.6. Furthermore, it can be concluded that all power predictions based on nacelle sonic anemometer data show lower $|P_{dev}|$ than the reference WPF-based wind power prediction. Generally, this improvement is slightly larger for the shorter forecasting horizon of 15 min. Furthermore, the improvement is larger for underestimations, which is a side-effect of the P_{dev} distributions for sonic based power predictions being centered closer to 0 MWh. However, even for overestimations the sonic-based power predictions show lower values of the mean absolute error than the reference for the forecasting horizon of 15 min and similar to refer-

ence values for the forecasting horizon of 30 min. The power prediction based on DL data shows higher values of the mean absolute deviation than the reference WPF-based power prediction for both under- and overestimations for the entire data set. The power prediction configurations with the best performance for the whole data set are ARIMA96, ARIMA192 and ARIMA288 for the forecasting horizon of 15 min ahead and the sonic based persistence approach for the forecasting horizon of 30 min. However, the difference between both sonic-based approaches is not statistically significant.

When evaluating only power predictions during ramp events, power deviations are approximately three times as large as during all time steps. Since the bias towards underestimations is stronger in this sub set of the data set, the analysis is not divided into under- and overestimations but only into up- and down-ramp events. For the reference WPF-based power prediction, the mean absolute deviation ranges from 1.49 MWh for up-ramp events to 1.52 MWh for down-ramp events. For all power prediction configurations based on observational data, $|P_{dev}|$ is on average slightly lower during down-ramps than during up-ramps. For both forecasting horizons the sonic-based persistence power prediction is the best performing setup. The improvement compared to the WPF-based power prediction reference during ramp time steps is 60-100 % higher than for all time steps, highlighting the usefulness of observation-based power prediction especially during ramp events. The DL power predictions do show lower mean absolute errors than the WPF run during down-ramp events, but not during up-ramp events. Although the prediction based on DL data provides also information on wind velocity profiles (see Sect. 4.4), it yields mostly larger deviations than the reference run. This is presumably due to the distance between the observatory, where the DL is installed, and the wind farm, which cannot be compensated by the higher amount of input data. Filtering the data set of the DL-based configuration to wind directions for which the DL is directly upwind of the wind farm Beeskow reduces the mean absolute deviation to similar values as the sonic-based approaches for the same forecasting horizon. However, this is not further analysed here, because the potential operational application of this approach would warrant wind power predictions for all wind directions.

5.3 Cost-oriented analysis

To estimate the monetary value of the improvement of wind power predictions due to a certain instrument, the deviations of the different power prediction configurations P_{dev} evaluated in Sect. 5.2 will be analysed in conjunction with the balancing cost data set. As explained in Sect. 2.3.3, balancing costs are not the only factor affecting variable costs for wind farm operators or directs marketers. However, it is a good way to estimate the financial benefit of better wind power predictions. The data set SMARD (*SMARD*, 2021b), provided by the *Bundesnetzagentur*, includes the German wide uniform balancing price for every 15 min time step. With the power deviation in MWh and the balancing price in € MWh^{-1} for each time step, the balancing cost in € can be calculated.

Figure 5.10 shows the costs for balancing energy per time step for the same wind power prediction configurations shown previously. Approximately 10 000 time steps between 1 January 2019 and 23 September 2019 have been evaluated, choosing only those time steps where power predictions for every configuration are available. The mean balancing cost per time step is shown in the white labels. Negative balancing costs represent refunds by the transmission system operators (TSO), positive balancing costs mean necessary payments to the TSO for the provided balancing services.
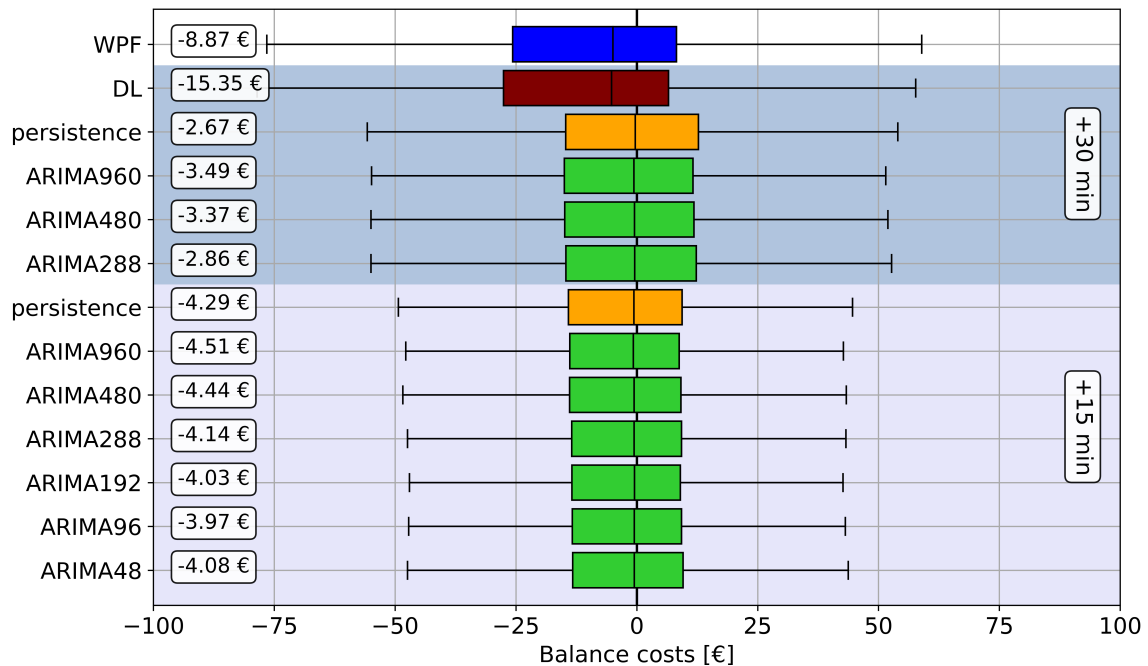
Figure 5.10: Costs for balancing energy per time step. The boxes represent data from the first to third quartile, the whiskers data from the 10th to the 90th percentile. The black vertical line is the median. Blue boxes are the reference WPF setup, brown the DL configuration, orange boxes refer to the persistence configurations and green boxes are ARIMA configurations (details in Sect. 4.4). The white labels show the mean balancing cost per time step for each wind power prediction configuration.

Surprisingly, all configurations show on average negative balancing costs. That means, regardless which power prediction setup the responsible wind farm operator or direct marketer would apply, they would on average not need to pay for balancing services and instead get money back, which is not very realistic. This effect is caused by (i) the small size (low amount of installed capacity) of the wind farm and by (ii) a systematical underestimation of the power generation due to the nonlinear relation between wind speed and power generation (see Sect. 5.2.1) and will be discussed in the next section. This systematic underestimation will be bias-corrected in Sect. 5.3.2 to enable a more meaningful estimation of balancing costs.

5.3.1 Reasons for negative balancing costs

The first reason for negative balancing costs is that the sign of the analysed wind farm's power deviation is independent of the nationwide grid situation. This is caused mainly by the relatively small size of the wind farm. For a larger balancing group, consisting of multiple wind farms, random errors of wind power predictions for individual locations would cancel out each other. Only systematical errors, due to widespread lower or higher wind speed than predicted or the delay of a large-scale frontal passage, which would affect all wind farms simultaneously, would remain. Assuming, as a thought experiment, that weather forecast errors cause widespread lower than predicted wind speed in Central Germany. For a large balancing group consisting of multiple wind farms at different locations in Central Germany, this weather forecast error causes an overestimation of the wind power prediction. Simultaneously, other wind farms in Central Germany belonging to other balancing groups are affected by the unexpected less windy conditions as well. Thus, they would also generate
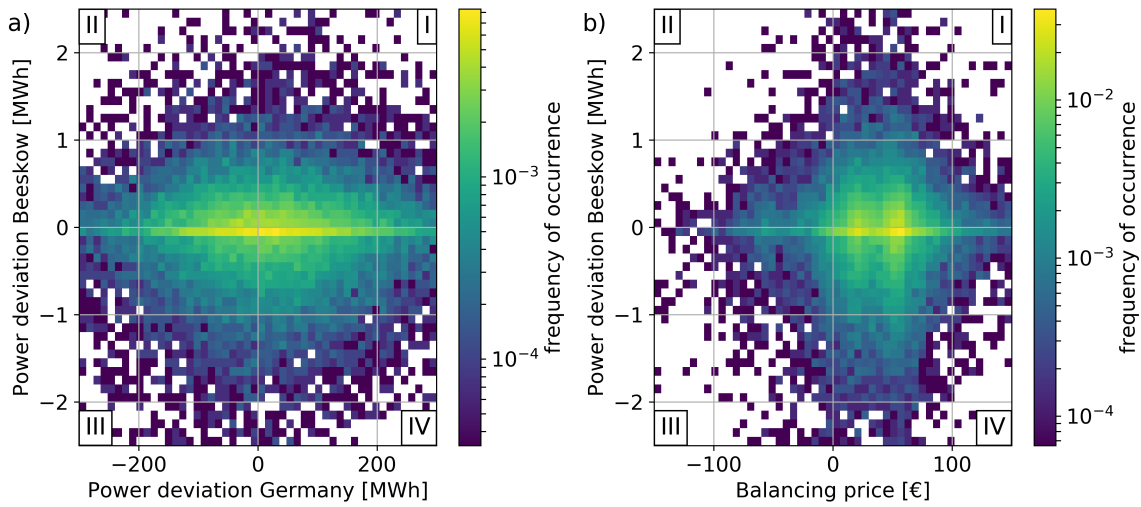
Figure 5.11: a) Power deviation P_{dev} of the wind farm Beeskow in comparison to P_{dev} of the whole German electricity grid. b) P_{dev} of the wind farm Beeskow in comparison to the nationwide balancing price. Roman numerals label the four quadrants.

less power than expected. If a large share of wind farms feed less power than expected into the otherwise unperturbed nationwide grid, the grid would then be undersupplied, too. Therefore, weather forecast errors causing power deviations of either direction for a large balancing group likely affect the whole grid in a similar manner as this exemplary large balancing group. Although not all imbalances in the German power grid are caused by deviating wind power predictions as this consideration implies, power generation by renewable energies and their variations are an important contribution.

Figure 5.11a shows P_{dev} of the wind farm Beeskow in comparison with the nationwide power deviation, to check if overestimations (underestimations) in the wind power prediction for Beeskow correlate with an undersupplied (oversupplied) national electrical grid. This is the case, when the power deviations have the same sign, i.e., the majority of the data points would be in the first (I: overestimation for Beeskow and undersupplied German grid) or third quadrant (III: underestimation for Beeskow and oversupplied German grid). However, Fig. 5.11a does not show such a distribution, instead the data points are distributed symmetrical around the axes. Therefore, there is no clear correlation between over- and underestimations for the wind farm Beeskow and the imbalance status of the national electrical grid. As the wind farm Beeskow is rather small and located near the Eastern border of Germany, random errors in the wind power prediction, which affect only the power generation of this wind farm, cause P_{dev} to be independent of the German wide grid status.

The second reason for on average positive balancing costs illustrated in Fig. 5.10 is the systematical underestimation of the generated wind power due to the non-linear relationship between wind speed and generated power. Figure 2.5 introduced the balancing price between 2017 and 2019 in Germany in comparison to the balancing volume of the German electrical grid. The balancing price depends to an extent on the nationwide electrical grid situation. However, P_{dev} for the wind farm Beeskow is largely independent on the nationwide situation, as shown above. Therefore, the 2D histogram of balancing prices against the balancing volume is reduced to a 1D histogram in Fig. 5.12, showing only the distribution of balancing prices between 2017 and 2019 in Germany. This distribution shows the bias towards positive prices more clearly. Even during situations with a power deficit in the grid, balancing prices are often
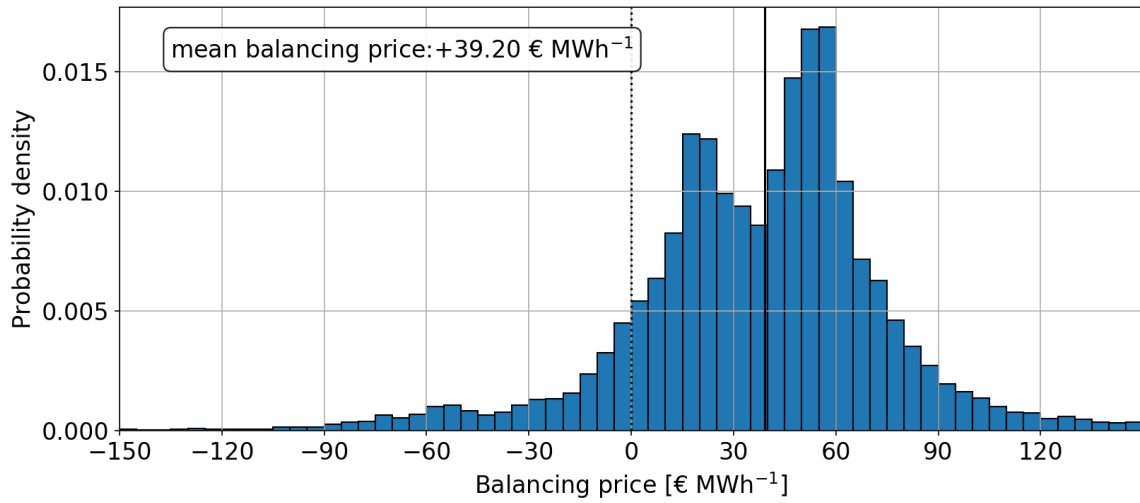
Figure 5.12: Histogram of balancing prices from 1 January 2017 to 31 December 2019. The mean balancing price is indicated by the black vertical line.

positive. The average balancing price is approximately $39 \, \text{€} \, \text{MWh}^{-1}$. During phases with positive balancing prices, excess power can be sold. The more frequent underestimation of the wind power generation by the LoadManager® means that the wind farm more frequently generates situations with excess power than with power deficits. In combination with the on average positive balancing price this causes on average negative balancing costs as illustrated in Fig. 5.10.

As shown in Sect. 5.2, the amount of mean underestimation $\overline{P_{dev}}$ decreases for power prediction configurations based on observational data. However, since the power deviations on average cause refunds of varying amounts (see Fig. 5.10), the reduction of $\overline{P_{dev}}$ in turn reduces the amount refunded. Therefore, a power prediction with reduced power deviations would financially be disadvantageous. This effect is mainly caused by the bias of the power prediction tool towards underestimations, which has to be eliminated for future analyses and future operation of the LoadManager®.

5.3.2 Bias-correction of the wind power predictions

To still evaluate balancing costs in the scope of this project, the bias is adjusted in a post-processing step. For each power prediction configuration, the mean power deviation $\overline{P_{dev}}$ is calculated. This value is $< 0 \, \text{MWh}$ for all power predictions considered in this thesis, ranging from -0.43 MWh for the DL configuration up to -0.07 MWh for the ARIMA288 configuration. The mean power deviation is then subtracted from the power prediction time series, unless the power prediction forecast a power generation equal to the wind farms maximum capacity. Thus, the original distributions (blue) in Fig. 5.6 are shifted by the mean power deviation to the right (orange) in Fig. 5.13.

In addition, the moderate total capacity of the wind farm Beeskow means that it is prone to large outliers in balancing costs, both, towards high costs and high refunds. When considering financial risk and reward aspects for this small balancing group, the outliers absolutely have to be taken into account. However, if the evaluated time period is not long enough, they affect the statistical analysis considerably. Theoretically, these outliers should balance out for a sufficiently long time period. Therefore, they are temporarily excluded for a statistical
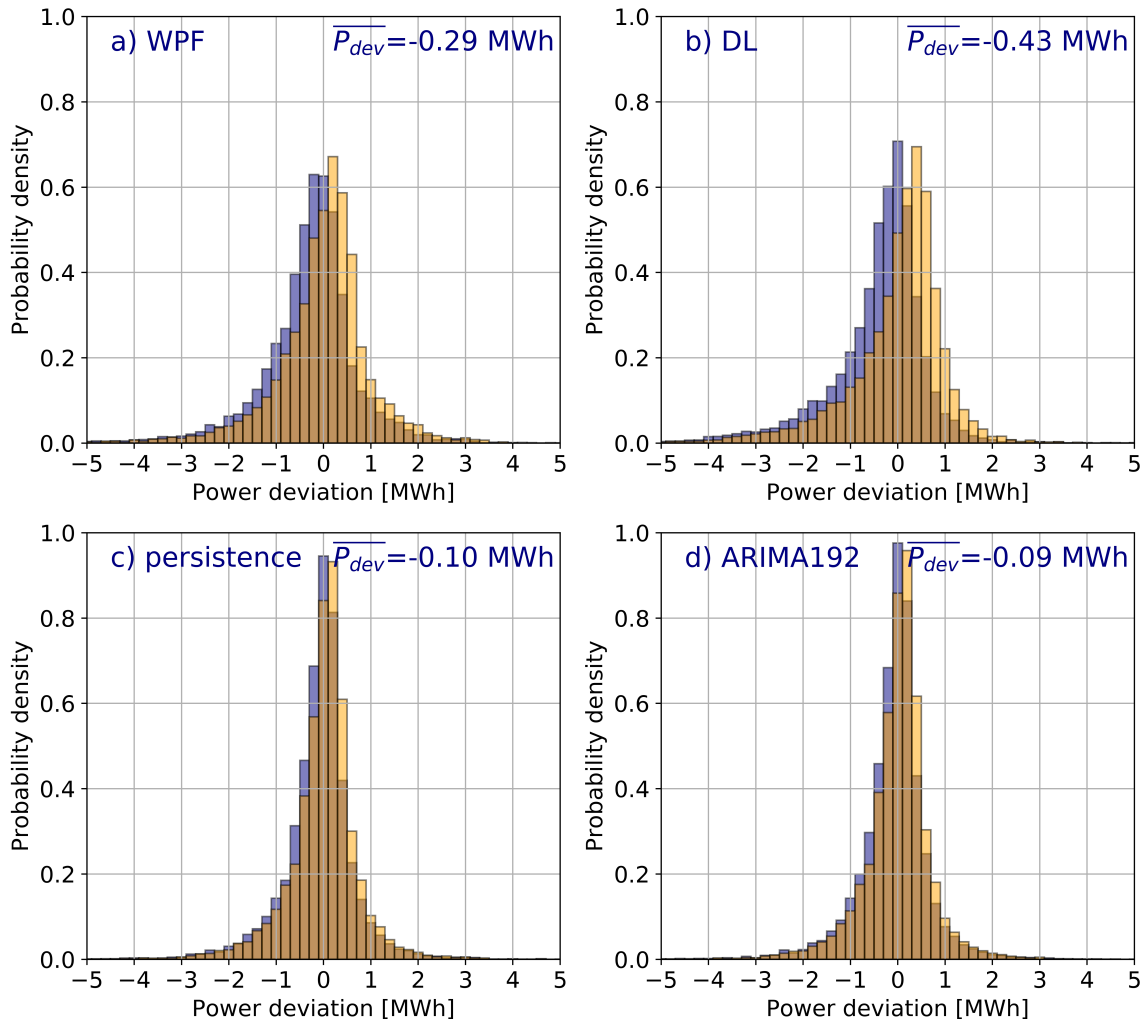
Figure 5.13: Histograms of power deviation P_{dev} for four wind power prediction configurations based on different wind velocity data: a) reference WPF, b) Doppler lidar data (30 min ahead), c) nacelle ultrasonic anemometer observations with persistence approach (15 min ahead), d) nacelle ultrasonic anemometer observations with ARIMA192 approach (15 min ahead). Blue bars show the original histograms, the orange bars show the distributions shifted by their mean value.

analysis. Excluding time steps with very high balance costs above $300\,€$ eliminates about 90 time steps for the total time period of approximately $10\,000$ time steps.

Figure 5.14 shows the same analysis of balancing costs between the different power predictions as Fig. 5.10, but with the correction for the power deviation bias and the removal of large outliers. The median balancing costs for these adjusted inputs are positive - meaning on average the wind farm operator actually has to pay for balancing services - which seems to be more reasonable. Additionally, the mean balancing costs for each power prediction configuration can be divided into three groups. The first group includes the configurations with the highest power deviations, WPF and DL, which exhibit the highest average of positive balancing costs. The second group, persistence and ARIMA power predictions for $+30\,$min, features lower power deviations and, thus, on average a slightly lower spread in the balancing costs. The third group is comprised of power predictions with a forecasting horizon of $15\,$min, which show on average the lowest balancing costs. For the first group, the mean
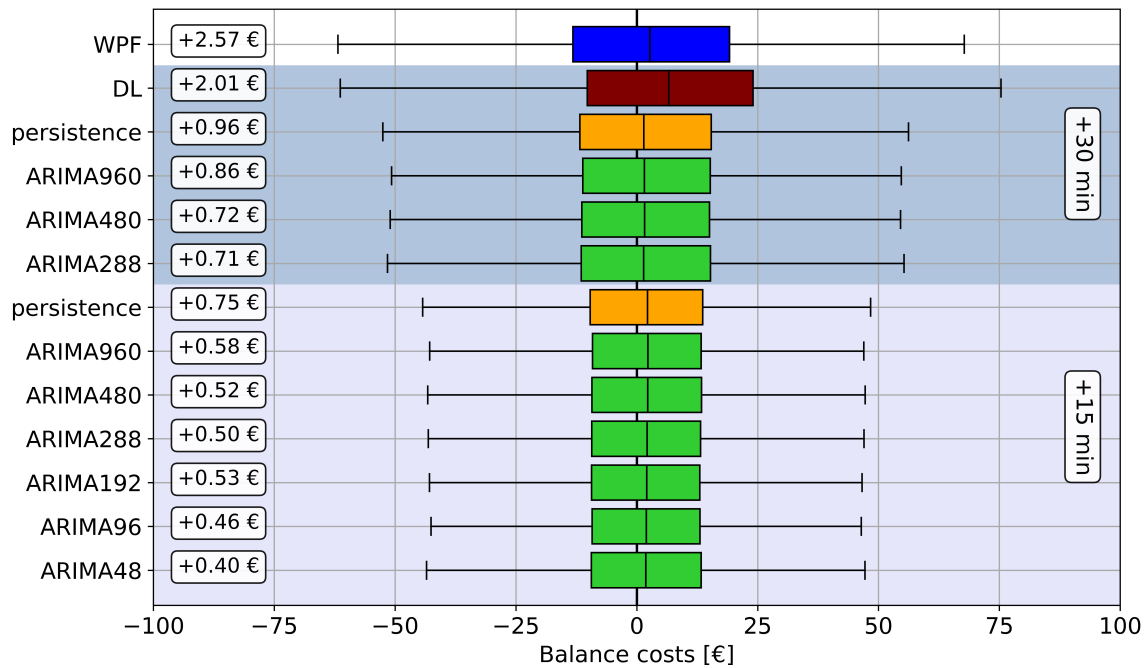
Figure 5.14: Same analysis as Fig. 5.10, but for the adjusted balancing costs, as described in Sect. 5.3.2.

balancing costs for the DL power prediction are lower than for the WPF reference run, because of a few more moderately high outliers towards negative balancing costs in the DL configuration compared to the WPF configuration. Average balancing costs for the second group (in Fig. 5.14 referring to moderate balancing costs) are 1-2 € lower than for WPF and DL configurations, and range from 0.71 € to 0.96 € per time step. The three shown ARIMA configurations on average perform better than the persistence power prediction, resulting in lower mean balancing costs. The average balancing costs for the third group - the +15 min power predictions - are up to 0.31 € lower than for the second group, ranging from 0.40 € to 0.75 € per time step. This is expected, because the prediction time is shorter. In addition, a narrower spread of the balancing costs around the median can be seen for group three. With the aforementioned corrections applied, the balancing costs qualitatively agree significantly better to expectations in (i) that power predictions with higher power deviations also cause higher balancing costs and (ii) that balancing costs are slightly positive, meaning the balancing group operator has to pay a fee for the service.

To complete the cost-oriented analysis, another commonly used metric to evaluate balancing costs is introduced, namely the balancing costs in relation to the generated power. Depending on their specific business contract, some direct marketers bill the wind farm operators a flat rate for balancing costs depending on the amount of generated power. Contracts for small wind farms, where the variability of balancing costs and, therefore, the risk is higher, typically feature higher rates. Contracts for large or multiple wind farms feature lower rates. To give an estimate, for the power prediction period in 2019, this flat rate was typically in the range of 0.50 € per generated MWh. That means the higher the power generation of the wind farm, the higher the available budget for balancing costs. Consequently, if the direct marketers power schedules are sufficiently accurate to not use up the whole rate allotted for balancing costs, they earn extra profit. However, calculating this metric per time step is statistically not meaningful, as single time steps with almost no generated power dominate
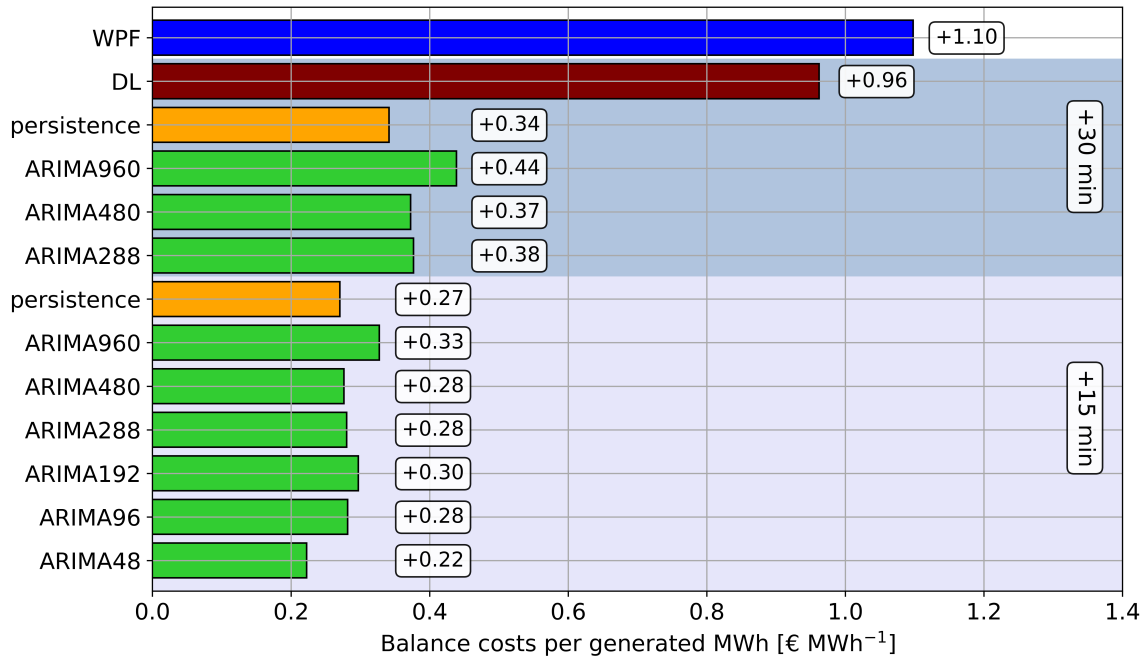
Figure 5.15: Balancing costs per total generated power for the power prediction period. The white labels show the exact value.

the statistics. Thus, Fig. 5.15 shows the total balancing costs of the power prediction period divided by the total generated power during this time. These data are adjusted for the bias and filtered for outliers ($>$300 €) per time step as well. Generally, they are on the same order as the typical flat rate of $0.50 \, € \, \mathrm{MWh}^{-1}$ allocated for balancing costs. When comparing the balancing cost per generated MWh, the reference WPF run and the DL run exceed this rate, while the other configurations, which are based on nacelle sonic observations, are lower than this rate. Again, the wind power predictions with a forecasting horizon of 15 min result in lower balancing costs per generated power than power predictions with a forecasting horizon of 30 min. Using any of the configurations based on nacelle sonic observations for their power predictions, balancing group operators would not exhaust the allotted example flat rate of $0.50 \, €$ per generated MWh^{-1} provided for balancing costs.

5.3.3 Cost-benefit calculation

This section briefly discusses a cost-benefit calculation for the acquisition of the used instruments. For this consideration, the values of the adjusted balancing costs per time step (Fig. 5.14) are used, as they are more realistic than the non-bias corrected costs. Nacelle sonic anemometers were found to provide the most benefit in terms of reducing power deviations and reducing balancing costs. As introduced in Sect. 3.2.3, anemometers are already required at every wind energy turbine, although they do not necessarily have to be ultrasonic. However, cup anemometers measure with similar accuracy as ultrasonic anemometers (e.g., *Sotelino et al.*, 2012; *WMO*, 2014) and, thus, are expected to provide similar results for power deviations and corresponding balancing costs. That means, no additional expenses incur for their acquisition and any financial benefit by using their observations results in a profit. Thus, using nacelle sonic observations to improve wind power prediction is the most prominent choice for the balancing group operator. Furthermore, for the application of this

approach to operational use, even disadvantages like wake effects as explained in Sect. 4.1.3 can be mitigated.

If no nacelle sonic observations are made available to the balancing group operator and they aim to improve their power predictions by the means of acquiring a DL instrument, its benefits have to outweigh the costs. Although the DL power prediction yields higher average balancing costs per time step than the nacelle sonic configurations, they are still 0.56 € or roughly 20 % lower than for the reference WPF run. The DL used in this thesis is a HALO Photonics Streamline Doppler lidar, introduced in Sect. 3.2.2, with initial costs of 276 000 €. Amortising this instrument with an average benefit of 0.56 € per time step would take approximately 5134 days, which translates to just over 14 years. This time probably exceeds the typical life span of a DL. However, the Streamline DL is a highly capable instrument for scientific purposes. Even a less powerful and, therefore, cheaper DL instrument like the Leosphere Windcube can provide competitive measurements in the typical height ranges of wind energy turbines (see Sect. 4.2.3). The current Leosphere Windcube v2.1 is priced at 100 000 €. Assuming similar performance of the DL power prediction with Windcube input data, the amortisation time is reduced to 1860 days (approximately 5 years), which is in the typical life span of DL instruments.

On the one hand, the acquisition costs do not represent the total costs of setting up the data processing flow. Although DLs usually do not require costly maintenance, additional expenses would include, e.g., personnel costs for the setup of the instrument and data infrastructure or the lease and safeguarding of a suitable DL deployment location near the wind farm. On the other hand, the setup of a DL specifically for the purpose of wind power predictions provides the opportunity to fit the devices location and scan pattern to this particular application, which was not the case for this study. Thus, DL data may even be more potent than estimated here for wind power predictions, reducing the amortisation time.

5.4 Volume-oriented vs balance-cost-oriented analysis

As shown, the calculation of balancing costs on the raw power prediction yields moderate refunds for all configurations instead of costs, which is not realistic on a larger scale. However, correcting for (i) the identified slight underestimations in all power prediction runs by shifting the predicted power generation by the mean power deviation $\overline{P_{dev}}$ and (ii) removing large outliers with balancing costs >300 € per time step, transforms the output to a more realistic result. With that correction, power predictions based on observational data require less balancing costs per time step and also per generated MWh. This reduction is most prominent for the persistence and ARIMA predictions based on nacelle ultrasonic observations. Naturally, the power prediction for a forecasting horizon of 15 min results in lower balancing costs than the 30 min ahead prediction.

Associating the improvement of the power prediction with clear financial costs or profits does make sense from an economic perspective. However, the systematic bias of the LoadManager$^{\circledR}$ output and the high fluctuations of single wind farm balancing costs made the aforementioned adjustments necessary. This renders cost oriented statistics for the discussed individual wind farm less robust.

However, volume oriented analyses are still valuable. One of the major motivations of this thesis is to investigate how to provide more accurate wind power predictions to improve the stability of the power grid. The current balancing price does not matter for the amount of grid power deficit or excess. Instead, only the volume difference matters for the deployment

of balancing reserves and, therefore, the stability of the electrical grid. From this Chapter's analyses the following conclusions can be drawn:

- both nacelle sonic-based approaches - the persistence and the ARIMA approach - reduce the power deviation in comparison to the currently used WPF power prediction configuration for both analysed forecasting horizons of 15 min and 30 min significantly

- power predictions based on nacelle sonic anemometer observations perform best in both, the volume-oriented and the cost-oriented approach

- the DL based power prediction does not generally yield lower power deviations than the reference WPF-based power prediction, presumably due to the distance between wind farm and observatory

- power deviations during down-ramps are slightly lower than during up-ramps

- balance cost calculation on the raw power deviation data yields differently sized refunds for all power prediction configurations, which is unrealistic and needs to be corrected

- with the discussed corrections, balancing costs per time step can be reduced by up to 80 % with the 15 min ahead sonic-based power predictions and 20 % using DL observations for the power prediction

6 Summary, conclusions and outlook

6.1 Summary and conclusions

This thesis aimed to improve power predictions for wind farms to increase electric grid stability, minimise necessary balancing efforts and costs and to reduce involuntary curtailments.

Since the energy generation by renewable power plants depends mainly on the weather, it cannot exactly be planned in advance. Therefore, power plant operators predict the power generation mainly based on prognostic weather information. In many cases, this power prediction is based on Numerical Weather Prediction (NWP) model data. However, NWP models cannot perfectly describe the future atmospheric state due to necessary parametrisations and limited grid sizes within the model. Therefore, if the power prediction for a renewable energy plant is based on that weather prediction, the power prediction is also (at least slightly) erroneous.

The main goal of this thesis was to evaluate whether observational data can reduce deviations between the wind power prediction and the observed power generation. The investigation was done on historic data of the wind farm Beeskow in Brandenburg, Germany. Close to the wind farm is the Meteorological Observatory Lindenberg - Richard Aßmann Observatory located, which is operated by DWD and provides a multi-year data set of Doppler Lidar (DL) measurements. For the wind farm, wind speed observations from nacelle ultrasonic anemometers of five wind energy turbines are available.

The power predictions were performed with the LoadManager® tool, developed by LEM Software, Leipzig, Germany. The LoadManager® tool is a software to forecast time series of energy generation, energy consumption and prices. It is used to process input data of different sources to generate wind power predictions for the wind farm Beeskow via an artifical neural network. Currently, only Wind Production Forecast (WPF) model data are used for wind power predictions. In this study, the power prediction method is enhanced by replacing the model wind speed input with observational data from the nacelle sonic anemometers and the DL.

The first research topic was discussed in Chapter 4, which describes the methodology of this thesis, the study site and the wind speed data sets.

I. **Are the Doppler lidar observations representative for the location of the wind farm? How well do the NWP model data agree to the observations?**

The relationship between the sonic and the DL data sets was investigated to evaluate the viability of the DL measurements 6 km away from the wind farm for the wind farm's power predictions. This comparison shows generally very good agreements between the observational data sets. However, wake effects of the wind farm are clearly visible in the nacelle wind speed data. The investigation of these wake effects shows that although no single nacelle observation is representative of the wind speed across the whole wind farm, the mean of the five available nacelle data sets constitutes a good approximation to the reference tower measurements for most wind directions. Therefore, both the DL and the nacelle sonic data sets are used in the subsequent wind power predictions.

Furthermore, nacelle sonic observations, DL measurements and wind speed data from a nearby meteorological tower were compared to the WPF model wind speed data to evaluate the model output. All three observational data sets show linear correlation coefficients around $R^2 \approx 0.85$ and on average approximately $1.25\,\mathrm{m\,s^{-1}}$ lower wind speed values than the model, with the difference between model and observation increasing for higher wind speed values. If the wind speed error of the model is mostly systematic, the power prediction ANN is able to learn and, thus, neutralise the error. However, the shown overestimation of the wind speed by the model impacts the application of shutdown criteria, e.g., for the protection of bats. Shutdown criteria already exist as an optional feature within the LoadManager® but they are based on absolute wind speed values. Therefore, the bias of the WPF wind speed forecast negatively affects the application of these criteria. More realistic wind speed values for the LoadManager® input, e.g., from observations, would thus also improve the reliability of these criteria.

To prepare the observational data sets for use in the wind power prediction tool they have been shifted in time for the persistence approach and put into an Auto-Regressive Integrated Moving Average (ARIMA) model to forecast the next two time steps. The application of the ARIMA model improves the wind speed prediction in comparison to the persistence model for both forecasting horizons of 15 min and 30 min slightly. While this improvement is not statistically significant, these data support the thesis that the ARIMA wind speed prediction is better than the persistence method. Still, the ARIMA approach for this site needs to be evaluated further.

In Chapter 5, the occurrence of ramp events is discussed, in order to answer research question two:

II. **Which wind situations cause challenges for wind power predictions?**

Though wind power predictions have been consistently improved in the last decade, persistent reasons for remaining uncertainties are sudden large changes in wind speed, so-called ramp events. Ramp events are defined in a binary approach as a certain power change ΔP_r during a time window WL. Seasonal and daily patterns in the occurrence of up- and down-ramps were analysed for a wind farm in Eastern Germany. The annual maximum of ramp activity occurs in March and April with about 22 monthly up- and down-ramp events. The winter months exhibit the least ramp activity with a minimum in November (eight up-ramp events, ten down-ramp events). For most ramp thresholds the number of up-ramps during the year is marginally higher than the number of down-ramps, agreeing with previous studies (*Gallego-Castillo et al.*, 2015). In most cases stricter power ramp thresholds, due to higher power threshold values ΔP_r or shorter time windows WL, lead to proportionally more up- than down-ramp events (see Tab. 5.1), which agrees well with *Pichault et al.* (2021a).

The number of ramp events starting during the day is higher than during the night. Only 5 % to 15 % of ramp events start between 21 UTC and 5 UTC. Up-ramp activity builds up early at around 6 UTC, while down-ramp activity starts to increase later at 9 UTC and reaches its maximum later at around 15 UTC. The higher ramp activity during the day is likely caused by the thermally induced formation of the atmospheric boundary layer and, therefore, increased momentum transport through the lower part of the atmosphere. Comparing results for different power ramp thresholds, the seasonal and diurnal patterns are similar. However, it is noted that the share of down-ramps decreases for stricter power ramp thresholds. This can be contributed to large scale synoptic events, e.g., cold fronts, causing

a fast increase of the wind speed while the relaxation afterwards happens over a longer time period. Furthermore, it is observed that increasing the ramp power threshold ΔP_r by 10 % roughly halves the number of observed ramp events. This can be a helpful rule of thumb for wind farm operators when evaluating the occurrence of differently sized ramp events.

Power predictions using additional wind speed observation data are able to predict significantly more ramp events than the too-smooth WPF prediction. Using current sonic anemometer wind speeds as input into the power forecast module yields up to 70 % of the observed ramp events. Doppler lidar observations done 6 km away from the wind farm as input to the power prediction proved to be less useful, albeit they also increased the number of observed ramp events from <10 % (WPF prediction) to about 20 % (*Lochmann et al.*, 2022).

To answer the third research topic, the quality of the wind power predictions with different input data is analysed.

III. **Can wind power predictions be improved with the available observational data sets? Which instrument provides the best power predictions?**

Since the currently used NWP base data for wind power predictions with the LoadManager® could not be modified, the existing methods within the LoadManager® were enhanced to investigate the use of observational data in real-time. Firstly, measurements from the DL and the nacelle sonic anemometers were evaluated in a persistence model (naïve model). Secondly, the latter have been post-processed with the ARIMA method. These data sets have been used to train a multi-layer perceptron ANN for approximately two years (1 January 2014 to 31 December 2015) of observation and wind power generation data. To evaluate the potency of these data sets, wind power predictions based on observational data have been performed with the LoadManager® for a prediction period of approximately nine months (1 January 2019 to 23 September 2019) and compared to the power prediction configuration purely based on NWP data for forecasting horizons of 15 min and 30 min.

The quality of the observation-based power predictions in comparison to the reference has been evaluated in two ways. First, the *volume*-oriented approach to determine the power deviation P_{dev} has been applied. The closer P_{dev} is to 0, the less extensive balancing services need to be. For the entire data set, all power predictions based on nacelle sonic observations, i.e., the sonic persistence and the ARIMA configurations, reduce the mean absolute power deviation compared to the reference NWP-based power prediction. Power predictions based on DL data only represent an improvement for the wind sector, where the wind farm is downwind of the DL, otherwise rarely. Furthermore, the performance of the observation-based power predictions was evaluated specifically during ramp events. Power deviations during ramp events are on average three times as large as for the entire data set. Power predictions based on sonic observations reduce P_{dev} compared to the reference even further than for the entire data set, at least for moderately steep ramp events. For very steep ramp events, either by a short window length WL or a high power difference ΔP_r, the observation-based approaches do not provide an improvement. However, very steep ramp events happen rarely in the wind distribution of the study region. As for the entire data set, power predictions based on DL data only provide an improvement over the reference, if the wind farm is downwind of the DL instrument. However, for the potential operational use of this approach power predictions for all wind directions are necessary and, therefore, the performance of the DL power predictions has been evaluated for all wind directions. In summary, wind power predictions can be improved by including observational data sets. In the setup of this thesis, the in-situ measurements of the ultrasonic nacelle anemometers provided the largest improvement.

The discussion of the *cost*-oriented approach yields the answer to the fourth research topic:

IV. **Does the benefit of using observational data outweigh the acquisition costs of the instruments? How long is the amortisation time?**

The different power prediction configurations have been evaluated according to the balancing costs necessary to compensate occurring power deviations. To achieve realistic balancing costs, the underestimation bias has been corrected and extreme outliers have been removed. The resulting balancing costs per time step of the new observation-based power configurations are lower than the reference for both investigated forecasting horizons. For the 15 min ahead predictions balancing costs are 75 % to 80 % less than for the WPF-reference. For the forecasting horizon of 30 min, balancing costs for the DL-based power prediction are 20 % less than the reference and for the sonic-based approaches 60 %-75 % less than the reference. Therefore, using the nacelle anemometer observations for very short-term power predictions is clearly beneficial, as these observations are already mandatory. Nacelle observations with cup anemometers as power prediction input likely generate similar benefits as sonic anemometers, as they provide measurements with similar accuracy (e.g., *Sotelino et al.*, 2012; *WMO*, 2014). Acquisition of a DL would turn profitable after approximately five years under the conditions of this study, which is reasonably in the typical life span of DL instruments. However, DL observations performed specifically for the purpose of wind power predictions could provide further improvements, e.g., by observing the wind speed closer to the wind farm or at a higher vertical resolution, thus increasing their benefit and reducing the amortisation time.

6.2 Outlook

For the wind power predictions based on nacelle sonic observations, the mean value of the anemometer measurements of the five available nacelles has been used. Section 4.1.3 demonstrates that this value is preferable over the measurements of a single instrument. It is shown that Easterly winds are underrepresented in this data set. Since Easterly winds are less frequent and on average slower, the slightly worse representation of Easterly winds does not invalidate the results of this thesis. However, for the operational application of this approach, the responsible wind farm operators or direct marketers presumably have nacelle wind speed data of all their wind energy turbines available. An enhanced approach could utilise wind speed data only of nacelles at the upwind edge of the wind farm, weighted with the dimension of the wind farm. Additionally, a complete data set including in-situ wind direction measurements at the nacelles would presumably increase the prediction accuracy.

Furthermore, the benefit of operating a vertically-pointing DL within a wind farm could be evaluated in future studies. Contrarily, in this study, the DL instruments operational scanning pattern is optimised to best observe the vertical wind profile above the location of the DWD observatory. Although that is not the ideal scanning pattern to observe the wind profile at the location of the wind farm, it could not be altered in the scope of this thesis. For better use in this approach, a ground based DL could either be located directly at the wind farm, providing a vertical profile of the wind field there, or it could be located at the Southwestern edge of the wind farm scanning towards the main wind direction to observe the wind upstream of the wind farm. The first approach would in theory yield the similar benefits to the sonic observations, providing wind speed measurements at hub height. In addition, it can also provide measurements in the layers immediately above the hub height and wind direction information for the column. The second approach has already been successfully applied with horizontally pointing DL (*Pichault et al.*, 2021a,b) or Doppler radar

(*Valldecabres et al.*, 2020), used to predict wind power generation for wind farms offshore and onshore in moderately complex terrain with significant improvements compared to the benchmarks. A horizontally pointing DL has the advantage of observing a cone-shaped layer at preferably hub height and, therefore, evaluating the whole wind field before it arrives at the wind farm. However, if the DL is only located at one corner of the wind farm, it can only provide good upstream results for a specific wind direction sector.

One critically discussed point of the binary ramp definition is the definition of arbitrary power ramp and window length threshold values. *Gallego et al.* (2011) propose a continuous wavelet transformation approach as alternative, which was already applied (*Pichault et al.*, 2021a) and can be evaluated against this study's approach in future work.

Apart from refining the ramp identification method, there is potential to add more complexity to the wind speed and wind power prediction methods. The implementation of the ARIMA model requires more work to yield statistically significant improvements of the wind speed prediction, to which this study's results only hinted at. Using a data set with higher temporal resolution or including exogenous variables as, e.g., atmospheric pressure, can potentially provide the desired improvement. Recently, *Qian et al.* (2019) summarised hybrid methods for wind speed and wind power prediction, e.g., the coupling of an ANN or ARIMA model with wavelet-based methods used to decompose the original time series into different subseries, which could be modeled more effectively than the original time series. Furthermore, power deviations could be evaluated according to different synoptical situations, e.g., different wind speed classes, wind direction or atmospheric stability.

A Appendices

A.1 Wind direction dependency of WPF-sonic comparison

Figure A.1 shows the 2D histograms for the comparison between WPF model wind speed and nacelle sonic wind speed depending on the WPF wind direction. As shown in Sect. 4.2.2, the difference between WPF wind speed and observational wind speed of all three data sources (nacelle sonic anemometers, DL and MT) depends on the wind speed and increases for higher wind speed values. This is characterised by the slope m of the linear correlation and attributed to a fast bias by the WPF model. For the DL and the MT comparisons, the value of m does not change much for each of the four 90° wind direction sectors (see MT-WPF comparison in Fig. A.2). However, for the comparison between nacelle sonic wind speed and WPF model wind speed, m does depend on the wind direction (see Fig. A.1). For Easterly winds, m is considerably lower than for Westerly winds. This is because the nacelle sonic anemometer data set is affected by wake effects within the wind farm, explained in Sect. 4.1.3. Especially of Easterly winds, these wake effects cause the observed wind speed within the wind farm to be lower than it would be in an undisturbed flow, with the difference being higher for higher wind speed values. For the Easterly comparison between WPF and nacelle sonic measurement that means WPF wind speed values are higher than they should be. This fast bias by the model affects all comparisons. Also, for this wind direction the sonic wind speed values are lower than they should be due to wake effects, which only affect sonic comparisons and are most pronounced for Easterly winds. Therefore, the difference between both wind speed sources increases for higher wind speed values even more, which is characterised by the lower slope m for Easterly winds.
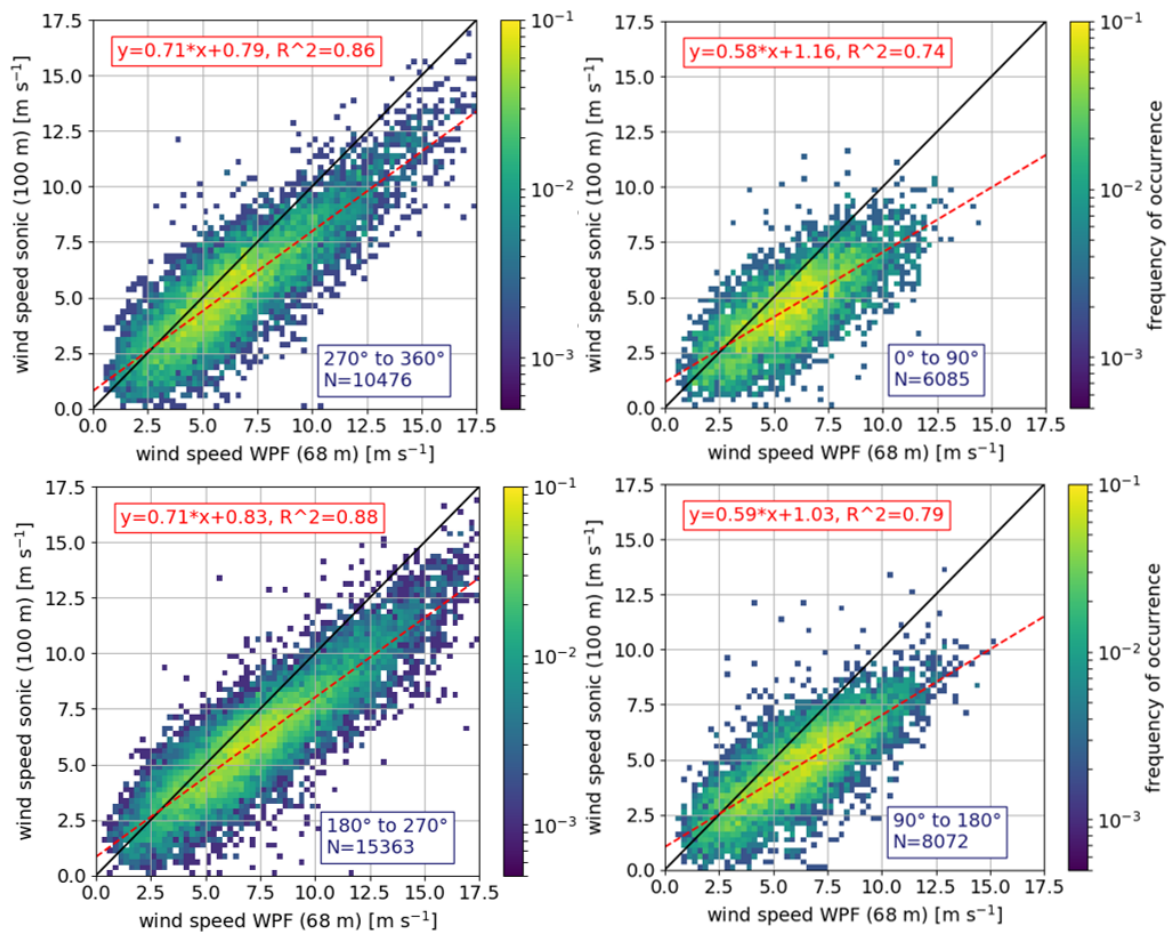
Figure A.1: Wind speed comparison between WPF model wind speed and nacelle sonic anemometer measurements for four wind direction sectors (blue box). 2D histograms of both wind speed sources. The solid black line marks the 1:1 line and the dashed red line represents a linear regression with its parameters displayed in the red box. The amount of data points N in this comparison is shown in the blue box. The bin width of the histograms is $0.25 \, \mathrm{m \, s^{-1}}$.
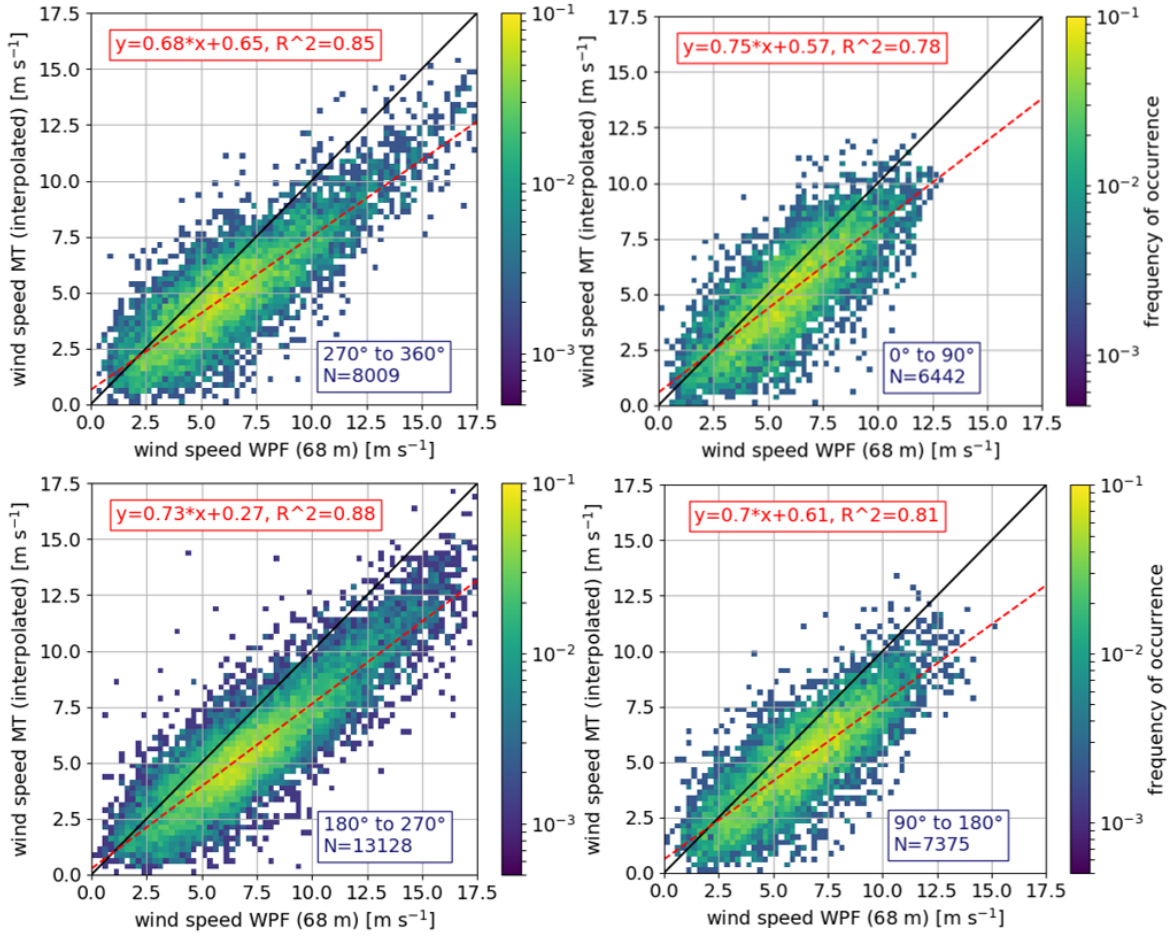
Figure A.2: Same analysis as Fig. A.1 but for the comparison between WPF model wind speed and MT measurements.

A.2 Occurrence of ramp events for different ramp detection thresholds

Figure A.3 shows the diurnal ramp occurrence at the wind farm Beeskow for six additional ramp detection thresholds (ΔP_r=30 %, 40 %, 50 % and WL=30 min, 60 min). As these thresholds are stricter than the previously discussed values, the absolute number of ramps detected is lower than in Fig. 5.2. However, the general shape of these additional histograms is similar to the previously shown distribution (Fig. 5.2).

Furthermore, Fig. A.4 shows the seasonal ramp occurrence at the wind farm Beeskow for the same six additional ramp detection thresholds. These stricter thresholds yield lower absolute numbers of ramp events, but the seasonal cycle resembles the previously shown distribution (Fig. 5.1).
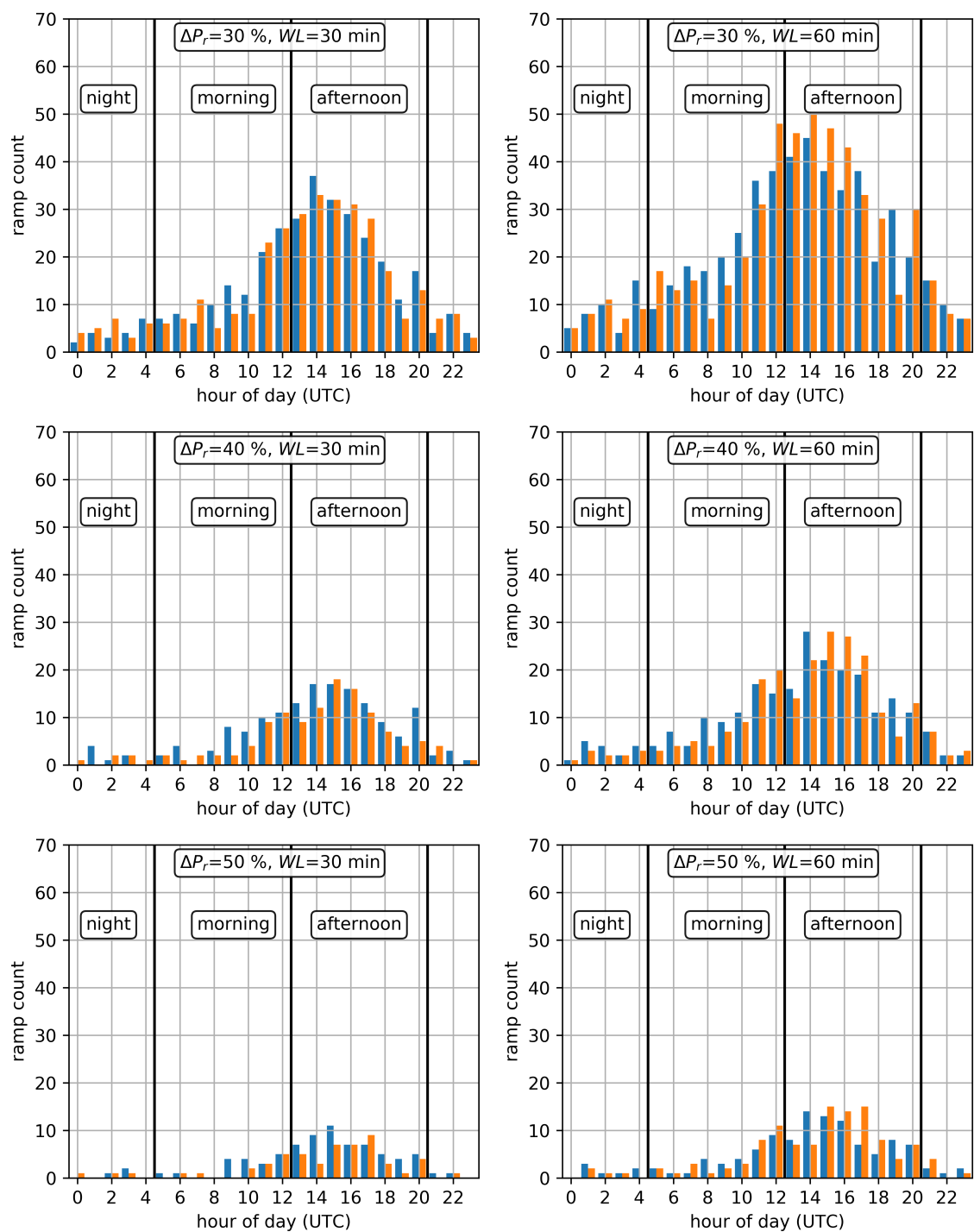
Figure A.3: Same analysis as Fig. 5.2 but for six additional ramp detection thresholds, indicated at the top of each figure.
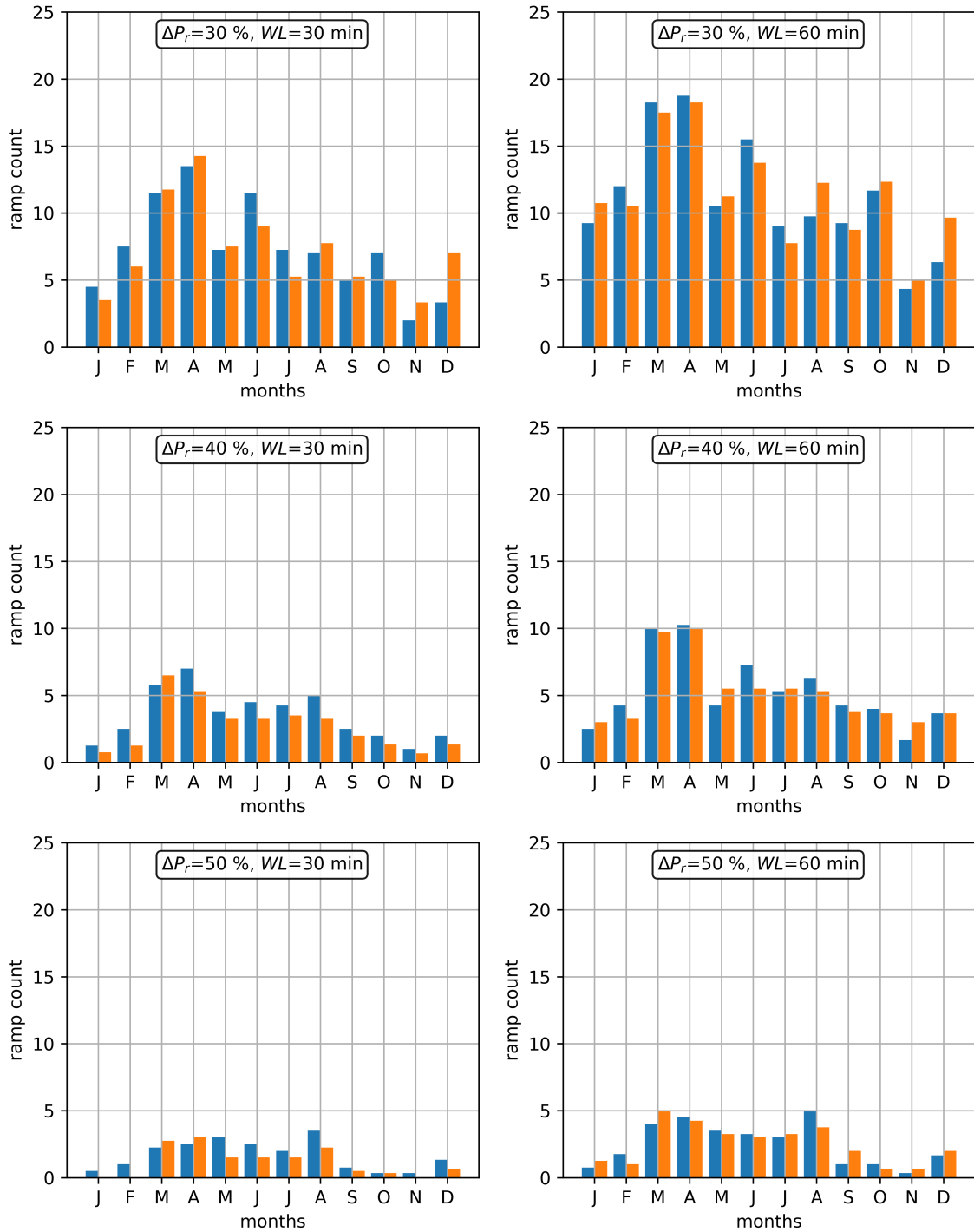
Figure A.4: Same analysis as Fig. 5.1 but for six additional ramp detection thresholds, indicated at the top of each figure.

List of Figures

List of Tables

Bibliography

Agora-Energiewende: Stromnetze Für 65 Prozent Erneuerbare Bis 2030. Zwölf Maßnahmen Für Den Synchronen Ausbau von Netzen Und Erneuerbaren Energien., Tech. rep., Agora-Energiewende, Berlin, 2018.

Allik, A., Uiga, J., and Annuk, A.: Deviations between Wind Speed Data Measured with Nacelle-Mounted Anemometers on Small Wind Turbines and Anemometers Mounted on Measuring Masts, p. 12, 2014.

Antoniou, I. and Friis Pedersen, T.: Nacelle Anemometry on a 1MW Wind Turbine. Comparing the Power Performance Results by Use of the Nacelle or Mast Anemometer, 1997.

AWS-Truewind: AWS Truewind's Final Report for the Alberta Forecasting Pilot Project, Technical Report, AWS-Truewind, Albany, New York, 2008.

Baas, P., Bosveld, F. C., Klein Baltink, H., and Holtslag, A. A. M.: A Climatology of Nocturnal Low-Level Jets at Cabauw, Journal of Applied Meteorology and Climatology, 48, 1627–1642, doi:10.1175/2009JAMC1965.1, 2009.

Betz, A.: Die Windmühlen im Lichte neuerer Forschung, Naturwissenschaften, 15, 905–914, doi:10.1007/BF01506119, 1927.

Bianco, L., Djalalova, I. V., Wilczak, J. M., Cline, J., Calvert, S., Konopleva-Akish, E., Finley, C., and Freedman, J.: A Wind Energy Ramp Tool and Metric for Measuring the Skill of Numerical Weather Prediction Models, Weather and Forecasting, 31, 1137–1156, doi:10.1175/WAF-D-15-0144.1, 2016.

Bossavy, A., Girard, R., and Kariniotakis, G.: Forecasting Ramps of Wind Power Production with Numerical Weather Prediction Ensembles, Wind Energy, 16, 51–63, doi:10.1002/we.526, 2013.

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J., and Shelly, A.: Unified Modeling and Prediction of Weather and Climate: A 25-Year Journey, Bulletin of the American Meteorological Society, 93, 1865–1877, doi:10.1175/BAMS-D-12-00018.1, 2012.

Bühl, J., Leinweber, R., Görsdorf, U., Radenz, M., Ansmann, A., and Lehmann, V.: Combined Vertical-Velocity Observations with Doppler Lidar, Cloud Radar and Wind Profiler, Atmospheric Measurement Techniques, 8, 3527–3536, doi:10.5194/amt-8-3527-2015, 2015.

Bundeswirtschaftsministerium: Zeitreihen Zur Entwicklung Der Erneuerbaren Energien in Deutschland, Tech. rep., Dessau, 2022.

Bushaev, V.: Stochastic Gradient Descent with Momentum, https://towardsdatascience.com/stochastic-gradient-descent-with-momentum-a84097641a5d, 2017.

Cheung, Y.-W. and Lai, K. S.: Lag Order and Critical Values of the Augmented Dickey–Fuller Test, Journal of Business & Economic Statistics, 13, 277–280, doi:10.1080/07350015.1995.10524601, 1995.

Cutler, N., Kay, M., Jacka, K., and Nielsen, T. S.: Detecting, Categorizing and Forecasting Large Ramps in Wind Farm Power Output Using Meteorological Observations and WPPT, Wind Energy, 10, 453–470, doi:10.1002/we.235, 2007.

Dolatabadi, A., Abdeltawab, H., and Mohamed, Y. A.-R. I.: Deep Spatial-Temporal 2-D CNN-BLSTM Model for Ultrashort-Term LiDAR-Assisted Wind Turbine's Power and Fatigue Load Forecasting, IEEE Transactions on Industrial Informatics, 18, 2342–2353, doi:10.1109/TII.2021.3097716, 2022.

Frehlich, R.: Estimation of Velocity Error for Doppler Lidar Measurements, Journal of Atmospheric and Oceanic Technology, 18, 1628–1639, doi:10.1175/1520-0426(2001)018<1628:EOVEFD>2.0.CO;2, 2001.

Gallego, C., Costa, A., Cuerva, Á., Landberg, L., Greaves, B., and Collins, J.: A Wavelet-based Approach for Large Wind Power Ramp Characterisation, Wind Energy, Wind. Energ. 2013, 16, 257–278, doi:10.1002/we.550, 2011.

Gallego-Castillo, C., Cuerva-Tejero, A., and Lopez-Garcia, O.: A Review on the Recent History of Wind Power Ramp Forecasting, Renewable and Sustainable Energy Reviews, 52, 1148–1157, doi:10.1016/j.rser.2015.07.154, 2015.

González-Longatt, F., Wall, P., and Terzija, V.: Wake Effect in Wind Farm Performance: Steady-state and Dynamic Behavior, Renewable Energy, 39, 329–338, doi:10.1016/j.renene.2011.08.053, 2012.

Grant, A. L. M. and Mason, P. J.: Observations of Boundary-Layer Structure over Complex Terrain, Quarterly Journal of the Royal Meteorological Society, 116, 159–186, doi:10.1002/qj.49711649107, 1990.

Grigonytė, E. and Butkevičiūtė, E.: Short-Term Wind Speed Forecasting Using ARIMA Model, Energetika, 62, doi:10.6001/energetika.v62i1-2.3313, 2016.

Hisscott, A.: Exploiting Wind Profiler Information, Weather, 74, 8–19, doi:10.1002/wea.3387, 2019.

Hooper, D. A., Edwards, D. M., Holmes, G., Linklater, K., Oakley, T., Parrett, C., and Turp, M.: The Usefulness of Model-Comparison Statistics for Wind-Profiling Radar Operators, in: Proceedings of the Thirteenth International Workshop on Technical and Scientific Aspects of MST Radar, pp. 141–150, Leibniz-Institute of Atmospheric Physics at the Rostock University, K¨uhlungsborn, Germany, 2013.

Hyndman, R. J. and Athanasopoulos, G.: Forecasting: Principles and Practice (3rd Ed), OTexts: Melbourne, Australia, 2021.

Illingworth, A. J., Hogan, R. J., O'Connor, E. J., Bouniol, D., Brooks, M. E., Delanoé, J., Donovan, D. P., Eastment, J. D., Gaussiat, N., Goddard, J. W. F., Haeffelin, M., Baltink, H. K., Krasnov, O. A., Pelon, J., Piriou, J.-M., Protat, A., Russchenberg, H. W. J., Seifert, A., Tompkins, A. M., van Zadelhoff, G.-J., Vinit, F., Willén, U., Wilson, D. R., and Wrench, C. L.: Cloudnet: Continuous Evaluation of Cloud Profiles in Seven Operational Models

Using Ground-Based Observations, Bulletin of the American Meteorological Society, 88, 883–898, doi:10.1175/BAMS-88-6-883, 2007.

Illingworth, A. J., Cimini, D., Gaffard, C., Haeffelin, M., Lehmann, V., Löhnert, U., O'Connor, E. J., and Ruffieux, D.: Exploiting Existing Ground-Based Remote Sensing Networks to Improve High-Resolution Weather Forecasts, Bulletin of the American Meteorological Society, 96, 2107–2125, doi:10.1175/BAMS-D-13-00283.1, 2015.

IPCC, ed.: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (Eds.)], Cambridge University Press, New York, NY, 2014.

IPCC: Summary for Policymakers: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, doi:10.1017/9781009325844.001, 2022.

Kamath, C.: Understanding Wind Ramp Events through Analysis of Historical Data, in: IEEE PES T D 2010, pp. 1–6, doi:10.1109/TDC.2010.5484508, 2010.

Kelley, N. D., Jonkman, B. J., and Scott, G. N.: Comparing Pulsed Doppler LIDAR with SODAR and Direct Measurements for Wind Assessment, p. 24, 2007.

Kost, C., Shammugam, S., Fluri, V., Peper, D., Memar, A. D., and Schlegl, T.: Levelized Cost of Electricity - Renewable Energy Technologies, Tech. rep., Fraunhofer Institute for Solar Energy Systems ISE, Freiburg, Germany, 2021.

Kraus, H.: Die Atmosphäre der Erde: eine Einführung in die Meteorologie, Springer, Berlin Heidelberg, 3., erw. und aktualisierte aufl edn., 2004.

Kraus, H.: Grundlagen der Grenzschicht-Meteorologie: Einführung in die Physik der atmosphärischen Grenzschicht und in die Mikrometeorologie, Springer, Berlin Heidelberg, 2008.

Lane, S. E., Barlow, J. F., and Wood, C. R.: An Assessment of a Three-Beam Doppler Lidar Wind Profiling Method for Use in Urban Areas, Journal of Wind Engineering and Industrial Aerodynamics, 119, 53–59, doi:10.1016/j.jweia.2013.05.010, 2013.

Larsen, M. F. and Röttger, J.: VHF and UHF Doppler Radars as Tools for Synoptic Research, Bulletin of the American Meteorological Society, 63, 996–1008, doi:10.1175/1520-0477(1982)063<0996:VAUDRA>2.0.CO;2, 1982.

Lochmann, M., Kalesse-Los, H., Schäfer, M., Heinrich, I., and Leinweber, R.: Analysing Wind Power Ramp Events and Improving Very Short-Term Wind Power Predictions by Including Wind Speed Observations, doi:(revised version submitted), 2022.

Lundquist, J. K., DuVivier, K. K., Kaffine, D., and Tomaszewski, J. M.: Costs and Consequences of Wind Turbine Wake Effects Arising from Uncoordinated Wind Energy Development, Nat Energy, 4, 26–34, doi:10.1038/s41560-018-0281-2, 2019.

Menzies, T., Kocagüneli, E., Minku, L., Peters, F., and Turhan, B.: Chapter 24 - Using Goals in Model-Based Reasoning, in: Sharing Data and Models in Software Engineering, edited by Menzies, T., Kocagüneli, E., Minku, L., Peters, F., and Turhan, B., pp. 321–353, Morgan Kaufmann, Boston, doi:10.1016/B978-0-12-417295-1.00024-2, 2015.

Ministry for Agriculture, Environment and Climate Protection Brandenburg: Handlungsempfehlung Zum Umgang Mit Fledermäusen Bei Der Planung Und Genehmigung von Windenergieanlagen in Brandenburg, 2010.

Monin, A. S. and Obukhov, A. M.: Basic Laws of Turbulent Mixing in the Surface Layer of the Atmosphere, p. 30, 1954.

Necker, T., Weissmann, M., and Sommer, M.: The Importance of Appropriate Verification Metrics for the Assessment of Observation Impact in a Convection-permitting Modelling System, Q.J.R. Meteorol. Soc., 144, 1667–1680, doi:10.1002/qj.3390, 2018.

Nfaoui, H.: 2.04 - Wind Energy Potential, in: Comprehensive Renewable Energy, edited by Sayigh, A., pp. 73–92, Elsevier, Oxford, doi:10.1016/B978-0-08-087872-0.00204-3, 2012.

Päschke, E., Leinweber, R., and Lehmann, V.: A One Year Comparison of 482 MHz Radar Wind Profiler, RS92-SGP Radiosonde and 1.5 Mm Doppler Lidar Wind Measurements, Preprint, Others (Wind, Precipitation, Temperature, etc.)/Remote Sensing/Validation and Intercomparisons, doi:10.5194/amtd-7-11439-2014, 2014.

Päschke, E., Leinweber, R., and Lehmann, V.: An Assessment of the Performance of a 1.5 Mm Doppler Lidar for Operational Vertical Wind Profiling Based on a 1-Year Trial, Atmos. Meas. Tech., 8, 2251–2266, doi:10.5194/amt-8-2251-2015, 2015.

Pichault, M., Vincent, C., Skidmore, G., and Monty, J.: Characterisation of Intra-Hourly Wind Power Ramps at the Wind Farm Scale and Associated Processes, Wind Energy Science, 6, 131–147, doi:10.5194/wes-6-131-2021, 2021a.

Pichault, M., Vincent, C., Skidmore, G., and Monty, J.: Short-Term Wind Power Forecasting at the Wind Farm Scale Using Long-Range Doppler LiDAR, Energies, 14, 2663, doi:10.3390/en14092663, 2021b.

Potter, C. W., Grimit, E., and Nijssen, B.: Potential Benefits of a Dedicated Probabilistic Rapid Ramp Event Forecast Tool, in: 2009 IEEE/PES Power Systems Conference and Exposition, pp. 1–5, doi:10.1109/PSCE.2009.4840109, 2009.

Qian, N.: On the Momentum Term in Gradient Descent Learning Algorithms, Neural Networks, 12, 145–151, doi:10.1016/S0893-6080(98)00116-6, 1999.

Qian, Z., Pei, Y., Zareipour, H., and Chen, N.: A Review and Discussion of Decomposition-Based Hybrid Models for Wind Energy Forecasting Applications, Applied Energy, 235, 939–953, doi:10.1016/j.apenergy.2018.10.080, 2019.

Quaschning, V.: Understanding Renewable Energy Systems, Routledge, Taylor & Francis Group, Earthscan from Routledge, London ; New York, revised edition edn., 2016.

Quaschning, V.: Erneuerbare Energien und Klimaschutz: Hintergründe, Techniken und Planung, Ökonomie und Ökologie, Energiewende, Hanser, München, 4., überarbeitete und erweiterte auflage edn., 2018.

Ramos, F., Vidal, M., Bone, A., Malon, H., and Aguirre de Juana, Á.: Analysis of the Air Flow Generated by an Air-Assisted Sprayer Equipped with Two Axial Fans Using a 3D Sonic Anemometer, Sensors (Basel, Switzerland), 12, 7598–613, doi:10.3390/s120607598, 2012.

Robinius, M., Stein, F., Schwane, A., and Stolten, D.: A Top-Down Spatially Resolved Electrical Load Model, Energies, 2017, doi:10.3390/en10030361, 2017.

Röttger, J. and Larsen, M. F.: UHF/VHF Radar Techniques for Atmospheric Research and Wind Profiler Applications, in: Radar in Meteorology: Battan Memorial and 40th Anniversary Radar Meteorology Conference, edited by Atlas, D., pp. 235–281, American Meteorological Society, Boston, MA, doi:10.1007/978-1-935704-15-7_23, 1990.

Schleussner, C.-F., Rogelj, J., Schaeffer, M., Lissner, T., Licker, R., Fischer, E. M., Knutti, R., Levermann, A., Frieler, K., and Hare, W.: Science and Policy Characteristics of the Paris Agreement Temperature Goal, Nature Clim Change, 6, 827–835, doi:10.1038/nclimate3096, 2016.

Shin, D. and Ko, K.: Experimental Study on Application of Nacelle-Mounted Li-DAR for Analyzing Wind Turbine Wake Effects by Distance, Energy, 243, 123 088, doi:10.1016/j.energy.2021.123088, 2022.

SMARD: SMARD - Strommarkt Erklärt, https://www.smard.de/en/strommarkt-erklaert/alle, 2021a.

SMARD, B.: SMARD - Data Download Portal, Tech. rep., Bundesnetzagentur, 2021b.

Smith, D. A., Harris, M., Coffey, A. S., Mikkelsen, T., Jørgensen, H. E., Mann, J., and Danielian, R.: Wind Lidar Evaluation at the Danish Wind Test Site in Høvsøre, Wind Energy, 9, 87–93, doi:10.1002/we.193, 2006.

Smith, T. G., et, and al: Pmdarima: ARIMA Estimators for Python, 2017.

Sotelino, L. G., Coster, N. D., Beirinckx, P., and Peeters, P.: Intercomparison of Cup Anemometer and Sonic Anemometers on Site at Uccle/Belgium, p. 7, 2012.

Standen, J., Wilson, C., Vosper, S., and Clark, P.: Prediction of Local Wind Climatology from Met Office Models: Virtual Met Mast Techniques, Wind Energy, 20, 411–430, doi:10.1002/we.2013, 2017.

Steiner, A., Köhler, C., Metzinger, I., Braun, A., Zirkelbach, M., Ernst, D., Tran, P., and Ritter, B.: Critical Weather Situations for Renewable Energies – Part A: Cyclone Detection for Wind Power, Renewable Energy, 101, 41–50, doi:10.1016/j.renene.2016.08.013, 2017.

Strauss, C. E. M.: Beating Betz' Law: A Larger Fundamental Upper Bound for Wind Energy Harvesting, arXiv:2110.14612 [physics], 2021.

Stull, R.: Practical Meteorology: An Algebra-based Survey of Atmospheric Science., version 1.02b edn., 2017.

Tay, K., Koh, T.-Y., and Skote, M.: Characterizing Mesoscale Variability in Low-Level Jet Simulations for CBLAST-LOW 2001 Campaign, Meteorol Atmos Phys, 133, 163–179, doi:10.1007/s00703-020-00736-3, 2021.

TenneT: Bilanzkreise - TenneT, https://www.tennet.eu/de/strommarkt/strommarkt-in-deutschland/bilanzkreise/, 2022.

Theuer, F., van Dooren, M. F., von Bremen, L., and Kühn, M.: Minute-Scale Power Forecast of Offshore Wind Turbines Using Long-Range Single-Doppler Lidar Measurements, Wind Energy Science, 5, 1449–1468, doi:10.5194/wes-5-1449-2020, 2020.

Valldecabres, L., von Bremen, L., and Kühn, M.: Minute-Scale Detection and Probabilistic Prediction of Offshore Wind Turbine Power Ramps Using Dual-Doppler Radar, Wind Energy, 23, 2202–2224, doi:10.1002/we.2553, 2020.

Vestas: Vestas Wind Systems Brochure for V80-2.0 MW, V90-1.8/2.0 MW and V100-1.8 MW Wind Turbines, 2011.

Walczak, S.: Artificial Neural Networks, in: Encyclopedia of Information Science and Technology, Fourth Edition, pp. 631–645, IGI Global, doi:10.4018/978-1-5225-2255-3.ch011, 2018.

Werner, C.: Doppler Wind Lidar, in: Lidar, edited by Weitkamp, C., vol. 102, pp. 325–354, Springer-Verlag, New York, doi:10.1007/0-387-25101-4_12, 2005.

Wharton, S. and Lundquist, J. K.: Atmospheric Stability Affects Wind Turbine Power Collection, Environ. Res. Lett., 7, 014 005, doi:10.1088/1748-9326/7/1/014005, 2012.

Wilczak, J. M., Olson, J. B., Djalalova, I., Bianco, L., Berg, L. K., Shaw, W. J., Coulter, R. L., Eckman, R. M., Freedman, J., Finley, C., and Cline, J.: Data Assimilation Impact of in Situ and Remote Sensing Meteorological Observations on Wind Power Forecasts during the First W Ind F Orecast I Mprovement P Roject (WFIP), Wind Energy, 22, 932–944, doi:10.1002/we.2332, 2019.

WMO: Guide to Meteorological Instruments and Methods of Observation: (CIMO Guide). 2014 Edition., Report, World Meteorological Organization, doi:10.25607/OBP-1533, 2014.

Würth, I., Valldecabres, L., Simon, E., Möhrlen, C., Uzunoğlu, B., Gilbert, C., Giebel, G., Schlipf, D., and Kaifel, A.: Minute-Scale Forecasting of Wind Power—Results from the Collaborative Workshop of IEA Wind Task 32 and 36, Energies, 12, 712, doi:10.3390/en12040712, 2019.

Zack, J. W.: Optimization of Wind Power Production Forecast Performance during Critical Periods for Grid Management, in: Proceedings of the European Wind Energy Conference EWEC, Milano (IT), vol. 8, 2007.

Zimmer, J., Raabe, A., and Lemberg, T.: Increasing Grid Stability through Accurate Infeed Forecasts of Renewable Energies, in: CIRED 2012 Workshop: Integration of Renewables into the Distribution Grid, pp. 381–381, IET, Lisbon, Portugal, doi:10.1049/cp.2012.0899, 2012.

Acronyms

ABL Atmospheric Boundary Layer

ANN Artificial Neural Network

ARIMA Auto-Regression Integrated Moving Average

DL Doppler Lidar

DWD German Meteorological Service (Deutscher Wetterdienst - DWD)

GCC Grid Control Cooperation

LIM Leipzig Institute for Meteorology

LR Learning Range

MOLRAO Meteorological Observatory Lindenberg - Richard-Aßmann Observatory

MT Meteorological Tower

NWP Numerical Weather Prediction

RR Recall Range

TSO Transmission System Operator

UM Unified Model

VMM Virtual MetMast

WET Wind Energy Turbine

WPF Wind Production Forecast

WPR Wind-Profiling Radar

Acknowledgements