Characterization and forecasting of wind conditions over the Baltic Sea

CHRISTOFFER HALLGREN
Dissertation presented at Uppsala University to be publicly examined in Hambergssalen, Geocentrum, Villavägen 16, Uppsala, Friday, 10 November 2023 at 13:00 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Julie Lundquist (University of Colorado, Boulder).

Abstract

To meet the increasing demand of sustainably produced electricity, the total number of installed wind turbines is rapidly increasing globally. Although onshore installations are dominating, offshore wind power is taking a greater share of the market every year. Offshore, the wind is generally stronger than onshore and with the possibility to construct bigger turbines the electricity yield is also greater per turbine. Furthermore, it is possible to build larger wind farms offshore than onshore. The Baltic Sea is an area of high interest to many stakeholders and a major expansion of offshore wind power is expected in the region in the coming decades.

As the Baltic Sea is a semi-enclosed sea with relatively short distances to the coast from anywhere in the basin, there are many mesoscale meteorological phenomena occurring, affecting the shape of the wind profile and the preconditions for wind power. This thesis focuses on these wind profiles, utilizing multi-year lidar observations and state-of-the-art numerical models. Wind profiles with a local maximum, i.e., low-level jets, are of special interest as they are frequently occurring over the Baltic Sea. These non-ideal wind speed profiles are characterized in terms of frequency and effects on turbulent properties, and the best way to define the low-level jets is investigated. Furthermore, the change in wind direction with height is addressed and a new index to automatically identify sea and land breeze circulations in operational weather prediction models is created. Finally, different post-processing methods to improve short-term forecasts of wind power production are compared and a recommendation on how to combine the methods depending on the weather situation is presented. Altogether, the research in this thesis adds a piece to the puzzle in reaching further understanding of the Baltic Sea wind conditions. The findings will be useful also in other coastal areas in siting, farm layout, and load analysis as well as in creating improved power production forecasts for offshore wind turbines.

Keywords: Baltic Sea, forecasting, low-level jet, lidar, machine learning, mesoscale meteorology, numerical weather prediction, post-processing, reanalysis, sea breeze, turbulence, wind atlas, wind energy, wind power

Christoffer Hallgren, Department of Earth Sciences, LUVAL, Villav. 16, Uppsala University, SE-75236 Uppsala, Sweden.

© Christoffer Hallgren 2023

ISSN 1651-6214
URN urn:nbn:se:uu:diva-512116 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-512116)
The answer, my friend,
is blowin’ in the wind
– Bob Dylan
List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the publishers.
Author contributions

My contributions\(^1\) to Papers I–VII:

I carried out the conceptualization, formal analysis, investigation, methodology, software, validation, and visualization in all papers. I wrote the original drafts of all papers, apart from Paper III where J. A. Aird wrote most of the draft. J. A. Aird was also setting up the neural network in Paper III. For Papers I, II, IV, VI, and VII, I worked on the review and final editing of the papers. Papers III and V are submitted, but not yet reviewed, at the time of writing this thesis.

During my doctoral studies, I also contributed to the following papers, which are not included in this thesis:


---

\(^1\)According to the Contributor Roles Taxonomy (CRediT), https://credit.niso.org/
Abbreviations and acronyms

AMJJ April – July
AROME Applications of Research to Operations at Mesoscale
ASIT Air Sea Interaction Tower
A.S.L. Above Sea Level
ASON August – November
BLH Boundary-Layer Height
CAPE Convective Available Potential Energy
CMIP6 Coupled Model Intercomparison Project Phase 6
CSI Critical Success Index
D1/D2 A combination of 0–23 h and 24–47 h forecasts
DJFM December – March
ECMWF European Centre for Medium-Range Weather Forecasts
ERA5 ECMWF Reanalysis version 5
HARMONIE HIRLAM–ALADIN Research on Mesoscale Operational NWP in Euromed
IEA International Energy Agency
IFS Integrated Forecasting System
LBI Land Breeze Index
LCC Low Cloud Cover
LIDAR Light Detection And Ranging
LLJ Low-Level Jet
LLM Low-Level Minimum
LWT Lamb Weather Type
MAE Mean Absolute Error
ML Machine Learning
MMIJ Meteorological Mast IJmuiden
MODIS Moderate Resolution Imaging Spectroradiometer
NASA National Aeronautics and Space Administration
NBH Neighbourhood method
NN Neural Network
NWP Numerical Weather Prediction
MERRA2 Modern-Era Retrospective analysis for Research and Applications, version 2
MSLP Mean Sea Level Pressure
NEWA New European Wind Atlas
PCHIP Piecewise Cubic Hermite Interpolating Polynomial
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>REWS</td>
<td>Rotor Equivalent Wind Speed</td>
</tr>
<tr>
<td>RF</td>
<td>Random Forest</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>Rn</td>
<td>Net Radiation</td>
</tr>
<tr>
<td>SBI</td>
<td>Sea Breeze Index</td>
</tr>
<tr>
<td>SDG</td>
<td>Sustainable Development Goal</td>
</tr>
<tr>
<td>SEDI</td>
<td>Symmetric Extremal Dependence Index</td>
</tr>
<tr>
<td>SHF</td>
<td>Surface Sensible Heat Flux</td>
</tr>
<tr>
<td>SODAR</td>
<td>Sound Detection And Ranging</td>
</tr>
<tr>
<td>SSP</td>
<td>Shared Socio-economic Pathway</td>
</tr>
<tr>
<td>SST</td>
<td>Sea Surface Temperature</td>
</tr>
<tr>
<td>TKE</td>
<td>Turbulent Kinetic Energy</td>
</tr>
<tr>
<td>UERRA</td>
<td>Uncertainties in Ensembles of Regional ReAnalyses</td>
</tr>
<tr>
<td>WRF</td>
<td>Weather Research and Forecasting</td>
</tr>
</tbody>
</table>
1. Introduction

It all starts with the Sun. The uneven heating of the surface of the Earth from solar radiation – due to, among other things, the Earth’s axial tilt relative to its orbital plane, clouds, and a big span in heat capacity of different surfaces – results in rising and sinking air which further causes horizontal gradients in the air pressure. To balance these differences, the air starts to move, a movement that we normally refer to as wind. Anyone who has been outside on a really windy day can testify about the force and energy in the air on such days. Of course, it comes as no surprise that mankind have tried to extract this energy, convert it to electricity, and use it to power electrical machines.

As we are now facing a potential future energy crisis (Chevalier 2009; IEA 2022) it might as well be that wind power is a big part of the solution to the problem. Among the seventeen Sustainable Development Goals (SDGs) stated by the United Nations (UN 2023), there is a target to strive for environmentally friendly, reliable, and affordable electricity production globally. Wind power, together with e.g., hydropower and solar power, fulfills all those requirements and the total installed capacity is expected to increase rapidly in the coming decades (IEA 2022). In 2001, the total installed wind power capacity, both onshore and offshore, was 24 GW. By the end of 2022 it reached 906 GW, see Fig. 1.1 (GWEC 2023). Projections are to reach 2 TW before the end of 2030. At the time of writing this thesis (2023), wind power accounts for a share of approximately 12–13% of the total global electricity consumption (IEA 2023). For early adopters, such as Denmark – where the first offshore wind turbines were in operation in the Vindeby Offshore Wind Farm already in 1991 – wind power nowadays accounts for 40–50% of the annual electricity consumption (IEA Wind TCP 2023).

1.1 Why offshore wind energy?

Although, to this date, most of the wind turbines have been installed on land – accounting for 93% of the global wind energy production in 2022 (Fig. 1.1, GWEC 2023) – offshore wind power is becoming increasingly attractive as technology advances, making it possible to construct wind turbines that extract the energy from the wind over water with higher revenue than before. The two main reasons for placing a wind turbine offshore instead of onshore are: (1) the generally higher wind speeds offshore than onshore (neglecting mountain
Figure 1.1. Total installations (in GW) of onshore and offshore wind energy for the period 2001–2022. The percentages on top of each bar represents the share of offshore wind energy installations for that year. Data from GWEC (2023).

The main reasons that offshore turbines can be bigger include that the change in wind speed over the rotor generally is lower – due to lower surface friction over water compared to land – causing lower shear stress on the turbines, and that the visual and noise disturbance for humans is almost negligible if the turbines are erected far offshore (Haggett 2008). In addition to that, it is also possible to plan for larger wind farms offshore as considerations to topography are minimized. However, both the bathymetry (the depth of the sea) and the distance to the coast play important roles when assessing the viability of a site (Schwartz et al. 2010; Martinez and Iglesias 2022). With the technology of today, fixed tower wind turbines are possible to install at depths down to approximately 50 m. Floating offshore turbines, on the other hand, can be deployed in the deep sea, and although most installations today are at 50–300 m depth, there is potential to go down to 1,000 m in the future (Bilgili and
Alphan 2022; Dutton et al. 2023). A longer distance to the coast implies higher costs for longer underwater cables to connect the park to the land-based grid and also increases costs for maintenance (Swider et al. 2008; Martinez and Iglesias 2022).

The generally higher wind speed and the larger rotor areas offshore gives an advantage over onshore turbines. Since the energy of the wind is proportional to the cube of the wind speed, as seen in the wind power equation

\[ P = \frac{\rho A u^3}{2} \]  

(1.1)

where \( P \) is the power of the wind across the rotor area \( A \) swept by the blades, \( \rho \) is the density of the air and \( u \) is the wind speed; a stronger wind and a larger area has a major impact on the available power (Emeis 2018). However, it is impossible to extract all the energy from the wind and the theoretical maximum, called the Betz limit, is to extract 59% of the wind energy content in the air upstream of the turbine. Due to technical limitations, the Betz limit can, unfortunately, never be reached and turbines of today are, at their best, extracting approximately 45–50% of the energy in the wind, see Fig. 1.2 for an example.

The expected power production of a wind turbine is given by a power curve (example in Fig. 1.2), provided by the turbine manufacturer. The rotor starts spinning and producing electricity at the cut-in wind speed, usually in the range of 3–4 m s\(^{-1}\). The power production is then more or less following the cubic relation to the wind speed (Eq. 1.1), until the rated output power is reached, typically at 8–15 m s\(^{-1}\). The increase in power production is related to an increase in rotational speed of the wind turbine blades and to avoid extreme loads and structural damages, the same rotational speed and power output is kept until the cut-out wind speed, usually between 20 and 25 m s\(^{-1}\), is reached and the turbine is stopped. When denoting a wind turbine in this thesis, it is the rated power that is given.

To maximize the revenue from a wind farm, finding the site with the best wind conditions is a key factor, and here numerical models come into play, creating wind maps of the wind resource on both a regional and local scale. As inferred by Eq. 1.1 and the power curve in Fig. 1.2, even small errors in the estimated wind resource in an area can be crucial to determine if a site assessed for new wind farm development is profitable or not. An error of less than 1 m s\(^{-1}\) in the estimate of the annual wind resource can lead to several millions of dollars lower annual revenue for a wind farm (Banta et al. 2013). Thus, it is important to validate the models and be aware of their accuracy in describing both the general wind climate, but also their performance in different weather situations. After construction, when a wind farm is in the operational phase, the stochastic behaviour of the wind and uncertainties in wind speed forecasts on a short time scale translates into a financial risk for traders on the energy market (Moreno et al. 2012; Xiao et al. 2021).
As noted in the above, the distance to the coast is one of the determining factors when planning for offshore wind farms and an increase in distance typically implies an increase in cost. As a result, many offshore wind farms are constructed in coastal waters, although winds are normally slightly lower than far offshore. The Baltic Sea, see Fig. 1.3b, is a semi-enclosed sea in northern Europe that in many ways is optimal for offshore/coastal wind power installations (Sempreviva et al. 2008). Autumn and winter are the windiest seasons and this time of year cyclonic activity increases. In summer however, high pressure systems can be long lived, causing blocking situations and extended periods of weak winds in the entire basin (Nilsson et al. 2016). The wind direction is most often from west or southwest. The average wind speed at 150 m, a typical hub height for modern offshore wind turbines, is around 8–10 m s\(^{-1}\) and the average depth of the basin is 50 m, however reaching a maximum of 450 m in the Baltic proper (Jakobsson et al. 2019). The distance to the coast from anywhere in the basin is always less than 150 km.

Nine countries (Denmark, Sweden, Finland, Russia, Estonia, Latvia, Lithuania, Poland and Germany) have political borders to the sea, making it an interesting area for green energy production for many stakeholders. To date, the Baltic Sea wind power potential has mostly been utilized in the southwest, with Denmark and Germany having the highest installed capacity (Costanzo et al. 2023). However, many wind farms are now in the planning stage in all parts of the basin.

Before getting all permissions needed to start construction, there are many aspects to take into consideration, apart from the economical analysis. Animal life must be protected (Leung and Yang 2012) and shipping and aerial routes as well as military interests taken into account. Also, noise and visual disturbances
Figure 1.3. Overview of all sites analysed in this thesis. Black dots: sites with lidar measurements; red dots: sites where only model data were used. From the global overview in panel (a), there is a zoom in on the Baltic Sea area in panel (b), and a further zoom in on the small island Östergarnsholm, just east of Gotland, in panel (c).

for the people living in the coastal zone must be addressed, see e.g., Bishop and Miller (2007).

Under ideal circumstances, the wind speed profile displays an approximately logarithmic form in the atmospheric boundary-layer, i.e., the lowest (approximately) 100–1000 m of the atmosphere. However, the large land areas surrounding the Baltic Sea give rise to many local and mesoscale wind phenomena and complex wind profiles are more common here than far offshore (e.g., Smedman et al. 1996; Barthelmie et al. 2007; Svensson et al. 2016). For example, wind speed profiles with a local maximum, referred to as low-level jets (LLJs), are frequently occurring over the Baltic Sea, especially in late spring and summer (physical explanation given in Sect. 2). An example of an
LLJ wind profile is illustrated in Fig. 1.4a, indicating that the jet core typically is interfering with the rotor plane, causing a sharp transition from positive to negative wind shear over the rotor. Apart from the LLJs, there are also other types of non-ideal wind speed profiles that differ from the conditions that could be expected from a logarithmic wind profile (Kettle 2014; Møller et al. 2020), such as low-level minima (LLM) and negative wind speed profiles.

![Figure 1.4. Conceptual illustrations of a low-level jet (panel a) and of a wind profile wind strong directional shear (panel b). Both profiles are compared to generic ideal wind profiles. Figure adapted from the graphical abstract of Paper II.](image)

Furthermore, not only the speed, but also the direction in the wind profile can behave differently in coastal areas compared to far offshore (Miller et al. 2003; Kalverla et al. 2017). Due to the general increase in wind speed with height, associated with lower drag forces from turbulent motions generated by the surface friction, there is often a gentle clock-wise turning (Northern Hemisphere) of the wind with height, typically referred to as the atmospheric Ekman spiral (Ekman 1905). However, the different processes occurring in the coastal zone can cause the wind direction profile to behave differently, as illustrated in Fig. 1.4b. The change of wind direction over the rotor causes the wind to not be perpendicular to the rotor plane everywhere, which in turn results in a lower power production than what could otherwise have been expected (Murphy et al. 2020).

Compared to ideal wind profiles, both the non-ideal wind speed and the non-ideal wind direction profiles result in a change in structural and aerodynamic loading on the turbines and also changes the behaviour of the wake behind the turbine (Gutierrez et al. 2017; Gutierrez et al. 2019; Doosttalab et al. 2020; Gadde and Stevens 2021). This in turn affects how much electricity that is generated from the wind farm as a whole.
1.2 Aim of this thesis

This thesis aims to answer the following three research questions:

1. What are the characteristics and predictability of non-ideal wind profiles over the Baltic Sea?

2. How good are state-of-the-art models in describing the Baltic Sea wind conditions?

3. Can short-term power production forecasts for an offshore/coastal site in the Baltic Sea be improved using post-processing methods?

To accomplish this, seven studies have been performed, in the remainder of this thesis referred to as Papers I–VII. Although the aim of this thesis primarily is to characterize the wind conditions relevant for offshore/coastal wind energy production in the Baltic Sea, also other sites – both onshore and offshore in other parts of the world – have been analysed, see Fig. 1.3. Model data from reanalyses/wind atlases describing past conditions and from forecasting models predicting the near future as well as observational data have been used for these purposes. For the observations, light detection and ranging (lidar) instruments have been the key source providing actual measurements of the wind speed profile at heights relevant for wind power. Below follows a brief overview of the main contents in the different studies, and in Table 1.1, the data sets that were used and the main objective of each study is summarized.

In Paper I, different non-ideal wind speed profiles (LLJs, LLMs, and negative profiles) were defined and characterized using lidar data from Östergarnsholm (Fig. 1.3bc). The conditions when the non-ideal wind profiles appeared were analysed using observations from a meteorological mast and a wave buoy located nearby. Furthermore, the turbulent spectra and the differences in spectral values at low frequencies (large turbulent eddies) were studied and discussed in terms of shear sheltering, see Sect. 2.1.

Paper II is a validation of four recent data sets that are commonly used to describe the present climate in Scandinavia. To validate the data sets in terms of wind conditions, lidar observations from four sites in the Baltic Sea were used. The study focused partly on the general wind profile, but also on the agreement between models and observations in case of LLJs.

Single-level variables from the most recent version of the European Centre for Medium-Range Weather Forecasts (ECMWF) global reanalysis (ERA5), were used in Paper III as a proxy to investigate if – by only having surface observations at hand – it is possible to predict the speed and shape of the wind speed profile. The study made use of measurements from Utö in the Baltic Sea, but also from meteorological platforms in the southern part of the North Sea and in the U.S. Northeastern Atlantic Coastal Zone, see Fig. 1.3ab.
Paper IV is an attempt to find the optimal definition of the LLJ, applying a global perspective, see Table 3.1 and Fig. 1.3ab. Data from both onshore and offshore sites as provided by ERA5 were used.

Papers V and VI focus on the change of wind direction with height. In Paper V, an ERA5 climatology of strong directional shear (i.e., the change of wind direction over a turbine) over Scandinavia and the Baltic Sea was generated and validated against lidar observations from Utö and Östergarnsholm and the meteorological variables that are the most important factors deciding the strength of the directional shear were identified. In Paper VI, a new metric to automatically identify sea and land breezes in output from operational numerical weather prediction (NWP) models is suggested and the performance of the metric was tested both in case studies and in a multi-year overview, using archived NWP data.

Finally, in Paper VII, different statistical methods were applied to post-process NWP data to try to improve predictions of wind speed at hub height and, consequently, of estimated wind power production. The performance of the different statistical methods were evaluated in terms of governing meteorological conditions, and a recommendation of which method to apply in which condition was given.
Table 1.1. Overview of the data sets and main objectives for Papers I–VII.

<table>
<thead>
<tr>
<th></th>
<th>Observational data</th>
<th>Model data</th>
<th>Main objective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lidar</td>
<td>Met mast</td>
<td>Buoy</td>
</tr>
<tr>
<td>Anholt</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>ASIT</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>FINO2</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>MMU</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Utö</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Östergarns-holm</td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>Östergarns-holm</td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>ERA5</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>MERRA2</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>UERRA</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>NEWA</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>HARMONIE/AROME</td>
<td></td>
<td></td>
<td>O</td>
</tr>
</tbody>
</table>

Paper I: O
Paper II: O
Paper III: O
Paper IV: O
Paper V: O
Paper VI: O
Paper VII: O
2. Theory

In this chapter, the physical processes associated with the formation of non-ideal wind speed and wind direction profiles are described. Also, a general overview of the two machine learning (ML) algorithms – the random forest (RF) and the neural network (NN) that are applied in Paper III (RF and NN) and Paper VII (RF) – are given.

2.1 Processes affecting the wind profile

In Fig. 2.1, an overview of different mechanisms that can affect the wind speed and wind direction profiles at heights relevant for wind power are presented. High up in the atmosphere, where the air flow is unaffected by the resistance from the terrain, the wind speed and direction is a result of the balance between the pressure gradient force and the Coriolis force. The resulting wind is called the geostrophic wind. However, closer to the surface of the Earth, the topography of the landscape, the vegetation, waves on the oceans and man-made constructions influence the wind. The frictional force exerted by the ground surface on the air above causes the wind to slow down in the surface layer of the atmosphere. As a consequence, there is a change in both wind speed and wind direction with height (Ekman spiral), see the illustration in Fig. 2.1a.

Turbulence, the stochastic distribution of different sized eddies within the air flow, characterizes the atmospheric boundary-layer and is crucial in vertical re-distribution of heat, moisture and other atmospheric variables (Stull 2012). Turbulence can be generated mechanically from the interaction of the air with the surface roughness elements and by buoyancy due to differences in density of air parcels having different temperatures. In the textbook example, conditions are unstable over land during a sunny summer day. When the sun heats the land, which in turn heats the air at the surface, the air starts to rise as the warmer air has lower density than the colder air aloft. This buoyancy creates a turbulent vertical mixing of the air. Throughout the day, as the heating continues, the turbulent eddies grow larger and push the top of the boundary-layer upward, making it grow. The onset of stable stratification is marked by the evening transition. When the sun sets, the net radiation at the surface switches from positive to negative, indicating a cooling of the surface, and the magnitude of the turbulent movements starts to decrease. Turbulence can still be generated mechanically, but is dampened by buoyancy, and existing larger eddies cascade down to gradually smaller eddies, until eventually ending
Figure 2.1. Illustrations of different mechanisms that can affect the vertical profiles of wind speed and direction: (a) surface friction causes the wind to slow down and turn close to the surface, (b) the decrease in surface friction going from land to sea can cause the wind speed to increase and a low-level jet can be created, (c) the sea breeze circulation changes the wind speed and direction, (d) effects on the air flow due to complex terrain, (e) swell waves can impact the wind profile, possibly creating a low-level minimum in the wind speed profile, (f) effects on the wind due to frontal passages. Panel (c) is reproduced from Paper VI, © American Meteorological Society. Used with permission.

up as heat. As a consequence, also the boundary-layer decreases in height. Near-neutral conditions are typically a transition state going from stable to unstable conditions, or vice versa, but can also be long lasting, e.g., in windy and cloudy weather. In neutral conditions, there is a balance between the generation of new turbulent eddies and degradation of old eddies. As the properties of the boundary-layer are different from the properties of the free atmosphere above, this also implies that the top of the boundary-layer can be associated with a change in both wind speed and direction. In stable conditions, when the boundary-layer height (BLH) can be as low as 300 m, or even lower, this transition can occur at heights swept by the wind turbine blades.

The transition between an unstable and a stable boundary-layer is not only associated with the diurnal cycle but can also be generated spatially by non-homogeneous surface conditions. A typical example – that is of high relevance for the Baltic Sea wind conditions – is when warm air is advected from the land sector out over the comparably cooler sea (Fig. 2.1b). This typically happens over the Baltic Sea in spring and summer when winds are from the south or west, bringing warmer continental air over the water that is still relatively cold after the winter. The change from unstable to stable conditions at the coastline leads to a suppression of turbulent transport of momentum, which causes a momentum unbalance between the forces exerted on the air. This, in turn, results in a speed-up of the wind, which can form an LLJ offshore. Similar physical processes appear onshore in the formation of the nocturnal LLJ, as a cause of the diurnal cycle and nighttime cooling of the surface. As the wind speed of the LLJ increases during the night, the Coriolis force deflects it to the
right, creating a spiraling movement of the wind; a phenomenon known as an
inertial oscillation. For more details on the physical processes involved in the
formation of the LLJ, we refer to Blackadar (1957), Holton (1967), and van de
Wiel et al. (2010).

The pronounced change in the wind speed with height, i.e., the wind shear,
associated with the LLJ, switching from positive shear below the core to
negative shear above the core (see Fig. 1.4a), can act as a blocking layer,
potentially limiting the turbulent transport through the layer. The theory of
shear sheltering, primarily developed for engineering purposes such as flows in
pipes (Hunt and Durbin 1999), have been tested for atmospheric applications
to see if LLJs have any impact on the turbulent transport, but results are
inconclusive. Smedman et al. (2004) found indications that the shear sheltering
theory was applicable for the Baltic Sea LLJ and Prabha et al. (2008) came to
similar conclusions for a forested site in Maine (USA). However, the results
were questioned by both Karipot et al. (2008) and Duarte et al. (2012) assessing
shear sheltering associated with LLJs for onshore sites in the U.S., even seeing
increases in turbulence intensity and enhancement of variances and covariances
at low frequencies (large turbulent eddies) in the layer with positive shear
below the LLJ core. In Paper I, the behaviour of the turbulent spectra at low
frequencies in case of LLJs compared to ideal wind speed profiles was analysed
and the conclusions from Smedman et al. (2004) revisited.

In the coastal zone, differences in heat capacity and different response to
heating and cooling of the land surface and the water surface by solar radiation
can create mesoscale wind patterns such as the sea breeze (Fig. 2.1c) and the
land breeze (Simpson 1994; Miller et al. 2003). In the textbook example,
during sunny summer days the land surface heats up much faster than the sea
surface. The heating of the land initializes convection and a local low pressure
is generated. Since the air pressure over the sea is higher in comparison, the
pressure gradient force is directed from the sea towards land, resulting in a
wind, the sea breeze, blowing from the sea towards the shore. At higher levels,
a return flow is established, directed from land towards the sea. During night,
when the land surface cools faster than the sea, there can be a switch in the
direction of the circulation, resulting in a wind directed from land towards the
sea at low levels, i.e., the land breeze. While the sea and land breezes mainly
are associated with a change in wind direction with height – which is used
for automatic identification of sea/land breeze circulations in Paper VI – the
mesoscale wind pattern can also create LLJs (Kottmeier et al. 2000; McCabe
and Freedman 2023). The change of wind direction with height, i.e., directional
shear, where the sea/land breeze is one of many possible causes, was studied in
detail in Paper V.

As the wind has to adapt to surface features, such as topography and vegetation,
complex terrain can also lead to complex wind profiles, as illustrated in
Fig. 2.1d. For example, in mountainous areas, the wind is squeezed to go either
around or above the mountains, leading to local increases in the wind speed
and wave patterns in the wind field propagating downstream of the mountain. The wind can be channelled in valleys and due to differences in air density, katabatic winds can be formed transporting cooler and heavier air down the sides of the mountains and into the valley (for a coastal perspective, see e.g., Li et al. 2007). The air can be trapped in cold air pools in the valleys, leading to long-lasting stable conditions (Luiz and Fiedler 2023). As such, the complex terrain in the mountains frequently causes wind profiles with high directional shear (Paper V) and also non-ideal wind speed profiles. Similarly, coastlines also locally affect the direction of the wind and coastal LLJs can appear (Grisogono et al. 2007; Ranjha et al. 2013).

The influence of sea waves on the wind speed profile, Fig. 2.1e, is discussed in Paper I. As the wind blows, the downward transport of momentum brings energy into the sea, creating waves. However, in the case of swell, i.e., waves that are not generated locally by the wind, there can be a positive momentum flux, directed from the sea to the atmosphere. As more energy is injected into the atmospheric surface layer, this can lead to an increase in wind speed from below. This, in turn, can cause non-ideal wind speed profiles to form, such as negative profiles, where the wind speed decreases with height, or LLM, where the wind speed decreases with height up to a certain level, and then starts to increase again, creating a local minimum in the wind speed profile. For a deeper discussion on the impact on the wind profile from swell, see e.g., Hanley and Belcher (2008), Sullivan et al. (2008), Högström et al. (2009), Semedo et al. (2009), Smedman et al. (2009), and Wu et al. (2017).

Apart from all the local and mesoscale phenomena that could cause non-ideal wind profiles, also weather patterns on the synoptic scale can be associated with these type of wind profiles (Fig. 2.1f). For example, the sharp vertical transition between air masses when a pronounced cold front passes can give rise to steep gradients in both wind speed and wind direction (Kotroni and Lagouvardos 1993). Also, trough lines or the presence of convective cells with a high degree of rapid turbulent motions can cause non-ideal wind profiles to appear.

2.2 Machine learning methods

In Papers III and VII, machine learning (ML) was applied to find patterns in the data to (1) create predictions of how the wind speed changes with height (Paper III), and (2) to improve forecasts of power production over the coming 24 hours (Paper VII). Here, a brief overview and a background to ML models are provided.

Generally speaking, ML models are advanced non-linear statistical models that identify characterizing patterns in the training data which, in this context, is a large set of different atmospheric variables (Wilks 2011; Dueben et al. 2022). A model is trained with a specific task in mind, for example to improve
the forecast of wind speed at hub height (Paper VII). The ML model can be supervised, where the correct answers are provided for the training data set, or unsupervised, where the model automatically creates data clusters based on similarities in the data. Only supervised ML models are implemented in this thesis. The time series of the predictors, i.e., the atmospheric variables used as input to the model, and the predictand, i.e., the correct answer, is divided into three groups: training, validation, and test data. Usually, the training subset accounts for 60–80% of the data and the validation and test subsets for 10–20% each. In Paper III, some data were excluded at the transitions between training and validation or test periods to minimize the problem with high temporal auto-correlation of the atmospheric variables.

The ML model is generated to optimize for a cost function or a selected score, e.g., the root mean square error (RMSE), based on the training data set. The model is then validated using the validation data set that was kept hidden from the model during training. To optimize the model, it is important to find the best set of predictors to include and to fine-tune the hyperparameters, which are model specific settings (more details below). In this thesis, the optimal set of predictors was found using the forward selection method (Papers III and VII) and the backward elimination method (Paper III). In short, using the forward selection method, predictors are added one by one in the order that they help to improve the performance of the ML model the most when validated. For the backward elimination method, starting with all predictors, the predictors that deteriorate the performance of the model the most are removed one by one. For more details, see Kohavi and John 1997; Mao 2004. After reaching an optimal set of predictors and after fine-tuning the hyperparameters, the ML model is ready to be put to the test and to predict the results for data in the test period, that have been kept hidden during both training and validation. The result for the test period is what is finally presented as the results for the study.

Two different ML methods have been applied in this thesis, the random forest (RF) in Papers III and VII and the neural network (NN) in Paper III. Below follows a brief overview of the two techniques.

The RF builds on a large group of individual decision trees (Breiman 2001) and has been used before for wind power forecasting applications by e.g., Lahour and Slama (2017) and Vassallo et al. (2020), with promising results. Each tree sees a random subset of the data and at each node in the tree a decision is made if a new branch should be created or if this is a final node (a leaf). The decision is based on the number of observations at the node, i.e., the leaf size. If there is a decision to create a new branch, then the threshold of the variable that creates the best split is found, sending all data in the time series associated with higher values of the variable along one branch, and all data associated with lower values along the other branch. To get the final prediction from the RF, the results from all the individual trees are combined, typically taking the median or mean value of all trees.
The RF can be set up for classification problems, i.e., to predict a category (as in Paper III), or for regression problems, i.e., to predict a value (as in both Paper III and Paper VII). For classification problems, a cost matrix can be established. This was done in Paper III, penalizing the model differently for predicting that the wind profile was ideal when it in reality was an LLJ, than for predicting an LLJ when in reality it was an ideal profile. To optimize the performance of the RF, different values of the minimum leaf size and the number of trees in the forest can be adjusted. The effect on the results for improving power production forecasts by changing these parameters was thoroughly tested in Paper VII.

Just like the RF, the NN can model complex and non-linear relationships in the data and has proven to give good results for wind power applications (e.g., Wang et al. 2021). The NN consists of an input layer, an output layer, and a number of hidden layers in between (LeCun et al. 2015). The layers consist of neurons that are connected with synapses for transferring information. The information propagates forward through the system, from the input layer towards the output layer, and different weights for the neurons along the way decides how the information should be combined. The NN is typically considered to be a strong ML method that results in more accurate predictions than the RF, but requires a large data set to be fully trained and can be unstable and difficult to tune for smaller problems (Wang et al. 2017). This was also seen in Paper III.

As both the RF and the NN are associated with some stochastic behaviour, it is interesting to compare the models, not only to see which model that gives the most accurate results, but also to study if the results, e.g., of which are the important predictors, are consistent between models. This was investigated in Paper III.
3. Materials

This thesis builds upon both observations and model data, sometimes used in combination and sometimes separately, as shown in Table 1.1.

3.1 Observations

A few different techniques exist to measure the wind profiles at heights relevant for wind power, for example by (1) direct measurements such as on high meteorological masts equipped with anemometers, by using radio soundings, or by using drones, and (2) by remote sensing such as lidar or sound detection and ranging (sodar) instrumentation placed on a platform and scanning the air volume above. In this thesis, only lidar measurements were used for observations of the wind profile. However, in Paper I, wind observations performed at 10 m height in a meteorological mast were also used.

3.1.1 Wind profile observations

A lidar wind profiler works by emitting a laser beam in at least three different directions, typically at 30° from the vertical. The electromagnetic signal is backscattered to the device from interaction with aerosols and by calculating the Doppler shift in the signal, the wind speed can be inferred. The lidar can have either a continuous or a pulsed laser beam, both with their own advantages and disadvantages. For the continuous lidar, the laser beam is focused at one specific height at a time, allowing a narrow focus with a low uncertainty at lower heights, but a wider focus and a larger measurement volume (and thus a greater uncertainty) higher up (Mikkelsen 2009; Peña et al. 2015; Svensson et al. 2019a). For example, for the Östergarnsholm lidar, two thirds of the backscatter is expected to be within ±15 m at 100 m height, while at 300 m height, the measurement volume extends to ±150 m. A pulsed lidar, on the other hand, emits laser pulses and is able to keep the same measurement volume, typically ±15 m, at all heights.

In general, the technique of measuring the wind speed using a lidar instrument works well in most weather conditions, but the signal can be attenuated by fog and low clouds, causing a risk of systematic under-representation of those weather conditions. If there is a thin cloud layer somewhere within the measurement volume, it could be expected that the received signal will be
Table 3.1. Overview of the six lidar data sets used in this thesis.

<table>
<thead>
<tr>
<th>Lidar model</th>
<th>Anholt</th>
<th>ASIT</th>
<th>FINO2</th>
<th>MMIJ</th>
<th>Östergarnsholm</th>
<th>Utö</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leosphere WindCube V2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Halo Photonics StreamLine</td>
<td></td>
</tr>
<tr>
<td>WindCube V2 / ZX300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporal resolution (min)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Average vertical resolution (m)</td>
<td>25</td>
<td>14</td>
<td>22</td>
<td>23</td>
<td>34</td>
<td>8</td>
</tr>
</tbody>
</table>

dominated by the backscatter from the cloud, resulting in a temporary bias in the data. All lidar data used in this thesis were quality controlled by the data provider. Additional quality control was then applied to ensure the quality of the data in respect to the aim of each study. In all studies, the observed values from the lidar were considered as the ground truth, neglecting inherent uncertainties in the measurements.

As a lidar instrument is expensive, needs a platform, power supply, and regular maintenance, the availability of long-term measurements of high quality are sparse, especially offshore. In this thesis, lidar data from six sites have been analysed (Fig. 1.3), four in the Baltic Sea (which were all available lidar observations in the area at the time of the studies), one in the Dutch part of the North Sea, and one in the U.S. Northeastern Atlantic Coastal Zone. The sites are briefly described below and in Table 1.1 an overview of the measurements is presented.

**Anholt**
Measurements of the wind speed profile, performed by the pulsed Leosphere WindCube v2 lidar during 2013 and 2014 on a platform 1.5 km west of the Anholt wind farm, were utilized in Paper II. The distance from the lidar to the closest shoreline was at least 20 km. However, the distance to the closest wind turbine in the Anholt offshore wind farm was only approximately 1.5 km, and data when the wind was directed from the north-east-south sector (i.e., when the wind was directed from the farm towards the lidar) were filtered from the analysis, as possible wake effects from the wind farm could potentially have altered the wind speed profile. The lidar measured the wind profile at ten heights, ranging from 65 to 315 m above mean sea level (a.s.l.). Although data were recorded as 10 min temporal averages, only hourly data were studied in Paper II.
ASIT
The Air Sea Interaction Tower (ASIT) platform is operated by the Woods Hole Oceanographic Institution and is located just south of Martha’s Vineyard in the U.S. Northeastern Atlantic Coastal Zone (Kirincich 2020). From October 2016 to September 2021, a Leosphere WindCube v2 pulsed lidar was used, but then switched to a ZX300 continuous wave lidar in October 2021. As of September 2023, the measurements are still ongoing, but only data until the end of 2022 were used in Paper III. Measurements were performed on eleven heights, in the interval 51 to 200 m a.s.l. The 10 min temporal data were down-sampled to hourly data to match the temporal resolution of ERA5.

FINO2
In 2007, the meteorological mast FINO2 was deployed in the German part of the Kriegers Flak reef in the southern part of the Baltic Sea. From the location, it is at least 35 km to the nearest land surface in any direction. During a one year long campaign, from July 2012 to July 2013, the mast was accompanied by a Leosphere WindCube v2 pulsed lidar measuring the wind profile at ten levels, from 62 to 280 m a.s.l., with 10 min resolution (Schwenk et al. 2013). However, only hourly data from this campaign were used in Paper II.

MMIJ
The Meteorological Mast IJmuiden (MMIJ) is located on a platform in the Dutch part of the North Sea. In the period November 2011 to March 2016, the platform was equipped with a ZX300 continuous-wave scanning lidar, measuring the wind profile at ten levels between 90 and 315 m a.s.l. with 10 min resolution, later downsampled to hourly resolution for use in Paper III. Lidar data from the site has been extensively used to assess the wind profile for LLJs, see Kalverla et al. (2017), Kalverla et al. (2019), and Kalverla et al. (2020).

Östergarnsholm
Lidar data from the continuous ZX300 lidar on Östergarnsholm was studied in Papers I, II, and V. Östergarnsholm is a mostly flat island, approximately 2 km² in size and located 4 km east of the larger island Gotland in the Baltic Sea, see Fig. 1.3c. On the southern tip of the island, a meteorological mast is located and from December 2016 to June 2020 the lidar was measuring the wind profile on eight levels up to 300 m a.s.l. with 10 min resolution. The data were downsampled to 30 min in Paper I and to hourly resolution in Papers II and V. For a comparison of lidar measurements against the mast measurements at 30 m a.s.l., see Svensson et al. (2019a).

Utö
In the southwestern part of the Finnish archipelago, approximately 60 km from the mainland, the island Utö is located (Hirsikko et al. 2014; Laakso et al. 2018).
The area of the island is 1 km$^2$ and the highest point is less than 20 m above the sea level. Approximately 10 km away, the nearest islands of similar size are located. Since February 2015, a Halo Photonics StreamLine pulsed Doppler lidar has been measuring the wind profile with 8 m vertical resolution, starting at 35 m a.s.l. Data from the Utö lidar has been utilized in Papers II, III, V, and VII, and although the measurements reach much higher, the profile has been limited to 300 m a.s.l. in all studies as those are the heights most relevant for offshore wind power production. Also, the quality and availability of the data decreases rapidly above 300 m height. Data were stored as 15 min averages, but down-sampled to hourly data in all studies. For an extensive analysis of the Utö LLJ and more details about the lidar setup, see Tuononen et al. (2017).

3.1.2 Observations of air turbulence and sea waves

In addition to the lidar data, observational data from a Campbell CSAT three-dimensional anemometer, setup to measure the turbulence at 20 Hz resolution, mounted in the meteorological mast at Östergarnsholm were analysed in Paper I. In Paper I also the wave conditions as measured by a Directional Waverider buoy, located 3 km southeast of Östergarnsholm, were analyzed. The water depth at the site is 39 m. 30 min averages of the data were used, for the time period when all three instruments (lidar, sonic anemometer, and wave buoy) had simultaneous measurements of good quality. All measurements from Östergarnsholm, data handling and, quality control are described in Paper I. For further details about the Östergarnsholm instrumentation, see Rutgersson et al. (2020).

3.2 Models

As observations are limited in time and space, they have to be complemented by models providing three-dimensional data sets of atmospheric variables and covering several decades, both allowing for studies on regional scales and creating climatologies. Data from three types of models are utilized in this thesis: reanalyses, wind atlases and weather forecast models.

A reanalysis is a three-dimensional gridded description of the atmosphere at any given time in the past decades, providing information about the historical weather conditions at the grid points. As input, all available meteorological observations – ranging from meteorological masts to measurements performed by instruments on satellites – are combined, and a model is run to shape this information into an optimised and consistent three-dimensional gridded data set. As there is a high computational cost of running the calculations to create detailed global reanalyses, regional reanalyses can be run at a higher resolution than global reanalyses as they only cover a part of the globe.
A wind atlas is in many ways similar to a reanalysis, providing a gridded description of the atmospheric conditions. However, as the main objective for a wind atlas is to create a reliable data set for the wind climate, the wind atlas is optimised for this. This means for example that wind observations at heights relevant for wind power are crucial in validation when developing the model and that the climatological conditions are more important than trying to exactly match the reality hour-by-hour.

A weather forecast is generated using an NWP model, calculating the evolution of the weather based on the current conditions. Similar to the reanalysis, multiple sources of observations are assimilated to create a background field of high quality. By using the governing equations of the atmosphere and parameterize processes that are difficult to calculate explicitly, a forecast of the weather conditions in the near future is created, typically reaching up to a few days or two weeks ahead. As for the reanalyses, the NWP can be set up either as a global or a regional model. For a regional model, output from a global model is needed as input along the borders of the domain. The main benefit of the higher resolution, provided by regional models, is the possibility to resolve local and mesoscale processes in a more accurate way, allowing for sharper spatial contrasts due to e.g., coastlines. Both reanalyses and NWP models are normally set up to be able to represent the general weather situations well, which also means that they suffer in resolving extreme weather conditions. In the context of this thesis – the representation of stable conditions, which is one of the main ingredients for formation of non-ideal wind profiles – is of particular interest, and several studies have pointed out the difficulties of weather and climate models under these conditions, see e.g., Holtslag et al. (2013) and Sandu et al. (2013).

Below descriptions are provided for the three reanalyses, the wind atlas, and the NWP used in this thesis.

**ERA5**
The global reanalysis ERA5 (Hersbach et al. 2020) has a horizontal resolution of $0.25\degree \times 0.25\degree$, corresponding to approximately $17 \text{ km} \times 31 \text{ km}$ over the Baltic Sea. In the vertical, 137 hybrid sigma levels are used from the surface of the Earth to the top of the atmosphere (80 km height). The vertical distance between model levels is shorter close to the surface but increases with height. In the standard atmosphere, there are 10 levels lower than 300 m and 14 levels up to 500 m. However, as the exact heights of the model levels vary in time, their height at any given time step was calculated, based on the surface pressure and profiles of temperature and specific humidity. Data are available on hourly time steps from 1940 to present. However, the data set is known to have significantly higher quality from 1979 onward, as this year marks the start of the era when satellite imagery improved the input data for reanalyses dramatically (Bell et al. 2021).
In Papers II and V, the wind speed and wind direction profiles from ERA5 were, respectively, validated against lidar observations in the Baltic Sea. In Paper III, single-level data, i.e., data with spatial variation but no vertical extent, from ERA5 were utilized as a proxy for surface observations, and used to predict the wind speed profile. In Paper IV, ERA5 wind profiles from 1979 to 2022 were used to investigate different definitions of the LLJ

**MERRA2**
The global reanalysis from The National Aeronautics and Space Administration (NASA) is called the Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA2) and has a horizontal resolution of $0.625^\circ$ (longitudinal) $\times$ $0.5^\circ$ (latitudinal), corresponding to approximately 40 km $\times$ 55 km over the Baltic Sea (Gelaro et al. 2017). Data are available for download on 72 vertical levels from the surface up to 0.01 hPa. Apart from standard model output such as the 10 m wind speed, data are only provided on five heights in the lowest 500 m of the atmosphere. Data are presented in steps of three hours from 1980 onward. For comparison with lidar measurements in Paper II, MERRA2 data were linearly interpolated to hourly time steps.

**UERRA**
With a spatial resolution of 11 km $\times$ 11 km the regional reanalysis Uncertainties in Ensembles of Regional ReAnalyses (UERRA) covers all of continental Europe (Ridal et al. 2019). Wind data on eleven height levels from 15 to 500 m above the surface are available from 1961–2019 with hourly temporal resolution. Data assimilation was performed four times per day and short forecasts were run in between to generate the hourly data. UERRA wind speed profiles were validated against Baltic Sea lidar observations in Paper II.

**NEWA**
The New European Wind Atlas (NEWA) has a spatial resolution of 3 km $\times$ 3 km and data can be downloaded on 30 min temporal resolution on eight height levels between 10 and 500 m for the period 1989–2018 (Dörenkämper et al. 2020; Hahmann et al. 2020). To generate the data set, the Weather Research and Forecasting (WRF) model was run for ten separate regions covering the European Union plus Turkey, and then welded together. For each region, the model was optimized to give as accurate climatological wind conditions as possible at heights relevant for wind power. ERA5 data were used along the boundaries of the model domains and eight day forecasts (including one day of spin-up) were calculated and combined into a continuous time series. As for ERA5, MERRA2, and UERRA, the wind profiles from NEWA were assessed and validated against lidar observations in the Baltic Sea in Paper II.
Figure 3.1. Visualization of the resolution of the different numerical models studied in the thesis, showing the land/sea-mask for an area centered around the island Gotland in the Baltic Sea. Each black dot represents a grid point in the model. Figure modified from Paper II.

HARMONIE–AROME
The NWP model HIRLAM–ALADIN Research on Mesoscale Operational NWP in Euromed (HARMONIE) Applications of Research to Operations at MEsoscale (AROME) is a convection permitting operational weather forecasting model used for short-range forecasting by meteorological institutes in several countries around the Baltic Sea, including Sweden, Denmark, Finland, Estonia and Lithuania (Bengtsson et al. 2017). Scandinavia and the Nordic Seas are covered by the domain that has a resolution of 2.5 km × 2.5 km. The vertical resolution provides eleven levels in the lowest 300 m. Boundary conditions for the domain are provided by the ECMWF Integrated Forecasting System (IFS) global forecasting model. HARMONIE–AROME is run four times daily with a data assimilation cycle of three hours. In Paper VI, a metric to automatically identify sea and land breezes in HARMONIE–AROME was developed and in Paper VII different statistical post-processing methods were applied to improve the wind speed forecast and, subsequently the resulting power production forecast, from HARMONIE–AROME.

In Fig. 3.1, the horizontal resolutions of all models studied in this thesis are visualized, showcasing differences in the land/sea-mask for the island Gotland in the Baltic Sea (see also Fig. 1.3bc).
4. Methods and metrics

This chapter describes the main different methods that have been implemented in the studies that this thesis summarizes. Some of the methods are originally developed for the purpose of the studies, such as methods to identify non-ideal wind speed profiles and how to identify sea and land breezes in NWP model output, while other methods, such as e.g., the Jenkinson and Collison (1977) method to calculate the Lamb Weather Type (LWT) have been extensively used and validated over decades (Lamb 1972; Jones et al. 2012; Fernández-Granja et al. 2023). After the methods are described, an overview of some of the metrics used to quantify the performance of the models is presented.

4.1 Definitions of non-ideal wind profiles

A variety of different definitions of the LLJ have been used in the papers that this thesis consists of, and to this date, there is still no consensus which is the best definition to use to identify LLJs, see e.g., Baas et al. (2009), Kalverla et al. (2020), Aird et al. (2021), and Weide Luiz and Fiedler (2022). However, all definitions agree that an LLJ is a local maximum in the wind speed profile, see illustration in Fig. 1.4a, somewhere in the lowest hundreds of meters in the atmosphere and that – at least – the change in wind speed above the core is a determining factor if the profile should be considered an LLJ or not.

In Paper II, two thresholds (1 m s\(^{-1}\) and a 2 m s\(^{-1}\)) for the falloff above the core in the wind speed profile up to 300 m height were tested to define the LLJ. Also, in Paper VII, the 1 m s\(^{-1}\) falloff above the core was used. In Paper I, wind profiles were classified as strong LLJs if the change in wind speed was both 2 m s\(^{-1}\) and 20% of the core speed, both above and below the core. Similarly, for weak LLJs, the thresholds were set to 1 m s\(^{-1}\) and 10%, and the conditions for a strong LLJ should not be reached.

In Paper I, also transition profiles were defined. Transition profiles are considered transition states between ideal wind speed profiles and LLJs. Although they have a local maximum, the absolute and relative falloff above the core only had to exceed 0.5 m s\(^{-1}\) and 5% of the core speed. The exact same definitions of weak LLJs, strong LLJs, and transition profiles were also used in Paper V.

In Paper III, three criteria of increasing levels of strictness were applied to identify LLJs. For the least strict criterion, that was the one mostly used in Paper III, an LLJ was identified if the decrease in wind speed both above and below the jet core was at least 1 m s\(^{-1}\) and 10% of the core speed, i.e., the
same as the definition for weak LLJs in Papers I and V. In Paper III, also 2 m s\(^{-1}\) and 20%, i.e., the same definition as for strong LLJs, and 3 m s\(^{-1}\) and 30% LLJ definitions were tested.

**Paper IV** is an attempt to thoroughly investigate which is the best definition of the LLJ for wind energy applications. Two qualitatively different definitions were compared; one is the same as the definition to identify weak LLJs described above using the absolute and relative falloff above and below the core, the other using a threshold for positive and negative wind shear that has to be reached below and above the jet core, respectively. The shear threshold was set to ±0.01 s\(^{-1}\). The difference in which profiles that are identified by a falloff threshold but not a shear threshold is illustrated by the following two examples. Assume that wind profile data were provided with 30 m vertical resolution. If the wind speed at a given point in time, decreased by 0.6 m s\(^{-1}\) between one height level and the next, resulting in a shear of −0.02 s\(^{-1}\), the profile would be classified as an LLJ according to the shear definition, assuming that the shear criterion below the core was met, but not according to the falloff definition, since the 1 m s\(^{-1}\) falloff criterion was not met. However, if the wind speed decreased linearly by in total 1 m s\(^{-1}\) over five height levels (150 m) above the core, the shear was only −0.007 s\(^{-1}\) and the profile was classified as an LLJ according to the falloff definition, assuming that the falloff criterion below the core was met, but not according to the shear definition.

In Papers I and V, also definitions of LLM and negative profiles were implemented. An LLM was defined as a wind profile exhibiting a local minimum in the profile, with wind speeds at least 1 m s\(^{-1}\) and 10% stronger than the wind speed at the local minimum, both above and below the "core". If the wind profile was not classified as any other type of profile, and if the wind speed decreased by at least 1 m s\(^{-1}\) with height between the maximum wind speed below and the minimum, the profile was classified as a negative profile. All profiles in Papers I and V not classified as weak or strong LLJs, transition profiles, LLMs or negative profiles were classified as ideal wind speed profiles.

For a fair comparison of wind speed profiles from lidar observations and model data, the modeled wind speed profiles were interpolated to the heights of the lidar measurements by using Piecewise Cubic Hermite Interpolating Polynomials (PCHIP) on a logarithmic height scale.

4.2 Maximum directional shear over the rotor

In Paper V, the maximum change in wind direction over the rotor of the International Energy Agency (IEA) 15 MW offshore reference turbine (Gaertner et al. 2020) at wind speeds of operation (see Fig. 1.2), was investigated. The hub height of the turbine is 150 m, and with 120 m long blades the wind direction profile was thus assessed between 30 and 270 m a.s.l. The cut-in wind speed
of the turbine is 3 m s⁻¹ and the cut-out speed is 25 m s⁻¹, both thresholds referring to the wind speed at hub height.

The maximum change in wind direction over the rotor is bounded to 0–180° and the change of direction with height, i.e., if the winds were veering or backing, was not considered. For example, if the winds were veering up to a certain height, but then started backing, only the maximum change in wind direction for either the veering or backing part of the profile was registered, whichever gave the highest value. In the example in Fig. 1.4b, the maximum directional shear over the rotor is marked.

4.3 Sea and land breeze indices

In Paper VI, the sea breeze index (SBI) and the land breeze index (LBI) were developed for automatic operational identification of sea and land breeze circulations for coastal sites based on HARMONIE–AROME model output. The direction of the coastline is decided by the angle φ, as illustrated in Fig. 4.1. Apart from the direction of the coastline, the wind direction of the 10 m wind (α) and the direction of an upper wind (β) are also needed in order to calculate the indices. The SBI/LBI were only calculated if the following three criteria were all met:

1. **SBI**: the 10 m wind was at least partly directed from the sea towards land (i.e., α was within φ ± 90°)
   **LBI**: the 10 m wind was at least partly directed from land towards sea (i.e., α was within ±90° from φ + 180°)

2. **SBI**: the upper wind was at least partly directed from land towards sea (i.e., β was within ±90° from φ + 180°)
   **LBI**: the upper wind was at least partly directed from sea towards land (i.e., β was within φ ± 90°)

3. **SBI** and **LBI**: the upper wind was at least partly in the opposite direction as the lower wind (i.e., β was within ±90° from α + 180°)

If the above criteria for either SBI or LBI were fulfilled, the index was calculated as

\[ SBI = \cos(\alpha - \phi) \cos(\alpha + 180° - \beta) \]  \hspace{1cm} (4.1)

or

\[ LBI = -\cos(\alpha - \phi) \cos(\alpha + 180° - \beta) \]  \hspace{1cm} (4.2)

respectively. Note that the SBI and the LBI cannot be calculated simultaneously for a grid point, as criteria for both indices are never fulfilled at the same point.
in time. Values for SBI and LBI are in the range 0 to 1. As is seen in Eqs. 4.1 and 4.2, the highest values of SBI and LBI appear when the 10 m wind is perpendicular to the coastline, and when the upper wind is opposing the 10 m wind. For identifying the SBI in output from HARMONIE–AROME, model levels 51–38, roughly corresponding to 500 to 2000 m height and representing typical heights of the return flow in the sea breeze circulation, were considered for the upper wind, and the maximum SBI among all those levels was stored. For the LBI, model levels 46–62 were considered for the upper wind, corresponding to 100 to 900 m as the land breeze circulation typically is shallower than the sea breeze circulation (Holmer and Haeger-Eugensson 1999; Miller et al. 2003).

### 4.4 Classification of atmospheric stability

As described in Sect. 2.1, the atmospheric stability is one of the fundamental characteristics of the boundary-layer. Across the studies, two methods were used to classify the stability: (1) the bulk Richardson number $R_i^b$ (Paper VII), based on the vertical gradients of temperature and wind speed, and (2) the
stability parameter, \( z/L \) (Paper I), requiring only measurements at one height (\( z \)) and calculation of the Obukhov length (\( L \)).

For the bulk Richardson number, a positive (negative) value indicates stable (unstable) conditions, while values close to zero imply near-neutral stratification. Five stability classes were considered in Paper VII: strongly stable (\( \text{Ri}_b \geq 0.25 \)), stable (\( 0.05 \leq \text{Ri}_b < 0.25 \)), neutral (\( -0.05 \leq \text{Ri}_b < 0.05 \)), unstable (\( -1 \leq \text{Ri}_b < 0.05 \)), and strongly unstable (\( \text{Ri}_b < -1 \)).

To calculate the Obukhov length, used to classify the stability in Paper I, high frequency measurements of turbulent motions in the atmosphere are required, see details in Paper I. A five class stability classification was applied in Paper I: stable (\( 0.2 \leq z/L \)), weakly stable (\( 0.02 \leq z/L < 0.2 \)), near neutral (\( -0.02 \leq z/L < 0.02 \)), weakly unstable (\( -0.2 \leq z/L < -0.02 \)), and unstable conditions (\( z/L < -0.2 \)).

Note that the stability can change within different layers of the lower atmosphere, see Argyle and Watson (2014) for an offshore discussion, and thus, the stability classification might not be the same all the way from the surface up to 300 m, which is the top height of the wind profiles in most papers in this thesis. The \( \text{Ri}_b \) stability classification was calculated based on data from the two lowest levels in HARMONIE–AROME, approximately 12 and 37 m a.s.l., and the classification of the stability parameter \( z/L \) was based on measurements in the surface layer (10 m a.s.l.).

The BLH is strongly associated to the stability. Lower BLH generally appears in stable conditions and higher BLH in unstable conditions. The BLH as an output variable from ERA5, calculated from the bulk Richardson number following conclusions from the review by Seidel et al. (2012), was used in the analysis in Papers III and V. The BLH can – to some extent – be considered as a rough proxy of the general stability of the atmospheric boundary-layer, see Vogelezang and Holtslag (1996) for a discussion on these matters.

4.5 Turbulent spectra

In Paper I, the turbulent \( u \) and \( w \) power spectra were analyzed, where \( u \) is the horizontal component of the wind speed in the dominant wind direction during the averaging period of 30 min, and \( w \) is the vertical component of the wind speed in the averaging period. The size of turbulent eddies is closely linked to their frequency (\( n \)), with large eddies having low frequencies and vice versa. A normalized frequency (\( f \)) was calculated, taking the height of the measurements (\( z =10.4 \) m) and the wind speed into account,

\[
f = \frac{nz}{u} \tag{4.3}
\]

The normalized \( u \)-spectra, \( S_u(n) \) – normalized by removing all stability dependence so that all spectra coincide in the inertial subrange which allows
for easier assessment of variations in spectral values at low frequencies for different wind profiles – were calculated using the formula

\[ \hat{S}_u(n) = \frac{nS_u(n)}{u^*_2\phi_e^{2/3}} \]  

(4.4)

where \( u_* \) is the frictional velocity and \( \phi_e \) the non-dimensional dissipation rate of energy (Kaimal et al. 1972). For further details, see Paper I. Spectra for different types of wind profiles were compared at the selected low normalized frequency \( f = 0.01 \). Spectral values were interpolated to this frequency from neighbouring frequencies using linear regression in a log-log representation.

### 4.6 The Jenkinson and Collison method

The Jenkinson and Collison (1977) method is an objective way to classify any given synoptic weather situation into one of the 27 Lamb Weather Types (LWTs, Lamb 1972), using only the mean sea level pressure at 16 grid points in a specified layout. The Jenkinson and Collison method was used in Papers V, VI, and VII. In Paper VII the 27 LWTs were reduced to a set of only 11 weather types, following Demuzere et al. (2009).

In brief, using the Jenkinson and Collison method, the synoptic wind speed and vorticity representative for a focus area are determining the LWT. If both the wind speed and vorticity have low values, the LWT is classified as U (weak synoptic forcing). If the magnitude of the vorticity is lower than the magnitude of the wind speed, the LWT is classified as pure directional flow, i.e., one of the eight cardinal/inter-cardinal classes, depending on the wind direction. On the other hand, if the magnitude of the synoptic vorticity is greater than the synoptic wind speed, the LWT is assigned as either one of the eight directional anti-cyclonic hybrid classes (if the vorticity is negative) or one of the eight directional cyclonic classes (if the vorticity is positive). However, if the magnitude of the synoptic vorticity is more than twice the value of the synoptic wind speed, the LWT is classified as either pure anti-cyclonic conditions (negative vorticity) or pure cyclonic conditions (positive vorticity). A more detailed description of the Jenkinson and Collison method is given in Appendix A accompanying Paper VII.

The focus area where the classification is valid was set to the main part of the Baltic Sea in Paper V, to southern Sweden in Paper VI and over the northern part of the Baltic Sea, including the Bothnian Sea and Bothnian Bay in Paper VII. Deciding factors in the selection of the different focus areas were partly which region that was studied, but also limitations by the model domain.
4.7 Predictor importance

The main objectives in Paper III were to find the optimal set of variables to predict the wind speed in the profile and the shape of the wind profile using ML methods and to estimate how important the individual variables were in making an accurate prediction. As described in Sect. 2.2, both the forward selection and backward elimination methods were tested to find the optimal set of predictors.

Knowing the optimal set of predictors, the importance of each predictor was calculated by excluding predictors one at a time, re-training the ML model for each new set of predictors and quantify the corresponding decrease in performance. A large decrease in performance omitting a predictor indicated high importance of that predictor and vice versa. If $A$ is the score of the evaluation metric using the optimal set of predictors, and $\hat{a}$ is the score if a predictor is omitted, the predictor importance, $PI$, is given by

$$PI = \left| 1 - \frac{\hat{a}}{A} \right|$$

High values of $PI$ indicates important predictors and values close to zero indicates low importance. The formula to calculate $PI$ can be used both for metrics where an increase of the metrics marks an improvement, such as the Symmetric Extremal Dependence Index (SEDI), see details in Sect. 4.10, and for metrics where a decrease of the metric marks an improvement, such as the RMSE.

In Paper V, the predictor importance was calculated in a somewhat different way. Here a single classification tree was set up, not splitting the time series into training, validation, and testing but using all available time steps and all predictors at once. The predictor importance of each predictor was then estimated by analysing the risk associated with each split in the tree where the predictor was selected.

Also, in Paper VII, the predictor importance was analyzed in a different way. Here the forward selection method was used as an indicator of the importance of a predictor, i.e., in each iteration, the most important predictors were the first to be selected.

4.8 Calculating the power production

By using the power curve of a wind turbine, see Fig. 1.2 for an example, and knowing the wind conditions at a site, the theoretical power production can be calculated. Either the wind speed at hub height could be used straight away – as was done in Paper VII, considering a Siemens SWT-3.6-120 2.6 MW wind turbine with a hub height of 120 m (Siemens 2011) – or by calculating the rotor equivalent wind speed (REWS) taking an area-weighted average of the wind speed over the full span of the rotor, as was done in Paper IV.
The REWS have proven to give significantly different results in conditions with strong shear over the rotor, i.e., non-ideal wind speed profiles, and in cases of low turbulence intensity, i.e., in stable conditions (Wagner et al. 2014). In Paper IV, the REWS method was used, considering the IEA 15 MW offshore reference turbine (Gaertner et al. 2020) with a rotor disc covering heights from 30 to 270 m for the offshore sites, and a Vestas 150-4.2 MW (Vestas 2023) turbine covering 30 to 230 m for onshore sites. The change in wind direction over the rotor was not considered, see e.g., Murphy et al. (2020) for a discussion on the effect of directional shear on the REWS.

4.9 Methods for post-processing of HARMONIE–AROME forecasts

In Paper VII, several methods to improve power production forecasts, using the forecast wind speed at hub height, 120 m, from HARMONIE–AROME were tested:

- The D1/D2 mix forecast was created by combining the most recent forecast from HARMONIE–AROME with the forecast issued 24 hours earlier, but valid for the same time steps. Based on the performance in the training period, the two forecasts were assigned individual weights and then combined.

- In a similar way, the neighbourhood (NBH) forecast was created by combining forecasts from the four grid points closest to Utö, based on the performance of the individual forecasts in predicting the power production in different wind directions at the site of the lidar.

- Temporal smoothing was applied by averaging the wind speed in the forecast, using a running mean with a temporal window of three hours.

- Several linear regression methods were created by fitting first order polynomials to either all data, by comparing the forecast wind speed at hub height to the wind speed measured by the lidar, or to subsets of the data based on different variables, such as atmospheric stability, synoptic conditions or time of year.

- A random forest (RF) statistical model was fitted to atmospheric variables output from HARMONIE–AROME, as described in Sect. 2.2.

Also, a persistence forecast was implemented for comparison, keeping the last lidar observation of the wind speed from the day before (at 2300 UTC) as a constant prediction for the next 24 hours of evaluation. In the summary of Paper VII in this thesis, results from the linear regression methods and the
persistence forecast were omitted, as they were not fruitful in improving the forecasts.

4.10 Evaluation metrics

Apart from commonly used statistical metrics such as bias, the RMSE and standard deviation, a few metrics that are less frequently applied have been implemented in this thesis. Here follows a summary of two of them, the SEDI and the mean absolute error (MAE) skill score.

The SEDI was implemented in Paper III to optimize and evaluate the ML methods (RF and NN) regarding the task of accurately making a binary prediction of if an LLJ was present in the wind profile or not. Validating a binary prediction against observations, the result can be expressed in terms of number of hits ($h$), false alarms ($fa$), misses ($m$) and correct rejections ($cr$). Based on these numbers, the hit rate ($H$) and the false alarm rate ($F$) can then be calculated as

$$H = \frac{h}{h + m} \quad (4.6)$$

and

$$F = \frac{fa}{cr + fa} \quad (4.7)$$

and then the SEDI follows as

$$SEDI = \frac{\log F - \log H - \log(1 - F) + \log(1 - H)}{\log F + \log H + \log(1 - F) + \log(1 - H)} \quad (4.8)$$

The SEDI is suitable for deterministic forecasts of rare binary events (such as the occurrence of LLJs), is non-trivial to hedge and is independent of the base-rate of the relative frequency of the event (Ferro and Stephenson 2011). A SEDI of 1 indicates a perfect result, while 0 represents no skill compared to a random prediction following the climatology of the event. A negative SEDI indicates a prediction of lower value than a climatological prediction.

In Paper VII, the MAE skill score was used to assess the performance of the different post-processing methods, see Sect. 4.9. The skill score is defined as

$$\text{MAE skill score} = 1 - \frac{\text{MAE}_{pp}}{\text{MAE}_{NWP}} \quad (4.9)$$

where MAEpp is the MAE of the post-processed forecast and MAE_NWP is the MAE of the original NWP, i.e., HARMONIE–AROME in the case of Paper VII.
As is clear from Eq. 4.9, a MAE skill score of 1 represents a perfect forecast, i.e., exactly following the observations, and a MAE skill score of 0 implies that the post-processing resulted in no improvement over HARMONIE–AROME. A negative MAE skill score indicates that the forecast quality was deteriorated by the post-processing. In order to calculate the MAE skill score, all forecasts lengths (0–23 h) were combined and the error at each time step treated equally as no substantial decrease in forecast performance was identified over the length of the forecast.
5. Results

In this section, the main results from the studies that this thesis summarizes will be presented.

Paper I

In Paper I, observational data from Östergarnsholm (lidar, mast and wave buoy data) were analyzed to characterize non-ideal wind profiles and their properties. In Fig. 5.1a, the average wind speed lidar profiles for the different types of profiles, as described in Sect. 4.1, are shown. From the figure, it is clear that the average LLJ profiles appeared at similar wind speeds as the ideal profiles and that the speed at the core usually was somewhat higher than the average conditions for ideal profiles at the same height. Strong LLJs, did not only display a more pronounced local maximum in the profile, but were also typically associated with stronger winds. As the wind speeds for transition profiles, LLJs and ideal profiles at typical offshore hub heights are in the steep part of a typical power curve (Fig. 1.2), this implies that it is of high importance to have good knowledge about these profiles and their morphology to accurately estimate the power production. LLMs and negative profiles typically occurred in weaker winds, just above the threshold for cut-in wind speed.

In Fig. 5.1b, the average shear profiles are plotted. The shear profile affects the aerodynamic and structural loads on a wind turbine as well as the properties of the wake behind a turbine, as mentioned in Sect. 1.1 (Gutierrez et al. 2017; Gutierrez et al. 2019; Doosttalab et al. 2020; Gadde and Stevens 2021). The figure shows the switch from positive to negative shear with height for transition and LLJ profiles and from negative to positive shear with height for LLMs. As expected from the definitions, strong LLJs had higher absolute shear values than weak LLJs, both below and above the core. In turn, weak LLJs displayed higher absolute shear than transition profiles. Ideal profiles had on average positive shear, with the strongest shear on lower levels due to the effect from surface friction, see Sect. 2.1.

For the time period studied (December 2016 to June 2020), strong LLJs appeared 2%, weak LLJs 6%, transition profiles 8%, LLMs 1%, and negative profiles 4% of the time. During 80% of the time, the profiles were classified as ideal. Looking at the seasonal distribution, Fig. 5.1c, it is apparent that non-ideal wind speed profiles tended to appear in the spring and summer months. In the season April–July (abbreviated AMJJ), non-ideal wind profiles occurred 35% in common, with a peak of approximately 40% in May and June. In their peak months, strong LLJs appeared up to 7%, weak LLJs up to 12%, transition
Figure 5.1. In panel (a), the average wind speed profiles for profiles classified as strong LLJs, weak LLJs, transition profiles, LLMs, negative profiles, and idealized profiles are plotted. Panel (b) shows the average shear profiles for the different profiles. The seasonality of the non-idealized profile types is displayed in panel (c). In the legend, the total number of occurrences per profile is presented. Shaded areas around the lines in panels (a)–(c) mark the 95% confidence interval of the mean. In panels (d)–(f) the distributions of 10 m wind speed (all wind direction sectors) when the different non-idealized profiles were occurring are plotted, for the three different seasons DJFM (panel d), AMJJ (panel e), and ASON (panel f). The grey bars show the distributions for idealized profiles. In panels (g)–(i), the relative occurrence of all types of profiles in the three different wave age classes (only when the wind was directed from the open sea) is plotted for the three seasons. For each season, the bars for each type of wind profile add up to 100%. For panels (d)–(i), a minimum of 20 profiles in a season was used as a threshold, and thus data for LLMs are missing in ASON (panels f and i). Figure adapted from Paper I.
profiles up to 14%, LLMs up to 3%, and negative profiles up to 8% of the time. This is related to the high frequency of stable stratification of the atmospheric boundary-layer during this time of the year, see Sect. 2.1 for details. However, the governing synoptic conditions yield a large year-to-year variability and e.g., in May 2018 – dominated by atmospheric blocking and weak winds – non-ideal profiles were observed 60% of the time. In the fall season August–November (ASON), there was a decrease in relative occurrence while in the winter/early spring season December–March (DJFM), the relative frequency of non-ideal wind speed profiles increased month by month.

In the 10 m wind speed distributions, Fig. 5.1d–f, it is also clear that the average wind conditions were weaker in AMJJ than in DJFM and ASON. LLMs and negative wind profiles typically occurred when the 10 m wind speed was weaker than the median of the distributions for ideal profiles, as expected from Fig. 5.1a. Transition profiles, as well as weak and strong LLJs, typically appeared at wind speeds closer to the median of the ideal profiles distributions. However, if the 10 m wind speed exceeded 10 m s\(^{-1}\), it was very unlikely that a non-ideal wind speed profile would appear.

Based on the 10 m wind direction, the conditions at Östergarnsholm were classified as either representative for the open sea sector (if winds were from 45°–220°), affected by Gotland (if winds were from 220°–295°) or affected by Östergarnsholm itself (if winds were from 295°–355°), as shown in Fig. 1.3. The sector 355°–45° was excluded, as the turbulence measurements in the mast were disturbed by the mast itself when the wind was directed from this sector. When the wind was directed from the open sea sector, the wave conditions were classified as either growing sea, mixed sea or swell – depending on the wave age calculated from the phase speed of the dominant waves and the wind speed. In Fig. 5.1g–i, the relative occurrence of each type of wind profile in different wave conditions was calculated for the three seasons. Mixed sea and swell conditions were the most common sea states, and naturally it was in these conditions non-ideal wind profiles occurred most frequently. However, the frequency of LLMs and negative profiles during swell was unproportionally high, indicating that the upward momentum flux associated with swell, see Sect. 2.1, is likely a key factor in the formation of this type of non-ideal wind speed profiles.

As explained in Sect. 4.5, the median normalized \( u \)-power spectra were calculated and in the top row in Fig. 5.2, the \( u \)-power spectra is plotted against the normalized frequency in a log-log representation. Data were split into different categories based on wind direction, but only results for the open sea sector are shown here. Furthermore, data were split into different stability regimes (Sect. 4.4). Spectra for all types of non-ideal profiles were compared to those from ideal profiles. As a consequence of the normalization procedure, spectra coincided in the inertial subrange, but there was some variation in spectral values at lower frequencies (larger eddies).
Figure 5.2. In the top row, the median of the normalized turbulent $u$-power spectra is plotted against normalized frequency for the different types of wind profiles. Results are for the open sea sector and are divided into five categories based on the stability of the boundary-layer. A threshold of at least 20 occurrences per profile and category was used as a limit to include the statistics. In the bottom row, the distributions of the normalized turbulent $u$-power spectra at the selected normalized frequency 0.01 is plotted. Bottom and top edges of the boxes mark the 25th and 75th percentiles respectively, and the line in the box the median value. The 95% confidence interval of the median is given by the notches. The dots mark outliers and the whiskers reach the most extreme spectral values not considered outliers. Figure adapted from Paper I.

Distributions of the spectral values at the selected normalized frequency 0.01 are plotted for the open sea sector in Fig. 5.2 (bottom row). In all stabilities, there was a significantly lower median value of the spectra for transition and LLJ profiles compared to ideal-profiles. For the LLMs and negative profiles, no clear pattern could be seen as the median of the spectral distribution at the selected frequency sometimes was lower and sometimes higher than that of the ideal profiles. The lower spectral energy for large eddies at time steps with LLJs could possibly be attributed to the theory of shear sheltering, see Sect. 2.1, but further analysis of the turbulence at higher vertical levels are needed to justify this statement.

Paper II
In Paper II, the wind speed profiles from three state-of-the-art reanalyses (MERRA2, ERA5, UERRA) and one wind atlas (NEWA) were compared to lidar observations from four sites in the Baltic Sea, see Table 1.1 and Fig. 1.3b for an overview. In Fig. 5.3, the average wind speed profiles for all models are plotted for the four sites and compared to the profile as measured by the lidar at the site. At first glance, it is apparent that MERRA2 struggled resolving the wind profile accurately, having a negative bias of approximately 1 m s$^{-1}$ on the average profiles. Also, the vertical resolution in data available from MERRA2
**Figure 5.3.** Comparison of average wind speed profiles up to 300 m a.s.l. between lidar observations and model output at four sites in the Baltic Sea. Height levels where model data are provided are marked. The 95% confidence interval of the mean in the lidar profile is marked by the shaded area. Figure adapted from *Paper II*.

was much lower compared to the other models, only providing data on two levels between 35 and 300 m. Based on this, it was concluded that MERRA2 is not suitable for wind power applications and the model was excluded from further analysis in *Paper II*.

ERA5, UERRA, and NEWA were less easy to separate, although for Anholt and FINO2, ERA5 was generally the best model at all heights. In general, all models gave the best agreement for Utö, which also was the site with the longest temporal record, and the largest deviations were found for FINO2. The average shear was too low in NEWA compared to the lidar observations, while ERA5 and UERRA performed better in this regard. The shear distribution was further investigated in *Paper II*, but results are not included here.

The wind speed profiles were also assessed for local maxima, and considering a falloff threshold of 1 m s$^{-1}$ above the core, the data could be characterized into hits, misses, false alarms and correct rejections, comparing model output to observations. For all sites, the hit rate (Eq. 4.6) ranged from 12–18% for ERA5, 28–41% for UERRA, and 15–25% for NEWA. The false alarm rate (Eq. 4.7) was in the interval 1–3% for ERA5, 4–7% for UERRA, and 1–4% for NEWA. According to observations, falloffs greater than 5 m s$^{-1}$ occasionally appeared for all sites. Although rare, UERRA and NEWA sometimes predicted falloffs greater than 5 m s$^{-1}$, but in ERA5 such strong falloffs were never seen.

In Fig. 5.4, the distributions of LLJ core speeds as measured by the lidars and modeled by ERA5, UERRA, and NEWA are plotted for Östergarnsholm and Utö. Based on this it is evident that there was a wide range of core speeds possible; from just a few meters per second up to well above 20 m s$^{-1}$. In general, for all sites, UERRA was the model that best captured the median
Figure 5.4. Distribution of core speed of LLJs identified in the lidar observations and in ERA5, UERRA, and NEWA for Östergarnsholm (top) and Utö (bottom). Left and right edges of the boxes mark the 25th and 75th percentiles, respectively, and the square in the box the median value. The 95% confidence interval of the median is given by the notches. The dots mark outliers and the whiskers reach the most extreme values not considered outliers. Figure adapted from Paper II.

and the interquartile ranges of the distributions, and also UERRA was best at describing the extreme values. Given this – and having the highest LLJ hit rate – UERRA was considered to be the best choice for a Baltic Sea LLJ climatology, despite the relatively high false alarm rate compared to other models. For the general wind profile, it was a tight race between ERA5, UERRA, and NEWA and the best choice depends on the application in mind. However, as was previously mentioned, the underestimation of the wind shear is important to keep in mind if NEWA is selected.

Paper III
Single-level data from ERA5 were applied in Paper III to predict the wind speed in the vertical profile and the shape of the profile for three offshore/coastal sites with lidar measurements: ASIT, MMJJ, and Utö. Among all ERA5 single-level variables, a set of nine variables with low cross-correlation and presumably high relevance for the predictions were selected; the wind speed at 10 m (ws10), the wind direction at 10 m (wdir10), the sea surface temperature (SST), the mean sea level pressure (MSLP), the total amount of precipitation (precip.), the convective available potential energy (CAPE), the surface sensible heat flux (SHF), the net radiation (Rn), and the percentage of low cloud cover (LCC).
In Fig. 5.5, the variables were tested if they had a significantly different median in the distribution when LLJs occurred, compared to when the wind profile was a non-LLJ (Mann-Whitney significance test, 5% significance level). The wind direction at 10 m was not included in this test since its circular coordinates disqualify it from this type of analysis. Three criteria of increasing strictness were implemented to identify LLJs – falloff both above and below the core of at least 1 m s\(^{-1}\) and 10%, 2 m s\(^{-1}\) and 20%, or 3 m s\(^{-1}\) and 30% – and the seasonality of the relative occurrence for the different definitions and sites is plotted in the bottom row in Fig. 5.5. For all sites, despite the fact that they are in different regions (the U.S. Northeastern Atlantic Coastal Zone, the Dutch part of the North Sea, and the Baltic Sea), the same pattern was seen with an enhanced frequency of LLJs in late spring and summer. However, comparing the sites, large differences were observed, with LLJs over the Baltic Sea (i.e., Utö) being more frequent than for the other sites.

In general, LLJs appeared when the ws10 was low compared to climatology, when the SHF was high, and when there was a low amount of low clouds (low LCC). Furthermore, for Utö, the MSLP was usually higher when LLJs were forming, and thus associated with weaker winds. Also SST was higher than the median value for that month. Outside months of peak occurrence, LLJs usually appeared when conditions were low in precipitation and low in CAPE, at least for Utö.

The predictor importance for predicting the wind speed in the profile was calculated for the random forest (RF) and the neural network (NN), as described in Sect. 4.7. In Fig. 5.6, only the results of predictor importance for the RF are shown, but results were similar also for the NN. The ws10 and the SHF were by far the most important variables to predict the wind speed higher up. Other variables could however contribute with minor improvements to the prediction. Comparing the predictor importance for ws10 with SHF, ws10 had the highest predictor importance, i.e., the wind speed at a lower level – 10 m – was the best indicator for how strong the wind was higher up. The SHF is a good proxy of the stability of the atmospheric boundary-layer, which in turn describes the amount of turbulence and how quickly the wind speed changes with height. As seen in Fig. 5.6, the predictor importance for ws10 and SHF were fairly constant throughout the entire profile.

In Fig. 5.7, the RMSE reductions in the wind speed profile compared to ERA5 for the RF and NN generated profiles are plotted for the three sites. For ERA5 itself (dotted line in Fig. 5.7), there was a general increase in RMSE with height. Individual ML models were trained separately on each height on the set of predictors shown in Fig. 5.6. The RF model was generally showing greater RMSE reductions than the NN, which was trained following a similar protocol. Although ERA5 is based on a full-scale NWP model and the ML models only used a small set of predictors both the RF and the NN managed to lower the RMSE. However, as the ML models were trained on the wind speed from the lidar measurements, which ERA5 was not, the ML models were
Figure 5.5. For the three sites ASIT, MMII, and Utö, the results of the Mann-Whitney significance test is plotted month-by-month, testing three LLJ criteria with increasing levels of strictness. A circle is drawn if at least 30 LLJ samples are found using a criterion, and the size of the circle is scaled by the monthly relative frequency of LLJ for the site (bottom row). The circle is colored if there is a significant difference between the median of the LLJ and non-LLJ distributions of the corresponding ERA5 single-level variable for that month. Plus and minus signs below the circles indicate if the median in the LLJ distribution of the variable is shifted towards higher (+) or lower (-) values. Figure reproduced from Paper III.

optimized for the sites. In general, the performance of the ML models was better at lower levels, resulting in a 0–10% reduction, related to that mostly surface variables were used as predictors.

The second task in Paper III was to predict the shape of the wind speed profile: if it was an LLJ or not (using the 1 m s$^{-1}$ and 10% criterion). The results are presented in terms of the Symmetric Extremal Dependence Index (SEDI, see Sect. 4.10) and are plotted in Fig. 5.8. A perfect prediction, i.e., 100% hits and 0% misses, would result in a SEDI of 1, marked by the black pentagon in the top right corner in the plots. As the relative occurrence of LLJs
Figure 5.6. Predictor importance for the selected set of single-level ERA5 variables in predicting the wind speed at different heights for ASIT, MMIJ, and Utö using the random forest model. Results are for the validation period. Figure adapted from Paper III.

Figure 5.7. For the three sites ASIT, MMIJ, and Utö the dotted lines mark the RMSE for wind speed profiles from ERA5 versus lidar observations. The RMSE reductions compared to ERA5 using the random forest (purple) or neural network (yellow) machine learning methods are also plotted. Figure adapted from Paper III.

in ERA5 is too low, as discussed in Paper II, ERA5 struggles with the task, having less than 10% hits across the sites. For the RF and the NN, the SEDI was in the range 0.4 to 0.7 for the three sites, with the RF consistently giving the highest SEDI.

The main advantage of the RF over the NN in the setup that was used, was the possibility to assign a cost matrix, penalizing missed LLJs. However, the drawback of optimizing for SEDI using a cost matrix is a high rate of false alarms; generally around 50% of the predictions were false alarms, with somewhat reduced numbers for Utö. Splitting the analysis into monthly subsets (not shown here) it was apparent that during the peak months in spring and summer, the ML models predicted LLJs almost all the time, resulting in a prediction of low value.
Classifying the wind speed profile that was generated by the ML models in the regression task as either an LLJ or a non-LLJ did not result in particularly good SEDI, typically in the range 0.1–0.2, but – in the case of the RF for Utö – 0.44. However, the information could be used as an extra training variable in the classification task, resulting in slight increase of the SEDI. In general, the surface sensible heat flux (SHF), i.e., the stability of the boundary-layer, was the variable with the highest predictor importance when predicting if there was an LLJ in the wind profile (results not included here).

**Paper IV**

Within the community of LLJ researchers, there is no clear guideline of which definition to use for LLJ identification and thus definitions vary between studies, making inter-comparisons difficult. A range of different, commonly used, LLJ definitions have already been presented in this thesis, see Sect. 4.1. The aim in **Paper IV** was to investigate differences between two qualitatively different LLJ definitions and suggest which definition to use for wind energy applications. The two definitions are described in detail in Sect. 4.1, and will here only be referred to as the falloff definition and the shear definition.

Using 44 years of hourly ERA5 data (1979–2022) up to 500 m above ground level at three onshore and three offshore sites, Fig. 1.3, the differences between the definitions were investigated. In Fig. 5.9, results are shown for one of the sites; the Baltic Sea.

In total, the falloff definition identified 130% more LLJs than the shear definition for the Baltic Sea. Results were similar also for the other sites, with the falloff definition identifying approximately 50–150% more LLJs. Among
Figure 5.9. Panels (a)–(h) shows, in order: (a) the average wind speed profile of low-level jets according to the two definitions, (b) the core height distribution, (c) the core speed distribution, (d) the normalized power production, (e) the seasonality (f), the diurnal cycle, and the symmetry in terms of (g) falloff and (h) shear above and below the LLJ core. Results are for the Baltic Sea, one of the six sites studied in Paper IV. For further details, see Paper IV, from which the figure is adapted.

In Fig. 5.9a, the average profiles of LLJs identified by the two definitions are plotted. In general, for all sites, the shear definition identified LLJs in stronger winds than the falloff definition. For the coastal/offshore sites, LLJs typically occurred in wind speeds close to the average conditions, but for the three onshore sites LLJs appear in somewhat weaker winds, especially above the rotor plane. The distributions of core height, Fig. 5.9b, shows that, for the shear definition, the distribution is flatter (offshore sites) or shifted higher up (onshore sites). The wind speed at the core, Fig. 5.9c – which to some extent is related to the core height, i.e., in general, the higher the core, the stronger the core speed – were generally higher, both in terms of the median core speed and the extremes, for the shear definition.

Calculating the REWS for the Gaertner et al. (2020) 15 MW reference turbine for the offshore sites and the Vestas 150-4.2 MW (Vestas 2023) turbine for the onshore sites, the distributions in expected power production during LLJs was calculated, showing higher production for LLJs found by the shear definition (Fig. 5.9d). Offshore, operational wind speeds occurred 90–92% of the time, and 97–99% of the LLJs (both definitions) appeared in those conditions. Onshore, LLJs identified by the shear definition were more likely to appear when the turbine was generating power, compared to those LLJs found by the falloff definition, that were more likely to appear at REWS below cut-in.
All three offshore sites displayed a clear seasonal pattern, Fig. 5.9e, with a peak in LLJ frequency in spring–summer for both definitions. The falloff definition indicated a weak diurnal cycle for all offshore sites, but this could not be seen using the shear definition. Onshore, however, there was a clear diurnal cycle for both definitions. There was a higher dispersion in falloff symmetry (i.e., falloff above the core divided by falloff below the core) for LLJs found by the shear definition, Fig. 5.9g, and there were also differences in shear symmetry, Fig. 5.9h, with LLJs identified by the shear definition being more symmetric.

Concluding from Paper IV, the shear definition is recommended for studies of LLJs in wind energy applications. As the height of the LLJ core also is a deciding factor for the properties of the wake behind a turbine (Gadde and Stevens 2021), an additional criterion dividing the LLJs into different subsets depending on the core height was suggested for analyses of wake properties. To allow for identifying core heights well above the top height of the rotor, profiles up to 500 m above ground level are recommended. However, using observational data, i.e., lidar measurements, there is a vertical limitation in the data set that is often lower than 500 m.

Paper V
The change in wind direction over the rotor of the 15 MW offshore reference turbine (Gaertner et al. 2020) at operational wind speeds was analyzed in Paper V, creating a climatology over Scandinavia using 43 years (1979–2021) of hourly ERA5 data. Lidar observations from Östergarnsholm and Utö were used for validation in the coastal zone. Time series of the change in wind direction over the rotor as measured by the Utö lidar are plotted in Fig. 5.10. For both Östergarnsholm and Utö, there was a clear seasonality with the major part of the occasions with pronounced directional shear found in the spring and summer. Directional shear exceeding 60° was rare but happened occasionally every year, in a few events for Utö even reaching beyond 90°.

In Fig. 5.11, the relative occurrence of strong, exceeding 15°, very strong, exceeding 30°, and extremely strong directional shear, exceeding 45°, in the

---

Figure 5.10. Time series of directional shear over the rotor of the International Energy Agency 15 MW reference turbine (Gaertner et al. 2020) at Utö, as measured by the lidar. Both the color coding and the size of the markers are proportional to the directional shear. Figure adapted from Paper V.
Scandinavian domain is plotted. Note that the 15 MW offshore reference turbine, with blades sweeping from 30–270 m above ground level, was also considered for onshore areas, to simplify the analysis. It is clear from Fig. 5.11 that all three categories of directional shear were more common onshore than offshore, related to the higher degree of surface friction resulting in a more pronounced Ekman spiral, see Sect. 2.1. The Bothnian Bay and the Bothnian Sea had a higher degree of strong directional shear compared to the Baltic Sea, which in turn experienced more directional shear than the North Sea and the Norwegian Sea. Very strong and extremely strong directional shear primarily occurred in the mountain range, along the northeastern coast of Sweden, and along the northwestern coast of Finland. The high degree of complex terrain and the slope from the highlands towards the shore create good preconditions for channelling effects and katabatic winds, Sect. 2.1, which in turn can be responsible for directional shear.

To analyze the main driving forces behind strong directional shear, ERA5 data were analyzed for the grid points closest to Östergarnsholm and Utö, respectively. A classification tree was fitted to the model data using BLH, wind speed and wind direction at hub height (ws150 and wdir150), the 2 m temperature (T2m), the LWT, the LCC, the amount of precipitation, and the profile type, defined as in Paper I, as predictors. From the classification tree, also the predictor importance was estimated, Sect. 4.7.

As seen in Fig. 5.12a, showing the normalized predictor importance for Utö, the BLH was by far the most important predictor for strong directional shear. In panel (b), the scatter plot clearly indicates that extremely strong directional shear was highly associated with very low BLH and also that strong

Figure 5.11. Climatology of how often the maximum directional shear over the rotor of the 15 MW offshore reference turbine (Gaertner et al. 2020) exceeds (a) 15°, (b) 30°, and (c) 45° for the Baltic Sea region. The climatology is based on ERA5 data for the period 1979–2021. Figure reproduced from Paper V.
Figure 5.12. In panel (a), the predictor importance for strong directional shear ($\geq 15^\circ$) is plotted, calculated using a binary classification tree. Scatter plots for the four most important predictors (BLH, ws150, wdir150, T2m) are presented in panels (b)–(e). The shade in the color in each panel marks the data density (number of hours that each dot represents), with darker shade for higher density. Results are for the grid point closest to Utö, all data from ERA5, 1979–2021. Figure adapted from Paper V.

and very strong directional shear mainly appeared in low BLH. The ws150 was the second-most important variable with weaker winds being more associated with strong directional shear than stronger winds. When it comes to wdir150 and T2m, patterns were less clear. However, the scatter plot with directional shear versus T2m indicates a shift in appearance of stronger directional shear in higher temperatures, resonating well with the seasonal pattern seen in the observations in Fig. 5.10, that strong directional shear typically happens in the warmer months.

Figure 5.13. Relative occurrence of strong directional shear, exceeding 15°, in four selected Lamb Weather Types: A, C, ANW, and CNE. Results are from hourly ERA5 data for the spring and summer season (March to August) for the period 1979–2021. Isobars are marked in white for every second hectopascal, with high (H) and low (L) pressure centers marked. The approximate synoptic wind direction is also marked (white arrows), but arrows are not scaled with the wind speed. Figure adapted from Paper V.
Figure 5.14. Boxplots of directional shear over the rotor in different wind speeds (left) and wind directions (right) at Utö. The width of each box is scaled with the amount of data in the corresponding bin and the ERA5 data are limited to match the time steps with available lidar observations. Bottom and top edges of the boxes show the 25th and 75th percentiles, respectively. The line within the box marks the median value. The 95% confidence interval of the median is given by the notches. Dots mark outliers and the whiskers reach the most extreme values not considered outliers. Figure adapted from Paper V.

The LWT, LCC, amount of precipitation, and profile type proved to be less important as predictors but can still provide interesting aspects for further analysis. For example, the frequency of strong directional shear for a few selected LWTs during the spring/summer half of the year (March–August) are plotted in Fig. 5.13. A clear difference in relative occurrence of strong directional shear over the Baltic Sea could be seen comparing anticyclonic (A) and cyclonic (C) conditions. Coastal effects were pronounced in some LWTs, for example in ANW when the air was advected over mainland Sweden before reaching the Baltic Sea, providing conditions for the onset of stable stratification and strong directional shear along the east coast of Sweden (see also Sect. 2.1). Similar patterns could also be seen in cyclonic conditions with a dominant wind from the northeast over the Baltic Sea (CNE) where strong directional shear primarily appeared along the west coast of Finland.

In Fig. 5.14, the distributions of directional shear over the rotor as measured by the Utö lidar and from corresponding time steps in ERA5 are plotted. The ERA5 underestimation of the directional shear distribution is evident in all wind speeds and in all wind directions. Results are similar for Östergarnsholm (not shown here). Analysing the relation between the type of wind speed profile and directional shear, Fig. 5.15, it is clear that, according to observations, non-ideal wind speed profiles are associated with a high degree of turning winds in the profile. For example, in the case of an LLJ, it is likely that also the directional shear is strong. There are some indications also in ERA5 that non-ideal wind speed profiles are associated with a high degree of directional shear, but because of the relatively low frequency of those profiles in ERA5 (as shown in Papers II and III), the profile type was not selected as an important predictor for strong directional shear (Fig. 5.12).
Figure 5.15. Boxplots of directional shear in different wind speed profile conditions for Östergarnsholm (left) and Utö (right). ERA5 data are limited to match the time steps with available lidar observations. For ERA5, the LLM boxes are missing due to the low occurrence of that profile type in the reanalysis. Bottom and top edges of the boxes mark the 25th and 75th percentiles, respectively, and the line within the box the median value. The 95% confidence interval of the median is given by the notches. Dots mark the outliers and the whiskers reach the most extreme values not considered outliers. Figure reproduced from Paper V. Color coding of the different wind speed profiles is the same as in Paper I.

Paper VI

In Sect. 4.3, the method to identify sea and land breeze circulations in NWP output from HARMONIE–AROME was presented. This method was put to the test in Paper VI, where both case studies and summarizing statistics are presented.

In the satellite image from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Terra satellite, taken at 1200 UTC (1300 local sidereal time) 25 July 2018, Fig. 5.16a, a sea breeze front is seen in the cloud pattern along both the east coast and the south coast of southern Sweden. Also on Gotland, convective clouds associated with the sea breeze were forming. The synoptic conditions were governed by a ridge of high pressure with weak winds and a heat wave caused temperatures to reach 30–35°C over land.

Three transects (A–B, B–C and D–E) were drawn, see Fig. 5.16a, and the transect-wise wind speed was plotted for each transect respectively. For the transect A–B on the south coast (Fig. 5.16b), a textbook sea breeze was occurring with winds directed from sea towards land on lower levels and in the opposite direction higher up. The white band in the cross-section, showing zero wind speed along the transect, was clearly located higher up over land than over sea, indicating a tilt of the circulation.
Figure 5.16. In panel (a), the satellite image shows southern Sweden on 25 July 2018 at 12 UTC (13 local sidereal time). Locations of the seven locations (Falkenberg, Helsingborg, Ronneby, Kalmar, Gotland West, Gotland East, Valdemarsvik) and three transects (A–B, B–C, D–E) analysed in Paper VI are marked. In panels (b)–(d), the cross-section of transect-wise wind speed is plotted for the three transects at the same point in time. Satellite imagery from the Worldview Snapshots application (https://wvs.earthdata.nasa.gov), part of the Earth Observing System Data and Information System (EOSDIS). Figure adapted from Paper VI. © American Meteorological Society. Used with permission.
The complex coastline with the narrow strait between mainland Sweden and the island of Öland affected the winds in transect B–C (Fig. 5.16c). Close to the surface, the winds were diverging at the east coast of mainland Sweden, but converging over Öland. Higher up, a return flow was seen, although not as clear as for transect A–B. Over Gotland, transect D–E (Fig. 5.16d), a double sea breeze circulation was occurring, with converging winds at low levels at the center of the island, and diverging winds aloft, both acting to enhance the convective conditions.

Using an archive of HARMONIE–AROME forecasts it was possible to generate a five year statistical overview of the occurrence of SBI and LBI for seven selected sites along the coastline of southern Sweden, locations marked in Fig. 5.16a. The SBI had, as expected (Sect. 2.1), a clear seasonal cycle (upper row in Fig. 5.17), while the occurrence of LBI was more constant throughout the year. For sites on the east coast (Kalmar, Gotland East, Valdemarsvik), the SBI was appearing more often than LBI, while for the west coast (Falkenberg, Helsingborg, Gotland West) and south coast (Ronnieby), differences were less pronounced and LBI could even be more frequent. Both the SBI and LBI showed clear diurnal patterns (see Fig. 5.17 bottom row) for all sites. As expected from theory, Sect. 2.1, peak occurrence of SBI was reached in the afternoon. For LBI, the frequency was higher in the evening, night and early morning compared to mid-day.

![Figure 5.17. Seasonal cycle of the sea and land breeze indices (SBI and LBI) at the seven selected sites (top row). Also the diurnal cycle is plotted (bottom row). Times are in UTC (local sidereal time minus 1 hour). Shaded areas mark the 95% confidence intervals of the frequency. Figure adapted from Paper VI. © American Meteorological Society. Used with permission.](image)

The method shows promising results of classifying the sea and land breeze from NWP output and with operational implementation, the SBI and LBI could serve as a tool for the weather forecaster, both in improved understanding of the weather situation and in communication with, e.g., media. In research, the indices could be used to classify the weather conditions and investigate systematic forecast errors in the coastal zone, as well as creating climatologies of the occurrence of sea and land breezes based on reanalysis data.
In Paper VII, different post-processing methods were applied to the wind speed forecast from HARMONIE–AROME, see Sect. 4.9. The performance of the post-processing methods in terms of improving predictions of expected wind power production was evaluated, comparing with the estimated power production converted from the wind speed measured by the Utö lidar (Sect. 4.8). The temporal smoothing, the RF and the NBH methods were the only methods that managed to get positive MAE skill scores overall, showing an improvement over the original forecast. Among these methods, temporal smoothing performed best, with the highest MAE skill score, the lowest bias, the highest correlation coefficient and the lowest RMSE.

The reason for the strong performance of the temporal smoothing method was attributed to a decrease in the risk of double penalty (Mass et al. 2002; Gilbert et al. 2020). For example, if – in reality – a cold front is passing at 1200 UTC, resulting in a short peak in the wind speed, but this peak is forecast at, let’s say, 1300 UTC in the NWP model, the MAE skill score will first punish the model for having missed the event, and then punish the model again for having too strong winds at the later time. By applying the temporal averaging filter smoothing the forecast using a window of $\pm 1\text{ h}$, the wind peak in the forecast is flattened and the wind speed at the prior hour increased. This results in both a lower penalty at the time step when the front actually was passing and also a lower penalty when the front was supposed to pass according to the forecast. It is however worth pointing out that, from a forecaster’s point of view, the smoothed forecast dampens information that could potentially be important, such as strong wind speeds and associated loads on the wind turbine.

To further analyse in which meteorological conditions the different post-processing methods performed well, the data were categorized hour-by-hour based on the HARMONIE–AROME forecast, assessing different wind speeds, synoptic weather conditions using a reduced set of LWT, stratification of the atmospheric boundary-layer and if an LLJ was forecast or not. Results are presented in Fig. 5.18. The temporally smoothed forecast performed best in all wind speeds, Fig. 5.18a, but both the RF and the D1/D2 mix performed well in weak to intermediate wind speeds, i.e., 2.5–10 m s$^{-1}$. Associated with the good performance in weak winds, the D1/D2 mix outperformed the other post-processing methods in LWT U (weak synoptic forcing, Fig. 5.18b). Generally, in all other synoptic situations, temporal smoothing was a safe choice, with remarkably higher performance in N and NE conditions, probably linked to cold air outbreaks and increased gustiness in these wind directions. This was also indicated by the high performance of the temporal smoothing method in unstable and strongly unstable conditions, Fig. 5.18c. However, also in stable and strongly stable conditions, temporal smoothing was best, although the RF almost gave the same MAE skill score in strongly stable conditions. The high performance of the RF in LLJ cases, Fig. 5.18d, is linked to the high
performance in strongly stable conditions, the stability class where LLJs are most frequent, see Sect. 2.1 and results in Paper I.

In total, based on these results, the main conclusion in Paper VII was to select which post-processing method to use based on the forecast weather conditions. In the specific case of Utö, the recommendation was to use temporal smoothing as a baseline improvement, but to switch to the D1/D2 mix in LWT U and in neutral stratification, and to the RF when an LLJ was forecast. This combination of post-processing methods would most likely lead to the best MAE skill score and improve the power production forecast, which in turn could be used for improved predictions of loads in the electrical grid and for enhanced trading on the energy market.
6. Discussion and outlook

The wind conditions over the Baltic Sea are complex. Due to the relatively short distance to the coastline from anywhere in the basin, local and mesoscale effects are frequently occurring. Although some answers about the Baltic Sea wind conditions have been given in this thesis, it also opens up for many new questions that will be important to research in the coming years. In this discussion and outlook section, a few of the most urgent issues are commented.

In Paper I, the decrease in variance in spectral energy of large turbulent eddies in case of LLJs compared to ideal profiles, Fig. 5.2, could possibly be attributed to the shear sheltering effect, Sect. 2.1. According to the theory of shear sheltering, the effect should only apply in stable stratification of the boundary-layer. However, as seen in Fig. 5.2, the suppression in the low frequency part of the spectra during LLJs occurred in all stability regimes, measured locally at 10 m. Turbulent characteristics can also be altered by e.g., the combined shape of the wind speed and temperature profiles (Holmboe instabilities), directional shear, shear production at the surface due to a positive momentum flux from swell, and dynamical changes in transport of turbulence due to horizontal gradients. For a perfectly fair analysis, weather situations that spatially are almost completely similar, only displaying different types of wind profiles at a site, would be needed to get a deeper understanding of the physical mechanisms. In addition to that, turbulent measurements in the profile, both above and below the core of the LLJ, and studies in different locations, both offshore and onshore in different terrain, would be beneficial for further investigation of the impact of LLJ on the turbulent spectra.

The formation of LLMs and negative profiles was explained by the positive, i.e., upward-directed, longitudinal $\overline{u'w'}$ momentum flux associated with swell. In total, these conditions occurred 3.5% of the time analyzed, but often the positive values were so small that the sign of the momentum flux could be considered to be within the uncertainty of the measurements. The Monin-Obukhov similarity theory – that was used to characterize the stability of the atmosphere, Sect. 4.4 – is not adapted for such conditions, and the validity to use the stability parameter $z/L$ to classify the stability during swell could be questioned. Implication of swell for offshore wind energy was assessed by Wu et al. (2020), discussing effects on the vertical profiles of turbulent kinetic energy (TKE) and mixing length, as well as changes in wind direction. As previously mentioned, it is also important to note that the stability was classified locally at 10 m a.s.l. in the meteorological mast, and it was not necessary that the same stability conditions were valid for the full profile. In a study by
Argyle and Watson (2014), it was shown that internal boundary-layers from a coastline could extend 80 km over the North Sea and that the classification of atmospheric stability over the sea was strongly dependent on the height of the measurements.

The uncertainties in reanalyses and wind atlases in accurately quantifying the offshore wind resource over the Baltic Sea, as discussed in Paper II, are important to keep in mind when planning for new wind power deployments in the near future. Uncertainties in the wind speed translates directly into an economic risk and uncertainties of the viability of the wind farm. Although not tested in the study, it should be possible to identify systematic biases in the models and correct for these using different types of statistical correction tools, such as e.g., ML models (compare with the method and results in Paper III). As mentioned in Sect. 3.1.1, lidar data were used as the ground truth in all studies, and uncertainties in the measured wind profile emanating from e.g., the large extent of the measuring volume (also commented in Sect. 3.1.1) were not considered. As new versions of reanalyses and wind atlases are periodically released, it is important to keep testing and evaluating their performance for wind power applications (see Gualtieri 2022 for a recent review). Higher spatial and temporal resolution is not always the key to better performance when it comes to some metrics of evaluation, such as the correlation coefficient and timing of events. This was also shown by the positive MAE skill score for the temporal smoothing and NBH, i.e., spatial smoothing, post-processing methods in Paper VII. However, the selection of validation measure depends on the application, and evaluating for other scores it is not necessary that the same methods would have been best.

What is also very important to consider for wind farm installations is not only the past and current wind conditions, but also the future conditions. Although there is a large variation in the model and the results are not significant, there is a weak indication of an increase of the 10 m wind speed over the Bothnian Sea and Bothnian Bay by the end of the century (see Fig. 6.1), considering the Coupled Model Intercomparison Project Phase 6 (CMIP6) results for the Shared Socio-economic Pathway SSP2-4.5 (indicating a 2.1—3.5°C global warming by the end of the century, IPCC 2023). The possible increase in wind speed over the extended Baltic Sea is in contrast to expected decreases over continental Europe and the Atlantic Ocean. An increase in surface winds is also expected in the Arctic offshore/ice sheet region, linked to the rapid decrease of ice cover (Jakobson et al. 2019; Vavrus and Alkama 2022).

Although climate models provide a general overview of expected average wind conditions in the future, it is also interesting to investigate the rate of change of mesoscale and local effects. Will non-ideal wind speed and wind direction profiles be more common in a future climate? Will there be a shift in the distribution of atmospheric stability, and thus also a shift in the distribution of turbulence and associated forces exerted on wind turbines? Will extreme wind events be more common? Also, as wind turbines continue to grow in size,
it is important that analyses of future climate not only assess wind conditions at 10 m height, but at typical hub heights of offshore turbines, and preferably at several heights across the vertical extent of the rotor.

In Paper III, it was shown that single-level data from ERA5 could be used as input to ML models to create accurate predictions of the wind speed profile. It is possible to build on these results to use actual real-time observations to predict the wind speed in the profile and, subsequently, also to estimate the power production. The set of training variables used as predictors for the ML models in Paper III were kept to a relatively small number to allow for a straightforward analysis of predictor importance, thanks to the low cross-correlation between the variables. However, to get the best prediction possible, a larger set of predictors should be tested. Since lidar campaigns are expensive to run for long time periods, a lidar could be installed in a wind farm together with a set of meteorological and oceanographic observations and – when the lidar is removed – the other observations could be used as input data to an ML model, trained on the period with overlapping measurements. As shown in Paper III, key variables to measure are the ws10 and the sensible heat flux.
(SHF). If meteorological data also are collected by the wind turbines, e.g., the incoming wind speed at hub height, this information is most likely even more valuable as input to the ML models, than e.g., the ws10. With longer time series and more training variables, it is also expected that the NN takes over as the best ML model of the two, as it generally is considered a stronger method for large problems (Wang et al. 2017). Depending on the application, a suitable metric or cost function should be selected in optimization of the ML methods. In Paper III, the RMSE was used to evaluate the wind speed profile and the SEDI for the shape of the profile. As commented in Sect. 5, the drawback of optimizing for SEDI is the large number of false alarms. To render high-quality predictions that can be trusted with high confidence, the number of false alarms has to be reduced. One way of doing that could be to optimize for a score that does not take correct rejections into account, as those are numerous. Examples of such scores are the critical success index (CSI) and the F1 score.

The difficulty of accurately predicting the timing of LLJ relates to the short persistence of LLJ events. The length of LLJ events was studied in Paper IV, but was not included in the summary in this thesis. An extensive analysis of the persistence of LLJ events and the time evolution of e.g., turbulence during the events would be of high interest. Also, the spatial development of LLJs over the Baltic Sea is interesting to further analysis, although some work on this topic already was performed by Svensson et al. (2019b).

The suggestion in Paper IV to streamline all future research within the field to use the same definition of the LLJ would simplify inter-comparisons between studies. Although the falloff and shear definitions of the LLJ give similar results in many aspects, the shear definition addresses the rate of change in wind speed with height and should be better suited for wind power applications (Gadde and Stevens 2021; Weide Luiz and Fiedler 2022). The shear definition is also less sensitive to the vertical window applied, which also simplifies comparisons of results from different studies.

The main drawback of both the falloff and shear definitions tested is that they build upon arbitrary thresholds, e.g., 1 m s\(^{-1}\), 10% and 0.01 s\(^{-1}\), without physical reasoning behind. If definite thresholds that relate to e.g., transitions in flow properties or turbulence behaviour could be identified, this would help to improve the definition. Statistical definitions could be used, e.g., classifying the 10% most extreme wind speed profiles with a local maximum as LLJs. This would generate more site specific definitions – of which there are both pros and cons. Furthermore, using a statistical definition it is possible that a wind speed profile that was previously classified as an LLJ is no longer considered to fulfill the requirements, as the data set grows.

Some studies have been performed on how the wake properties behind a wind turbine and within a wind farm are affected by non-ideal wind speed and wind direction profiles (Gadde and Stevens 2021), but more studies are suggested for further physical understanding and optimization of turbine yaw alignment, i.e., wake steering (see e.g., Fleming et al. 2017; Howland et al. 2018).
2019). Also, driving mechanisms of wind power ramps, i.e., rapid increases or decreases in wind speed, resulting in rapid changes in wind power production (Gallego-Castillo et al. 2015)) needs further attention, and studies of ramps associated with non-ideal wind profiles as well as sea and land breezes would be interesting.

The change of wind direction in the profile is an often neglected aspect of the wind conditions, although shown in Paper V that the directional shear can be of significant magnitude. As previously mentioned, both the effect on wake properties, structural and aerodynamic loading, and the risk of lower power production caused by directional shear are topics interesting for further investigation. Murphy et al. (2020) analyzed the change in REWS when including the effect of backing and veering winds in the calculation, but concluded that there was no statistically significant difference between the traditional REWS and the directional shear adjusted REWS at their onshore site in North America. Similar results are likely also for the Baltic Sea conditions, as the strongest directional shear typically occur in the lowest wind speeds, Figs. 5.12 and 5.14. However, taking a step back from the overall statistics, the directional shear adjusted REWS can still be important in trying to accurately predict the power production from a wind turbine in real-time.

The method to identify sea and land breezes that was proposed in Paper VI showed promising results capturing expected seasonal and diurnal cycles for coastal sites in southern Sweden (Fig. 5.17). However, for a thorough evaluation, the results should be compared both to results from other automated classification schemes and to a subjective classification performed by an experienced meteorologist. The method should also be tested in other regions, where other local and mesoscale wind patterns may affect the SBI and LBI.

What is important to note, and what was also discussed in Paper VI, is that – although the indices are called SBI and LBI, referring to the sea and land breeze respectively – they only identify if the circulation of the wind is in the direction of a sea breeze or a land breeze, not actually if the physical process causing this circulation is the sea breeze, the land breeze, or something else (e.g., a frontal passage, low-level katabatic winds towards a coastline, or convective cells with a high degree of directional shear). To address this and try to only identify sea and land breezes, additional criterion on e.g., the synoptic situation could be implemented, following the idea of Steele et al. (2015). Other possible extensions of the metrics are to include the integrated mass flow in the onshore and offshore directions (Holton 1990) or to include surface heat fluxes and vertical winds a few kilometers inland and over the water. The SBI and LBI are suggested to be used as a tool for the operational weather forecaster and can be used both to identify systematic bias in e.g., the temperature forecast during sea breezes, and to assist the forecaster in communicating the weather conditions in an accurate way to media and other stakeholders.

As computational power has increased rapidly in recent years, one of the fields that has been developing quickly is different techniques in applying
non-linear statistical methods to large data sets. Applying ML methods as post-processing for NWP output will remain popular and continue to improve the accuracy of the predictions. Building on the methodology in Paper VII, more advanced ML models could be implemented and compared with, e.g., temporal smoothing methods using longer averaging windows and NBH methods using a larger spatial distribution of grid points. The methods should also be tested onshore in different landscapes. Extending on the lagged D1/D2 mix method that was used, ensemble members of the NWP could also be used, either as input to create a deterministic prediction, as the forecasts studied in Paper VII, or to get a probabilistic prediction. Using ML methods it is also possible to generate probabilistic predictions based on deterministic input, to address uncertainties in the forecast, see e.g., Molinder et al. 2020.
7. Summary and conclusions

The aim of this thesis has been to characterize the peculiar wind conditions over the Baltic Sea and to investigate different post-processing methods to improve forecast for wind power production. In the introduction (Sect. 1.2), three research questions were stated, providing the stepping stones for the thesis. In this chapter, the answers to those questions will be summarised.

1. What are the characteristics and predictability of non-ideal wind profiles over the Baltic Sea?

There is no consensus in the community of exactly which definitions that should be applied for non-ideal wind speed and direction profiles. Throughout the studies in this thesis, a number of different definitions have been implemented. However, in Paper IV, the conclusion is that – for wind energy applications – a shear based definition is probably optimal for low-level jet (LLJ) identification.

Depending on the definition, and for different sites in the Baltic Sea, the LLJ frequency varies drastically, but in months of peak occurrence (May and June), LLJs typically appeared between approximately 10–15% of the time for the most strict definition/least common site, and 40–60% of the time for least strict definition/most common site, see Papers I and II. The vertical extent of the wind profile that can be assessed for LLJ identification is often limited by the vertical range of the instrumentation. When using models it is recommended to assess the wind profile well above the highest point of the rotor, to allow for identification of LLJ cores both below, within and above the rotor.

LLJs typically appear when warm air is advected from land over the relatively cold Baltic Sea in spring and summer. This results in stable conditions, a suppression of turbulent transport of momentum and a subsequent speed-up of the wind, creating the local maximum in the wind speed profile. Profiles with a low-level minimum (LLM) and negative profiles on the other hand typically appear in swell conditions. The upward momentum flux caused by the waves results in a boost in the wind speed at lower levels and therefore also a decrease in wind speed with height, up to a local minimum.

In the analysis from Östergarnsholm (Paper I) it is clear that there was a suppression in the low-frequency part of the turbulent power spectra when LLJs appeared compared to non-ideal wind speed profiles. This could possibly be related to the theory of shear sheltering, but further investigation is needed. For LLM and negative profiles, no clear patterns were seen in the low frequency part of the spectra.
The change in wind direction with height was analyzed in both Paper V, studying the change in wind direction across the rotor of a wind turbine, and in Paper VI, analyzing the sea and land breeze circulations. It was found that extreme cases of directional shear typically appeared in low boundary-layer heights and in weak wind speeds. Clear seasonal patterns were found for Östergarnsholm and Utö, with higher frequency of strong directional shear in the spring/summer half of the year (March–August). Along the coastline of southern Sweden, sea breezes were most common in summer, while for land breezes there was not much of a seasonal cycle. The sea breeze typically appeared in the afternoon while the land breeze was more common during night.

In Paper III, it was concluded that – by the help of machine learning (ML) models – it is possible to predict the wind speed in the vertical profile using ERA5 single-level data. Using actual surface observations it should be possible to predict the wind speed profile in real-time. Furthermore, ML models were used to predict if the wind speed profile had the shape of an LLJ or not. The ML models outperformed ERA5 in terms of the Symmetric Extremal Dependence Index (SEDI), but also resulted in a high number of false alarms.

2. How good are state-of-the-art models in describing the Baltic Sea wind conditions?

Three reanalyses and one wind atlas were validated against lidar observations from four sites in the Baltic Sea in Paper II. It was concluded that ERA5, UERRA, and NEWA all performed similarly and that the best choice depends on the application in mind. MERRA2 was not suggested for use in wind power analyses, as it lacks in both vertical and spatial resolution. In addition to that, MERRA2 displayed the – by far – greatest biases in the average wind speed profiles. In general, also ERA5, UERRA, and NEWA had a negative bias, underestimating the average wind speed in the profiles for most sites and most heights. This has a high impact on wind resource assessments as the available power scales as the wind speed cubed (Eq. 1.1), and uncertainties in the wind resource translates into a financial risk. Thus, applying bias correction to the model output is crucial to minimize systematic errors.

Due to the intermittent behaviour of LLJs, mostly appearing as very short events, models struggle in accurately predicting the timing of the events. However, in the overall statistics, UERRA proved to be better than ERA5 and NEWA in capturing the seasonality of LLJs and the distribution of core speeds (Paper II).
3. Can short-term power production forecasts for an offshore/coastal site in the Baltic Sea be improved using post-processing methods?

The rapid increase in the capability of ML methods can be used to enhance output for wind power predictions from numerical weather prediction models. However, also simplistic models such as temporal and spatial averaging can improve the forecasts. This was studied in Paper VII. In terms of the mean absolute error (MAE) skill score, the temporal smoothing post-processing method resulted in a higher score than the much more advanced random forest (RF) method. In the study, focusing on the wind conditions over Utö, it was concluded that the temporal smoothing could be used as the standard post-processing method to be applied, but changed to the D1/D2 mix in weak synoptic forcing (i.e., Lamb weather type U) and to RF if an LLJ was forecast. However, depending on the application, it could be more suitable to use another metric of evaluation, resulting in a different recommendation.

Although this thesis has provided answers to the above three research questions, it has also opened up for many new interesting aspects of the wind speed and wind direction profiles to investigate further. Improved reanalysis and NWP models as well as advanced non-linear statistical ML models are important in taking the next steps forward. However, maybe the most important factor to reach further understanding of the physical processes associated with non-ideal wind profiles in the coastal zone is extensive observational data of high quality, spanning multiple years and reaching higher up, as wind turbines continue to grow in size.
Du kanske inte tänker på det till vardags, men luften runt omkring dig är full av energi. Särskilt under riktigt blåsiga dagar, då virvlar det av energi överallt. Det är just det här som är grundprincipen för vindkraft; att dra nytta av energin i vinden och omvandla den till el. De senaste decennierna har elanvändningen i samhället ökat explosionsartat, och därför har också behovet av miljövänlig och långsiktigt hållbar elproduktion ökat.

Vindkraft erbjuder ett sådant alternativ och utbyggnaden av vindkraft har gått snabbt framåt, framförallt på land, även om havsbaserad vindkraft tagit allt större marknadsandelar de senaste åren. Fördelarna med havsbaserad vindkraft, om man jämför med landbaserad, är att vindförutsättningarna generellt sett är bättre ute till havs: det blåser mer och det blåser oftare. Massor med energi helt enkelt. Dessutom är det rent tekniskt möjligt att bygga större vindturbiner som står i havet, vilket gör att man kan utnyttja vinden ännu bättre och extrahera mer energi. Moderna havsbaserade vindkraftverk har en navhöjd på cirka 150 m och vingar som är 120 m långa, vilket innebär att de sveper genom luften från 30 m höjd upp till 270 m över vattenytan. Det är inte riktigt lika högt som Eiffeltornet, men nästan!

Vindarna kan variera en hel del över det höjdintervallet, och det är just det som den här avhandlingen handlar om. Hur ändras vindarna med höjden ute till havs och varför dyker märkliga vindprofiler upp? Hur bra är datormodellerna på att beskriva vindarna och går det att hitta några metoder för att förbättra väderprognoserna så att förutsägelserna om hur mycket el som kommer att produceras blir mer träffsäkra?


Den här avhandlingen utgörs av sju artiklar som övergripande beskrivs på följande sidor.


Artikel II är en validering av fyra modeller för att utvärdera hur bra de är på att beskriva vindförhållandena över Östersjön. Även om alla modellerna är toppmoderna skiller de sig ganska mycket åt i hur de beskriver vindprofilerna. Värdena från modellerna jämfördes med lasermätningar av vindprofiler på fyra platser i Östersjön: på Östergarnsholm som nämndes tidigare, på Utö långt ut i den finska skärgården, på den danska ön Anholt och på mätplattformen FINO2 i den tyska delen av södra Östersjön.

Resultaten visar att tre av modellerna (ERA5, UERRA och NEWA) lyckades återskapa den allmänna vindprofilen på ett bra sätt. Den fjärde modellen (MERRA2) underskattade däremot vindarnas kraftighet. När det kommer till jetströmmarna på låg nivå så var UERRA bäst på att beskriva på vilken höjd jetströmmarna låg och hur kraftiga de var. Den generella rekommendationen från studien är att ERA5, UERRA och NEWA alla kan användas för att karakterisera vindarna över Östersjön, men exakt vilken modell man ska välja beror på vilken tillämpning man tänker sig.


För att bygga vidare på studien kan man använda mer avancerade statistiska metoder och faktiska observationer istället för modelldata från ERA5. Om man optimerar modellerna på rätt sätt är det möjligt att de kan användas för att förutsäga vindarna i realtid, med större träffsäkerhet än en väderprognosmodell, vilket kan vara till stor nytta för att optimera vindkraftsproduktionen i en park och för de som handlar på elmarknaden.

Artikel IV fokuserar helt på de speciella jetströmmarna, närmare bestämt på hur vi bäst kan defniera dessa vindar. Problemet är nämligen att det finns många olika definitioner inom forskningsfältet, vilket gör det svårt att jämföra resultat från olika studier. Två definitioner testades: den ena fokuserar på hur mycket vindstyrkan minskar både över och under jetströmmen, den andra på hur snabbt vindstyrkan minskar över och under jetströmmen. Mer än fyrtio år av data från ERA5 användes för tre platser ute till havs och tre platser över land. Efter att ha jämfört och diskuterat många olika aspekter av definitionerna blir slutsatsen att den definition som tar förändringen med höjden i beaktande (det vill säga hur snabbt vindstyrkan minskar) är den fördefred om man vill studera hur jetströmmarna påverkar vindkraften.

Artikel V vridde blicken från vindens styrka och fokuserar istället på hur riktningen ändras med höjden. Även om vindriktningen oftast inte ändras särskilt mycket med höjden kan det ibland vara omöjligt att linjera upp turbinen så att vinden kommer in vinkelrätt överallt. Inte nog med att det påverkar elproduktionen, hela strukturen utsätts då också för stora belastningar.

Utifrån fyrtio år av data från ERA5 skapades en skandinavisk klimatologi av förekomsten av stora skillnader i vindriktning med höjden och orsakerna identifierades. Dessutom jämfördes resulterten från ERA5 med de tillgängliga lasermätningarna från Östergarnsholm och Utö och det var tydligt att ERA5 underskattade vindvridningen rejält på båda platserna. Observationerna avslöjar också att det finns ett samband mellan kraftiga förändringar i vindriktningen.
och svaga vindar och att de ofta förekommer i samband med jetströmmar på låg nivå.


I artikeln utfördes ett par fallstudier för att visa indexets styrkor och svagheter och utifrån fem års arkiverade väderprognoser skapas en översikt över sjö- och landbrisens förekomst på utvalda platser längs Götalandskusten. Potentiellt har indexet stort värde om det implementeras i väderprognosmodellerna. Meteorologen får då assistans i sin bedömning av väderläget och med en förbättrad förståelse kan även prognosen förbättras. Dessutom kan indexet användas till att identifiera systematiska fel i vädermodellerna och för att klassificera väderläget i framtida forskningsstudier.

Acknowledgements

When I was a kid, I used to sail the Stockholm archipelago with my parents. Every evening it was important to be quiet for a few minutes; dad was listening to the weather forecast on the radio. It’s fair to assume that this was where my fascination for the weather in the coastal zone started. Although I’ve always enjoyed watching the clouds transform, little did I know that one day I would become a meteorologist and get paid to do so.

After more than five years of working as a weather forecaster I decided to try the academic path, and when the right project appeared it was not much of a hesitation to apply. Thank you for believing in me from the start Erik! Thanks for keeping this trust, giving me the space to develop my skills and planning my work exactly how I wanted – and always knowing that your guidance was there if needed. Many thanks to my co-supervisors Heiner and Stefan! The three of you really live up to the term super in the word supervisor. With your different expertise in the field, I couldn’t have gotten a more complete team around me. Thanks for the academic training! Also, thanks for the trips to Erken Erik, thanks for the music Heiner, and thanks for Visby, Copenhagen, and Glasgow Stefan!

I’ve sent a lot of emails throughout my PhD. Most of them I’ve forgotten, but there is a special one that I remember. May 26th, 2021, I listened to the presentations in the online Wind Energy Science Conference. All of a sudden, another PhD candidate showed up, presenting her research about low-level jets. Of course I had to reach out! This was the start of a long and fruitful collaboration with you Jeanie. It has been so much fun having someone with the exact same research questions top of mind to discuss with and together we reached further than we would have done individually. Thanks for all the long zoom calls. Thanks for being my LLJ friend and colleague!

Writing papers is a teamwork, and I’m truly thankful for all the support I’ve gotten from professors and researchers both in Uppsala and abroad. Thank you Johan! I would have loved to have you as a co-author on every single paper. Your comments really improved my scientific writing. Thanks for being so easy-going and thanks for all the chats about running, cycling, and all types of sports. And for the amazing pastries. Thank you Erik for your expertise in the Östergarnsholm data and for helping out with the data handling. Thanks to Rebecca and Sara for your input on the papers, I was very delighted to meet you in Scotland! Thank you Meto for your love for mesoscale phenomena! I cannot express my gratitude to you Ville; without the Utö lidar data I could not have performed this work. Thanks for being so quick in replying and kindly
providing this excellent data set. Thank you Heidi for sharing the wave buoy data and thanks Semjon for your help and support with the UERRA data. By the way, thanks to Reviewer II of Paper VII! I printed your kind review and taped it to the wall in my office to look at every time I needed a boost.

I also want to acknowledge that this work was performed thanks to Energimyndigheten, VindEl program (grant no. 47054-1). The work also formed part of the Swedish strategic research program STandUP for Wind. The ICOS station Östergarnsholm is funded by the Swedish Research Council (Vetenskapsrådet) together with Uppsala University and with the help of the research infrastructure facilities provided by FINMARI (Finnish Marine Research Infrastructure network), the wave measurements used in Paper I were maintained.

To pursue a PhD is to travel, in space and in mind. Even though I spent most of the time in my beloved hometown Uppsala I made friends from all around the globe: Austria, Belarus, England, France, Germany, Italy, Poland, Spain, India, Japan, Australia, New Zealand, Panama, Mexico, and the US.

Lucia and Kevin!! Thank you for all the laughs, the coffee breaks with freshly baked croissants, the therapeutic runs and rides, the long dinners and the late night parties. Thanks for the trips to Östergarnsholm! Thanks for the hikes, the kayaking, the skiing, and the ice skating. Thank you for being such awesome friends! You made my PhD life a thousand times more fun than it would have been without you. Thomas and Ramona, thank you for trading your beans and for going to space with me! Thanks for the winter sports in Åre and the summer evenings around the barbecue. Elena, thank you for the fondue, the pizza, the tiramisu, for baking lussekatter and having glögg with me. Always accept all the cookies! Thanks for joining me singing when I played the piano at Geo, we went on for hours through the night!

I have been extremely lucky to be gifted with the best office mates I could have ever asked for. Chiara, thanks for being there, back to back with me in the office. Thanks for dimming the lights and taking a few moments to turn the office into a moon rise cinema those dark January afternoons. Thanks for the bandy, the tennis matches, and the hikes! One day I’ll come visit you in Austria! Gabriella, thank you for challenging me and making a mess on my desk, I guess I needed that! Thanks for “How to hide your hoodie!” and thanks for letting me help with your coding and plotting, you know how much I enjoy that! Emma! Thank you for being my go-to stats expert and for starting every day chatting to me in Swedish. It’s been great seeing you conquering a new language! Thank you for your care and for cheering on me!

There is nothing a meteorologist enjoys more than a nice atmosphere, and I’ve truly enjoyed the relaxed feeling in our department. To all my friends and colleagues at Geo, thank you! Thank you Ferran for rentaplats and for inviting me to join the Easter trip to Nynäshamn and the hike in Tyresta. Anastasiya, thanks for seeing my fascination with the sky and for bringing me a cloud from England. And the cloud socks of course, my favorites! Thank you Rich for
valborg and the weather discussions. Thank you Erik for jamming with me in your PhD party, it was so much fun! Faranak, you are such a superstar! Thanks for your energy! Thank you Hector for your hospitality and for the hikes up and down the Norwegian mountains! Thanks for your excitement whenever you talk about machines or the Mexican way of living! Thanks for droning with me! Leonie, thanks for the climbing course and letting me drive your ADCP around Sweden! Sara, thank you for your advice and kindness! Finishing our PhDs almost at the same time, it feels like we established a special bond. Thank you Alizée for the trip to Abisko, the blueberry picking, and the post-volleyball beer. Thanks for the chocolate and the popcorn, the runs, and all the hittaut hikes! Thanks for the chats about sailing Zibi and for trying fencing! Yunus, thanks for the vibe you are spreading in the corridor! Thanks for being an amazing coordinator Andra and a great teammate in the drawing-guessing game – let’s say we are both better at guessing than drawing! Thanks for your quirky sense of humour Nicola! Gianmarco, thanks for the fun times and the wood chopping! Thank you Leonardo and Cristiano for the basketball talks, how come we never played?! Thanks for the interesting scientific discussions Mirea and for joining the squad to concerts and bike races! Laura, thanks for checking in on me, providing chocolate, and small pep talks! Thanks for the company Makoto, Michael, Franziska, Georgios, Elisa, George, Riccardo, Taylor, Emma, Joshi, Sara, Merle, Venu – during coffee breaks, lunches, and different festivities. Thanks to the original Geo team – Jennie, Nina, Kristina, Audrey, Eduardo – you were here when I started and gave me a warm welcome! Thanks for being my teaching companion Linus! It was so much fun discussing teaching strategies with you and developing the online course in basic meteorology and climatology!

Anna, thanks for sending me up the tower on Östergarnsholm and for asking the important scientific questions, it helped me. Thank you Cicci for all the talks about the weather and how the fronts were drawn in today’s charts in the newspaper. Malin, thanks for involving me in fun outreach activities. I really enjoyed writing the popular science texts for the Celsius weather photo calendar! Giuliano, thanks for the fun conversations and sharing a taxi with me to Stockholm when the trains got stuck. Thank you Fatima for your support. Elisa, thanks for the yoga! Thanks for the walk on the cliffs in Herrvik John! Thank you Lichuan, Jacopo, Melinda, Shuping, Johanna, Monica, Marcus, Gabriele, Benjamin, Veijo, Antonio, Tom, Fritjof, Björn, Gunnar, Hans, and Sofia for the chats over Thursday fika. Thanks to the Gotland group – Henrik, Henry, Prabhu, Memo, Hugo, Huidong – you guys are fun! And thanks to the Reference group for your positive feedback on my work!

Thanks to all the students I’ve been teaching, it’s been a lot of fun! Especially thanks to the bachelor and master students in meteorology. I had the honor to supervise a few of you in your projects; thanks for giving me new perspectives on wind energy topics August, Tumi, Anabelle, and Nelly!
Tack till mina kursare Leon, Elisabeth, Jalle, Kristian, Malin och Nitzan! Ni har för alltid en särskild plats i mitt hjärta! Varje år när jag undervisar mätprojekten och kartanalysen tänker jag på er och på hur kul vi hade! Tack Forecagänget, Erik, Maria, Lina, och Ingrid – här kan man snacka om att stå stadigt när det blåser hårt! Vi byggde ett speciellt band under de där åren. Tack Tora för en bal på slottet och för att du svängt förbi för en kopp kvällste i väntan på Stockholmståget. Tack till alla meteorologvänner därute: Annie, Charlotta, Emma, Lasse, Maria, Micke, Páll med flera!


Till Olle och Igor som gick bort under min doktorandtid, ni var båda väldigt speciella personer. Jag lärde mig mycket av er, på olika sätt. Tack.

Characterization and forecasting of wind conditions over the Baltic Sea
Oil on canvas, Silja Wager 2023
Bibliography


Fernández-Granja, J. A. et al. (2023). Exploring the limits of the Jenkinson–Collison weather types classification scheme: a global assessment based on


Li, X. et al. (2007). Coastal katabatic winds imaged by SAR. *Geophysical research letters* 34.3. DOI: 10.1029/2006GL028055.


Mass, C. F. et al. (2002). Does increasing horizontal resolution produce more skillful forecasts?: The Results of Two Years of real-Time Numerical Weather Prediction over the Pacific Northwest. *Bulletin of the American Meteorologi-


A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)