UiT The Arctic University of Norway

Faculty of Science and Technology
Department of Physics and Technology

Offshore Wind Prediction with Graph Neural Networks

Patricia Asemann

FYS-3900: Master's thesis in Physics, June 2024

In loving memory of Klaus-Peter "KP" Haupt, a science teacher and educator, founder of the student research centre (SFN) in Kassel, Germany, and, most importantly, a mentor and role model to many.

His unwavering support and believing in me paved the way to a life I could not have imagined in my wildest dreams. I am forever grateful for the courage and boundless curiosity he instilled in me.

Not long ago, the Earth completed its first full orbit around the sun since he quietly bid farewell to the living. He passed away on May 8, 2023, long before any of us were able to give back to him what he gave to us.

Danke für alles, KP.

“Data is not information, information is not knowledge,
knowledge is not understanding, understanding is not wisdom.”
–Clifford Stoll

“But the wind and water know all the earth’s secrets.”
–Amy Harmon

Abstract

Offshore wind energy, especially in regions like the Norwegian Arctic, is a promising source of renewable energy. Graph Neural Networks (GNNs) have shown potential in modeling complex systems like weather, making them suitable for improving wind resource assessments. This thesis investigates the use of GNNs for predicting offshore wind patterns, utilizing high-resolution Synthetic Aperture Radar (SAR) data from Sentinel-1 and the Copernicus Arctic Regional Reanalysis (CARRA) data. The research begins with a thorough exploration of wind data sources to evaluate their reliability. The findings indicate that the SAR-based wind retrieval method offers superior spatial resolution and detail compared to traditional reanalysis products and in situ observations while maintaining an accurate representation of long-term wind resources despite its poor temporal resolution. Experiments with several graph and GNN architectures were conducted to assess their effectiveness in predicting wind fields. Simple GNN architectures generated reasonable two-dimensional wind fields but struggled to capture the detailed variations observed in SAR data. This suggests the need for more sophisticated architectures and additional data inputs to improve accuracy. Key findings highlight the importance of incorporating long-range spatial dependencies, refining performance evaluation methods, and expanding the training dataset with more comprehensive data sources. This thesis represents a first step toward integrating GNNs into offshore wind resource assessments and identifies areas for further exploration.

Acknowledgements

I would like to express my deepest gratitude to all those who have supported and guided me throughout my degree and in completing this Master's thesis.

A huge thank you to Eduard Khachatrian, for the excellent teamwork throughout this project. Thank you for the countless motivational speeches, encouragement, and all the chocolate. To Benjamin Ricaud, my supervisor, as well as Yngve Birkelund and Igor Esau, for many intriguing discussions and helpful advice. Your knowledge and experience have greatly contributed to my academic growth and understanding of scientific research. And I extend my heartfelt thanks to our student advisor, Kirsti Merete Johannessen, for supporting us throughout this degree so competently and attentively.

To my classmates, Hannah and Ronnie: this degree has been nothing short of a breeze with the two of you by my side. Thanks for the solidarity, the friendship, and for making even the most desperate study sessions enjoyable. To the most wonderful friends – Hajal, Tzak, Fenja, Julia, Matthias, and Lauren: thank you for always believing in me and for putting things in perspective when needed, and to Anselm, not only for your friendship but also for pointing out all my silly little mistakes. To Peter, for keeping me fed and caffeinated, for distracting me with philosophy papers or the climbing gym, and for everything else. To my family – Mama, Papa, and my sisters, Terri and Becci: your love and support have always been my foundation. I owe you everything and I love you.

This thesis would not have been possible without the collective support of all of you and many more. I am immensely thankful for each one of you.

Finally, my sincere gratitude to the German Academic Exchange Service (DAAD) that supported this work and this Master's degree with a fellowship.

Deutscher Akademischer Austauschdienst
German Academic Exchange Service

Preface

This thesis project was part of a collaboration between the Machine Learning group and the Renewable Energy group at UiT that commenced in the summer of 2023, supported by Equinor Akademiaavtalen. The objective of this initiative is the exploration of possible applications of Graph Neural Networks (GNNs) for offshore wind prediction, focusing on the Norwegian Arctic. The project group within the time period of this thesis project consisted mainly of Prof. Yngve Birkelund, my supervisor Prof. Benjamin Ricaud, Prof. Igor Esau, Eduard Khachatrian, PhD (postdoc), Lihong Zhou (PhD candidate), and myself.

The first project phase involved the evaluation of available regional wind data sources in order to understand how GNNs would have to be applied, built, and trained. This resulted in a peer-reviewed publication created mainly by Eduard Khachatrian and myself. Therefore, this thesis deviates from the traditional monograph format commonly employed for Master's theses and instead adopts a hybrid structure that incorporates both monograph and the following publication:

> Eduard Khachatrian, Patricia Asemann, Lihong Zhou, Yngve Birkelund, Igor Esau, and Benjamin Ricaud. Exploring the Potential of Sentinel-1 Ocean Wind Field Product for Near-Surface Offshore Wind Assessment in the Norwegian Arctic. *Atmosphere*, 15, no. 2: 146, 10.3390/atmos15020146, 2024.

To preserve the coherence of the thesis and maintain a logical progression of topics, this publication has been integrated into this thesis as Chapter 2. This placement ensures that the reader can follow the unfolding narrative without disruption, as from Chapter 3 onwards, the thesis transitions to the application of GNNs. The inclusion of the published work serves to provide a tangible example of my first attempts at a "proper" scientific contribution. One other work was produced during the period of this thesis project that is not part of this thesis:

> Eduard Khachatrian and Patricia Asemann. The Value of Sentinel-1 Ocean Wind Fields Component for the Study of Polar Lows, Under Review at *Frontiers in Earth Science*.

All code produced by me within this thesis project will be made available to Benjamin Ricaud for further development within the project group.

Contents

List of Abbreviations

CARRA	Copernicus Arctic Regional Reanalysis
CNN	Convolutional Neural Network
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF Reanalysis v5
ESA	European Space Agency
GAN	Generative Adversarial Network
GCN	Graph Convolutional Network
GNN	Graph Neural Network
kNN	k-Nearest Neighbour
MSE	Mean Squared Error
NORA3	3km Norwegian Reanalysis
NWP	Numerical Weather Prediction
OCN	(Sentinel-1) Ocean (products)
OWI	(Sentinel-1) Ocean Wind Fields (component)
PyG	PyTorch Geometric
ReLU	Rectified Linear Unit
RMSE	Root Mean Squared Error
SAR	Synthetic Aperture Radar
WRF	Weather Research and Forecasting (model)

1 Introduction

1.1 Background and Motivation

As the global demand for sustainable power solutions continues to grow, offshore wind energy has emerged as one of the most promising sources of renewable energy [14]. With its large relative offshore area, high average wind speeds and high capacity factors, Norway makes an ideal candidate for the expansion of offshore wind power production [3]. The Norwegian energy industry is a world leader in offshore operations and installations, and much of the expertise from the oil and gas industry can leverage the development of offshore wind power. When it comes to floating offshore wind technologies, Equinor alone already holds a 47% share of the global power production. In 2022, the Norwegian government launched an ambitious initiative, planning on allocating large areas in Norwegian waters to increase offshore wind production in Norway to 30 gigawatts by 2040 [39]. To achieve this goal, around 1,500 wind turbines will have to be installed over the course of less than two decades. The increase in produced power is supposed to provide affordable electricity to Norwegian households, be exported to other countries, and to help meet the electricity demands of the petroleum sector. Although several offshore wind parks and research initiatives are already operating in the South and West of Norway, the potential of Northern Norway remains untapped.

A critical first step toward harnessing this potential involves the identification of suitable locations for wind power plants. Evaluating wind power potential in offshore regions presents several challenges. One of the main limitations

is the sparse availability of reliable observational or numerical data in these remote areas. Observational data includes in situ measurements and remote sensing data, and is essential for accurate wind resource assessment. In situ measurements are extremely sparse in our area of interest due to the logistical difficulties and high costs associated with deploying and maintaining equipment in offshore and Arctic regions. Remote sensing technologies offer alternative methods for obtaining observations of (larger-scale) wind fields. Until recently, scatterometers were the primary tool used for this purpose, but lately, Synthetic Aperture Radar (SAR) has been on the rise as a valuable method for retrieving offshore wind data. SAR determines the near-surface wind speed by measuring the backscatter from the sea surface. Capillary waves and sea surface roughness are directly influenced by atmospheric forcing, i.e. wind above the surface. Thus, the backscattering can be interpreted as a function of wind speed, wind direction relative to the look angle, and incidence angle. In comparison to scatterometers, SAR-based offshore wind retrieval methods result in slightly less accurate wind speed readings and are limited in their ability to retrieve the wind direction [7, 45]. However, they provide wind fields with a much higher spatial resolution—improving it by an order of magnitude—which results in the ability to capture the complex wind patterns often present in coastal areas, and a better representation of small-scale weather events [7].

Traditionally, wind resource assessment has relied on numerical simulations, including reanalyses and numerical weather prediction (NWP) models. High spatial resolution NWP models like the Weather Research and Forecasting (WRF) model can provide valuable insights into wind patterns over various terrains, but they are computationally very expensive, making them less accessible for large-scale or long-term wind resource assessments [48]. Additionally, the uncertainties and errors present in NWP models due to our incomplete physical understanding of weather systems and the insufficient resolution of small-scale interactions still lead to inaccuracies [13].
Reanalyses are comprehensive datasets that integrate historical weather observations with modern forecasting models to provide a continuous, gridded dataset of atmospheric conditions over time. Global reanalyses like ERA5 from the European Centre for Medium-Range Weather Forecasts (ECMWF) offer extensive temporal coverage and are widely used for various climate and weather applications. However, their performance in offshore areas, particularly in the Norwegian Arctic, is most limited by their coarse spatial resolution, failing to resolve mesoscale features or the complexities of wind fields introduced by the interaction of ocean and land areas [12, 16].
Regional reanalyses tailored to our area of interest include the 3km Norwegian Reanalysis (NORA3) and the Copernicus Arctic Regional Reanalysis (CARRA). They have higher spatial resolutions than global datasets and are better adapted to the specific meteorological characteristics of the Norwegian Arctic. However,

even they have the tendency to underestimate the mean wind speed and, thus, wind resources [19, 28, 43].

To tackle these challenges, there is a pressing need for better data coverage that is both temporally extensive and spatially detailed. In this context, innovations in machine learning and deep learning technologies offer promising solutions[1]. Among these, graph neural networks have emerged as a powerful tool capable of modeling complex systems governed by partial differential equations, such as fluids, and thereby weather systems [35, 41]. By leveraging the strengths of GNNs, it should be possible to create models that can better understand and predict wind patterns and energy availability in regions with limited observational data and poor-performing reanalysis products. This approach has the potential to mitigate some of the limitations of traditional wind resource assessment methods, possibly providing more reliable and accurate wind resource assessments.

1.2 Objectives and Research Questions

The research project which this thesis is embedded in is at an early stage. This thesis aims to take the first steps toward wind power prediction with GNNs. The following two main objectives and underlying research questions were raised:

Data exploration Assessing the reliability of different wind data sources available for the Norwegian Arctic in order to determine suitable ground truth data and training procedures for supervised deep learning methods.

- To what degree can SAR-based wind retrieval methods compete with the more traditionally used reanalysis products and in situ observations in offshore wind resource assessment?
- Is the information contained in the currently available SAR wind data for our area of interest comprehensive enough to serve as a reliable training source for deep learning networks, particularly GNNs, in offshore wind prediction?
- How can wind data be translated into a graph structure that allows message passing in accordance with the physical properties of wind fields as dynamic systems?

1. Section 3.3 provides more details on previous developments of deep learning in meteorological research.

Experimenting with GNNs Providing a preliminary understanding of the potential benefits of using GNNs for offshore wind prediction.

- Can the limitations of traditional methods for offshore wind resource assessment be addressed and mitigated by GNNs?

- Is a simple GNN architecture fed with minimal meteorological data able to generate useful two-dimensional near-surface offshore wind fields consistent with the level of detail present in SAR observations?

- How do changes in network parameters or architecture choices (e.g., number and order of layers, graph setup) affect the predictive power and accuracy of GNNs?

Chapter 2 will dive into the first part, data exploration, and give an answer to the first research question; comparing five different wind data sources (in situ and remote sensing observations as well as three different reanalysis products). Chapter 3 and Chapter 4 will lay the theoretical groundwork for exploring GNNs and provide the technical information of our experiments, respectively. The results of those experiments are presented in Chapter 5. Lastly, Chapter 6 recapitulates the main results of both parts and gives an outlook to future research on this topic.

2

Offshore Wind Data Exploration: Publication

Objectives The primary objective of this paper is to evaluate strengths and weaknesses of the most prominent publicly available wind data sources covering the offshore region of the Norwegian Arctic. We aim to particularly compare the lesser-used SAR-based Sentinel-1 Level-2 Ocean Wind Fields (OWI) component with the more common reanalysis products and in situ measurements to determine its suitability for offshore wind resource assessment.

Summary This paper conducts a case study, analysing 238 SAR scenes from 2022 for the Goliat station, an oil platform located in the Barents Sea. The Sentinel-1 data is compared with in situ observations and three reanalysis products—ERA5, NORA3, and CARRA—to evaluate its accuracy in representing offshore wind conditions. Statistical measures such as RMSE, correlation coefficient, and standard deviation are used to assess the performance of Sentinel-1 against the other data sources. Additionally, the study analyses the Weibull parameters to characterize wind speed distributions. It underscores the importance of high spatial resolution data in capturing intricate wind patterns near complex coastlines, which is critical for the planning and development of offshore wind farms. The findings contribute to the broader understanding of integrating SAR-based wind products with traditional wind assessment methods, providing a robust understanding of the publicly available wind data sources in the Norwegian Arctic.

Main Findings

1. The Sentinel-1 OWI product demonstrates similar statistical properties to in situ and reanalysis data, with comparable RMSE, correlation coefficients, and standard deviations.

2. The Weibull parameters for Sentinel-1 closely match those of the reanalysis products, indicating its capability to represent wind speed distributions accurately. However, in situ measurements tend to underestimate wind speeds compared to other sources.

3. While Sentinel-1 offers unmatched spatial resolution, its temporal resolution limitations do not significantly hinder its ability to characterize wind speed distributions, making it a valuable tool for offshore wind resource assessments.

Contributions and Authorship Eduard Khachatrian (E. K.) and I (P. A.) contributed equally to the publication. Eduard had already acquired the datasets and performed parts of the formal analysis when I joined the project. I reviewed existing literature on the topic, contributing to the framework of the manuscript. Eduard and I analysed and interpreted the results together. Yngve Birkelund, Igor Esau, Benjamin Ricaud, and Lihong Zhou contributed valuable comments to the manuscript. We also thank the anonymous reviewers for their constructive feedback.

Section	*Author*
Abstract	P. A.
1. Introduction	P. A.
2. Theoretical Background	P. A.
3. Methods	
3.1 Datasets	E. K.
3.2 Temporal and Spatial Preprocessing	E. K.
3.3 Statistical Analysis	E. K.
3.4 Wind Speed Distribution	P. A.
4. Results and Discussion	
4.1 Wind Speed	both
4.2 Wind Direction	both
4.3 Wind Speed Weibull Distribution Comparison	P. A.
5. Conclusions	P. A.

atmosphere

MDPI

Article

Exploring the Potential of Sentinel-1 Ocean Wind Field Product for Near-Surface Offshore Wind Assessment in the Norwegian Arctic

Eduard Khachatrian [1,*,†], Patricia Asemann [1,*,†], Lihong Zhou [1], Yngve Birkelund [1], Igor Esau [1,2] and Benjamin Ricaud [1]

[1] Department of Physics and Technology, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway
[2] Nansen Environmental and Remote Sensing Center, NO-5007 Bergen, Norway
* Correspondence: eduard.khachatrian@uit.no (E.K.); patricia.asemann@uit.no (P.A.)
† These authors contributed equally to this work.

Abstract: The exploitation of offshore wind resources is a crucial step towards a clean energy future. It requires an advanced approach for high-resolution wind resource evaluations. We explored the suitability of the Sentinel-1 Level-2 OCN ocean wind field (OWI) product for offshore wind resource assessments. The SAR data were compared to in situ observations and three reanalysis products: the global reanalysis ERA5 and two regional reanalyses CARRA and NORA3. This case study matches 238 scenes from 2022 for the Goliat station, an oil platform located 85 km northwest of Hammerfest in the Barents Sea, where a new offshore wind park has been proposed. The analysis showed that despite their unique limitations in spatial and temporal resolutions, all data sources have similar statistical properties (RMSE, correlation coefficient, and standard deviation). The Weibull parameters characterizing the wind speed distributions showed strong similarities between the Sentinel-1 and all reanalysis data. The Weibull parameters of the in situ measurements showed an underestimation of wind speed compared to all other sources. Comparing the full reanalysis datasets with the subsets matching the SAR scenes, only slight changes in Weibull parameters were found, indicating that, despite its low temporal resolution, the Sentinel-1 Level 2 OWI product can compete with the more commonly used reanalysis products in the estimation of offshore wind resources. Its high spatial resolution, which is unmatched by other methods, renders it especially valuable in offshore areas close to complex coastlines and in resolving weather events at a smaller scale.

Keywords: offshore wind; renewable energy; synthetic aperture radar; reanalysis data; Sentinel-1; Arctic

Citation: Khachatrian, E.; Asemann, P.; Zhou, L.; Birkelund, Y.; Esau, I.; Ricaud, B. Exploring the Potential of Sentinel-1 Ocean Wind Field Product for Near-Surface Offshore Wind Assessment in the Norwegian Arctic. *Atmosphere* **2024**, *15*, 146. https://doi.org/10.3390/atmos15020146

Academic Editor: William Cheng

Received: 26 December 2023
Revised: 17 January 2024
Accepted: 22 January 2024
Published: 24 January 2024

1. Introduction

The rapid exploitation of renewable and sustainable energy sources plays a pivotal role in the resolution of the current global environmental crisis. In this context, offshore wind power production has emerged as one of the most promising technologies available. According to the International Energy Agency, offshore wind has the potential to cover the current global energy demand many times over [1]. Innovations in turbine capacities, foundations, and energy transmission continue to promote this technology and drive down the costs of offshore wind energy production. Despite its high regional advantage, Norway's offshore wind resources remain largely untapped [1].

The first step towards the realization of new offshore wind power plants is the selection of high-quality locations, that is, those with high mean wind speed and minimal wind variability while being relatively close to the shore. Preliminary wind resource assessments are particularly difficult in offshore regions due to the sparseness of in situ observations. Reanalysis products offer a more consistent and long-term coverage. However, they also entail a number of drawbacks such as low spatial resolution, poor performance in coastal

areas, and a general underestimation of wind speed compared to in situ measurements [2,3]. Remote sensing data from synthetic aperture radar (SAR) offers a complementary source of information on near-surface offshore wind.

In light of the ambitious goals of the Norwegian government in terms of developing offshore wind, a comprehensive evaluation and intercomparison of all sources of wind data available for the Norwegian offshore region is crucial. Most offshore wind energy projects commissioned today are placed within 50 km of the shore; thus, reliable wind data sources for coastal regions are needed. When predicting wind energy production, an accurate estimation of the wind speed distribution is of essential importance as the wind power output is proportional to the cubed wind speed. Although the offshore wind resources in the North and the Norwegian Sea have previously been assessed, to the best of our knowledge, the wind resources in the Barents Sea have not been addressed in the existing literature.

This paper aims to evaluate the suitability of the Sentinel-1 OWI component for near-surface offshore wind energy assessments. We compared it to in situ measurements and three reanalysis products: the global reanalysis ERA5 and the regional reanalyses CARRA and NORA3, within a one-year case study for the Goliat station in the Barents Sea, presenting a comprehensive assessment of wind data sources available for the Norwegian Arctic offshore region. Wind data were compiled from all sources for the year 2022 and a statistical analysis was performed. This study is of particular interest to GoliatVIND, a floating offshore wind demonstration project currently in the planning stage that will be located near Goliat to supply renewable energy to the oil platform, reducing its carbon emissions [4].

The paper is organized as follows: Section 2 provides the theoretical background containing relevant research. Section 3 presents the utilized wind datasets and applied analysis methodology. Section 4 provides the results and discussion. Finally, the conclusions are presented in Section 5.

2. Theoretical Background

2.1. Reanalysis Products

Reanalysis data are produced by assimilating historical weather observations into a numerical weather prediction (NWP) model. The reanalysis datasets that are publicly available are often characterized by a low spatial resolution, typically ranging from 3 to 50 km. ERA5, CARRA, and NORA3 have each been validated against observations in previous studies.

Although ERA5 performs well in offshore areas and has been shown to generally outperform other global reanalysis products, its low spatial resolution of 31 km causes a significant performance decrease in coastal areas due to the changes in surface roughness and topology not being discernible in the dataset [2]. This is important in the coastal areas of northern Norway, where the complex mountainous coastlines and fjords can produce intricate wind fields that extend far out over the ocean. In Figure 1, where the wind comes from inland, this is evident in the long black shadows in the Sentinel-1 backscatter caused by the complex terrain. Beyond that, the low spatial resolution of ERA5 restricts its ability to adequately resolve mesoscale features. Several studies illustrated the weaknesses of ERA5 in the representation of weather phenomena characteristic to the Arctic climate system, such as polar lows, which would have a considerable effect on potential wind farms due to the occurring high wind speeds [5,6].

Reanalysis products generally exhibit the tendency to underestimate mean wind speed as compared to in situ measurements. The strongest tendency for the reanalysis products utilized here was found for ERA5 [3]. While this underestimation seems to be proportional to the ruggedness of the terrain, it is still present for terrain with low ruggedness, e.g., the ocean surface. In addition, a negative correlation of this underestimation with latitude was found. Both of these results reduce the impact of this tendency on the Norwegian Arctic offshore region. Regional reanalysis products tend to be more adequate to represent

smaller-scale and region-specific weather events and the spatial and temporal variability of wind. This was shown to be the case for both CARRA and NORA3; although a slight underestimation of mean wind speed was still found, they improved the representation of the spatial variability, temporal variability, and climatology of extremes in both offshore and coastal regions [5–7]. For both regional reanalysis products, a substantial part of the improvements compared to ERA5 were attributed to the higher spatial resolution.

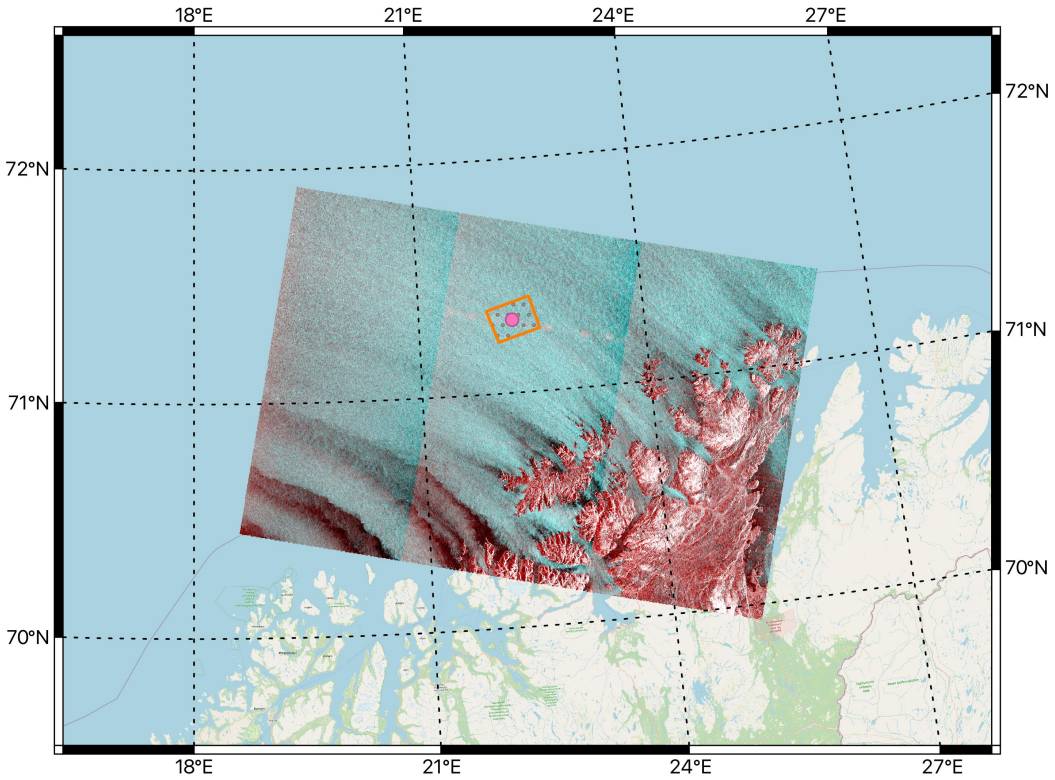


Figure 1. Illustration of the geographical area investigated in this study. The Goliat station (pink dot), located in the Barents Sea, 85 km north of Hammerfest; the search polygon for SAR data (orange rectangle); and false-color composite representation (HH, HV, and HV polarizations as RGB) of Sentinel-1 SAR images for 3 January 2022. The false-color composite scene illustrates the intensity of the backscatter in dB, where higher values, i.e., bright areas, correspond to the rougher surface and, therefore, higher wind speed, and darker areas correspond to a smooth ocean surface and, thus, lower wind speed.

2.2. Sentinel-1 Level-2 OWI Product

Wind data retrieval from satellite observations offers valuable complementary insights into offshore wind resources. Although a number of studies have attempted to utilize satellite scatterometers for offshore wind energy assessment, their low spatial resolution provides no advantage over reanalysis products [8]. SAR is capable of producing near-surface offshore wind speed data with a grid size of 1 km, a significantly higher resolution than all reanalysis products presented here. SAR operates at a variety of characteristics, such as frequency bands, polarization channels, and spatial resolutions, and responds to dielectric properties, geometry, and roughness, and to an object's surface or volume structure, depending on the penetration depth of the signal. One of the most crucial advantages of this type of sensor is the complete independence of solar illumination and weather conditions due to wavelengths that can penetrate dense clouds [9]. This is especially important for polar areas since dense cloud cover and long periods of darkness prevail there for several months of the year.

Wind speed can be estimated from SAR sensors through the measurement of the backscatter of the sea surface. As capillary waves and, thus, sea surface roughness, are directly dependent on atmospheric forcing, the sea backscattering in the C-band can be understood as a function of wind speed, wind direction (relative to the look angle), and incidence angle. However, other influences such as oil slicks, strong surface currents, or reflection from nearby coastlines can distort this proportionality. Furthermore, different polarization signals inhibit different limitations in the accurate representation of either low or high wind speeds: while co-polarized signals saturate with increasing wind speeds, cross-polarized signals cannot be distinguished from instrument noise at low wind speeds. Even a combined signal wind retrieval approach requires a priori information about the wind direction most commonly provided by NWP models [10]. Lastly, as no information about atmospheric stratification is available, wind retrieval at 10 m above the surface from SAR utilizes a geophysical model function assuming neutral stability. Therefore, the method cannot respond to changes in wind speed due to atmospheric stratification [11].

The key limitation of wind products acquired from SAR technology is its temporal resolution. Therefore, even when collecting a long time series, SAR cannot provide the same amount of temporal information as reanalyses or in situ observations. Barthelmie and Pryor conducted a detailed study assessing the suitability of limited satellite time series data for estimating the stochastic parameters of the wind speed distribution [12]. They showed that when accepting an uncertainty of $\pm$10% at a confidence level of 90%, only 150 (randomly selected) scenes are required to characterize both the mean wind speed and the variance sufficiently accurate. However, in the case of Sentinel-1, the scenes are not randomly selected but exhibit a strong temporal selectivity, as the satellite revisits any specific area only at certain times of the day. This provokes a discussion of the diurnal variability present in the investigated wind fields. Although coastal areas exhibit diurnal cycles largely due to advective effects, no strong diurnal patterns are expected over the open ocean.

To the best of our knowledge, the Sentinel-1 Level-2 OWI component has only been assessed in a small number of case studies, each comparing it to in situ observations, e.g., around Ireland [8], around Cyprus [13], and in the Ionian Sea [14]. Similar to the reanalyses, a tendency of the Sentinel-1 OWI component to underestimate wind speed was found, where the bias linearly decreased with increasing wind speed. When estimating average wind power, the errors as compared to results obtained from in situ measurements were between 5 and 10%. However, none of these studies compares the Sentinel-1 data to the reanalyses, which are more commonly used for wind resource assessments than in situ observations. Off the coast of Norway, where measurement stations are especially sparse and the complex coastlines produce intricate wind field patterns, this intercomparison will be of great value for the future of wind resource assessments, further reducing project risk at the site-finding stage.

3. Methods

3.1. Datasets

In this study, five different wind speed and wind direction data sources were analyzed: in situ observations, three reanalysis products, including one global reanalysis product and two regional reanalysis products, and a SAR-based remote sensing product. The following subsections briefly present the technical details of each data source used for the analyses. Table 1 displays the datasets used in this study along with the main characteristics and their crucial differences.

Table 1. List of datasets used in this study along with their crucial characteristics. Even though in situ observations have a high spatial resolution, the singular point corresponds to a certain small area (point on a map), which does not allow any further information about the broader surrounding area to be obtained.

Dataset	Type	Resolution	
		Temporal	Spatial
Seklima	in situ observation	every 20 min	singular point
ERA5	Global reanalysis	hourly	31 km
NORA3	Regional reanalysis	hourly	3 km
CARRA	Regional reanalysis	every 3 h	2.5 km
Sentinel-1	SAR-based	once in 1–2 days	1 km

3.1.1. In Situ Observations

For this study, measurements provided by the Norwegian Center For Climate Services (Seklima) were used, particularly from a single Norwegian offshore station, Goliat Fpso (Station number: SN76956; Goliat), located in the Barents Sea (71.31° N, 22.25° E) [15]. The sensor is located 46 m above mean sea level and has been operating from 9 September 2015. It provides wind speed and wind direction information, extrapolated to 10 m.a.s.l. by employing a power law wind profile with a wind shear coefficient of 0.13 [16]. From all data sources investigated in this study, this method has the highest temporal resolution, providing measurements every 20 min, averaged over 10 min intervals.

As Figure 1 illustrates, the wind field around and at the Goliat station is influenced by the complex coastline if the wind comes from inland. The sea surface at this location is not expected to be affected by wave reflections off the coast, rendering this a suitable place for comparison. However, it also limits the applicability of the results of this study to areas closer to the shore or even within fjords. This specific location was recently selected for the construction of an offshore wind farm. Thus, comparing the measurement data from this site with other available wind data sources can provide unique complementary insights into the design and implementation of wind power, not only for the Goliat station, but also for future offshore wind power exploitation in general.

3.1.2. Reanalyses

ERA5

ERA5 is the fifth-generation global reanalysis model developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) [17]; it is publicly available through the Copernicus Climate Change Service. It has a spatial resolution of approximately 31 km, rendering it unable to resolve mesoscale features and complex topology along coastlines. Its temporal resolution, however, is very high, providing instant values of hourly wind speed and direction. In this study, we used the "10 m u-component of wind" and "10 m v-component of wind" from the ERA5 hourly data on single levels, calculating the wind speed via $\sqrt{u^2 + v^2}$ [18].

NORA3

NORA3 is a regional reanalysis model produced by the Norwegian Meteorological Institute covering the North Sea, the Norwegian Sea, and the Barents Sea (44.02° N–84.06° N, 30.17° W–85.79° E) with a spatial resolution of 3 km. The NORA3 data are obtained by downscaling ERA5 data using a nonhydrostatic convection-permitting numerical weather prediction model (HARMONIE-AROME) [19]. Along with the wind field, NORA3 provides various atmospheric and surface meteorological parameters, such as mean sea level pressure, air temperature and relative humidity, fog, wind speed, and direction [20]. Solbrekke et al. demonstrated that the wind field in NORA3 is much improved relative to its

host analysis, particularly in mountainous areas and along the improved grid-resolving coastlines. In this study, we used NORA3 as one of the reanalysis sources since it is specifically designed and validated for Norway and has a high spatial and temporal resolution, providing instant values of wind speed and direction at 10 m.a.s.l. for every hour.

CARRA

CARRA (Copernicus Arctic Regional Reanalysis) was the second regional reanalysis product used in this study. It contains 3-hourly analyses along with hourly short-term forecasts at a 2.5 km resolution of atmospheric and surface meteorological parameters, such as surface and near-surface temperature, precipitation, humidity, wind, pressure, and atmosphere fluxes. It is divided into CARRA-East and CARRA-West; the west domain covers Greenland along with the neighboring seas and territories. In this study, we were interested in the East domain, which covers Svalbard, Franz Josef Land, Novaya Zemlya, and the northern parts of Scandinavia [21]. Similar to NORA3, CARRA was also created by the HARMONIE-AROME with the ERA5 global reanalysis as lateral boundary conditions. In addition, some enhancements have been implemented in comparison to both ERA5 and the operational HARMONIE-AROME modeling systems, including extensive utilization of satellite data for the HARMONIE-AROME operational weather prediction system, significant augmentation in the surface observation datasets, and substantial improvements in the regional physiography and orography [7]. Furthermore, CARRA utilizes an improved data assimilation system compared to NORA3.

3.1.3. Remote Sensing

The OWI component from the Sentinel-1 Level-2 OCN product was acquired through the publicly available Copernicus Open Access Hub, the European Union's Earth observation program. Sentinel-1 operates at the C-band with a central frequency of 5.404 GHz and includes two polar-orbit Sentinel-1A and Sentinel-1B missions that provide multiple sensing modes, such as stripmap (SM), extra-wide (EW), wave (WV), and interferometric-wide (IW) swath modes in single (HH or VV) or dual polarization (HH + HV or VV + VH) at a 40 m spatial resolution. SAR reaches equilibrium within less than one minute. The Sentinel-1 Level-2 OCN OWI component is fully calibrated and is provided as an ocean surface wind vector, including wind speed and wind direction, estimated from Sentinel-1 Level-1 SAR images by inversion of its associated normalized radar cross section (NRCS). It is a ground range gridded estimate of the surface wind speed and direction at 10 m above the surface with a spatial resolution of 1 km.

In this study, we focused on the extraction of the available Sentinel-1 Level-2 OCN scenes for the Goliat station for 2022. For Northern Norway, the acquisition mode is IW, which is acquired using the TOPSAR technique, providing an improved quality product by enhancing image homogeneity. Sentinel-1 revisits the region of interest of this study once every 1–2 days, passing either at around 5:00 UTC or 16:00 UTC.

3.2. Temporal and Spatial Preprocessing

The 238 SAR scenes available from Sentinel-1 between January and December 2022 for the area covering the Goliat station were extracted, providing wind speed and direction information at 10 m above the ocean surface. Subsequently, the same information from all the above-mentioned data sources, including the Seklima in situ observations and the three reanalyses ERA5, CARRA, and NORA3 were acquired, overlapping with the Sentinel-1 data both spatially and temporally. In the case of low-resolution ERA5, this matchup was performed by extracting the single point of the output grid closest to the Goliat station. For CARRA, NORA3, and Sentinel-1, which have higher spatial resolutions, the mean value of the four closest points was computed. The Seklima data were employed as the reference source, as in situ measurements are commonly assumed to provide the most accurate observations.

The data sources differ in temporal and spatial resolution. The Seklima data have the lowest spatial resolution (singular) and highest temporal resolution, providing the wind speed and direction data every 20 min. The global ERA5 reanalysis provides wind data at a low spatial resolution (31 km); however, it has a high temporal resolution with data provided every hour. On the other hand, CARRA reanalysis has a significantly higher spatial resolution (3 km), but relatively lower temporal resolution (3 h). The optimal source from both spatial and temporal resolution points of view is NORA3 with a spatial resolution of 3 km, which is provided hourly. Sentinel-1 provides wind produce with the highest spatial resolution (1 km), but at the expense of temporal resolution, with a maximum of one scene per day.

3.3. Statistical Analysis

In order to properly compare the differences between all the above-mentioned wind data sources, and to quantitatively evaluate the accuracy of the Sentinel-1 SAR-based wind speed and direction product, we applied several commonly employed stochastic measures to the full datasets, such as root mean squared error (RMSE), Pearson correlation coefficient, and standard deviation [22]. RMSE is one of the most commonly used measures for evaluating the quality of forecasts/predictions. It illustrates the deviation of a prediction from the reference value for each point of a time series by calculating the Euclidean distance between the two. The Pearson correlation coefficient captures the amount of linear correlation between two variables, ranging from -1 to +1. The standard deviation displays the spread of the data points around the mean value from the dataset, thus showing the dispersion in a set of values.

3.4. Wind Speed Distribution

The stochastic character of wind speed allows its representation by a probability density function. For the assessment of offshore wind resources in a given region, an accurate representation of the wind speed distribution is crucial. The Weibull distribution $f(v, \lambda, k)$ is widely used for this purpose as it generally provides a good fit to offshore wind speed data [23]:

$$f(v, \lambda, k) = \frac{k}{\lambda}\left(\frac{v}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{v}{\lambda}\right)^{k}\right) \quad (1)$$

with the wind speed $v \in \mathbb{R}^+$, the scale parameter $\lambda > 0$, and the shape parameter $k > 0$. The scale parameter λ specifies the ratio of horizontal to vertical extent of the distribution, i.e., an increase in λ results in a wider and lower Weibull distribution. The shape parameter k indicates the skewness of the distribution; a larger value for k, therefore, indicates more frequent high-wind-speed events. In a continuation of their work, Pryor et al. evaluated the critical sample size of independent SAR scenes needed to characterize the Weibull parameters. They came to the conclusion that no more than 250 observations are needed for an uncertainty of $\pm 10\%$ at a confidence level of 90% for the characterization of both Weibull parameters [24].

For an estimation of the annual wind power production of a wind turbine at a given site, the wind speed would first have to be extrapolated vertically to the hub height of the turbine. There are a number of models commonly used for this purpose, each with its own limitations. In heterogeneous terrain, models such as the logarithmic law or the power law are recognized as applicable well up to heights of 200 m and 300 m above the ground, respectively [25]. However, as wind turbines are rapidly increasing in size, especially offshore and for floating foundations (e.g., the Vestas V236-15.0 MW model by Vestas Wind Systems A/S in Aarhus, Denmark reaches a rotor diameter of 236 m), and as the atmospheric boundary layer above the smooth ocean surface is much shallower than over complex or rough terrain, the case discussed in this study extends far beyond the applicability of these models. More complex models require knowledge about many more atmospheric parameters that were not available to us for the considered location. Therefore, in this study, we limit the analysis of the available wind power to the comparison of the

Weibull distributions fitted to the 10 m wind speed histograms of each data source, as the wind speed distribution and wind power output are closely related. Although it is outside of scope of this study, the authors note the necessity to explore this issue further and find alternative solutions in order to maintain the performance of offshore wind resource estimation in the future.

4. Results and Discussion

4.1. Wind Speed

As the wind power available at a given site is proportional to the cubed wind speed, a precise assessment of the wind speed distribution and of the uncertainty present in the available data is of critical importance for wind energy purposes. It is a crucial parameter that directly impacts wind energy potential, influences the capacity factor and profitability of a given location, and affects the choice of wind turbine. Since wind turbines are optimized for particular wind speed ranges, the wind speed distribution present at a given site plays a key role in the selection of the turbine type. Wind turbines used for offshore power production usually have a cut-in wind speed at 2–3 m/s (at hub height), a rated wind speed of 14–16 m/s, and a cut-off wind speed between 25 and 30 m/s. Therefore, the interval ranging from the cut-in to the cut-off wind speed, and especially between the cut-in and the rated wind speed, has to be represented with high accuracy for wind power assessment purposes. It is worth noting that, as most wind data sources provide the wind speed at 10 m above the ocean surface, the wind speed interval of interest is shifted relative to that at the hub height.

4.1.1. Scatterplots

Figure 2 shows scatterplots obtained from comparing the wind speed from the reference source, i.e., Seklima in situ observations, to each other source, including ERA5, NORA3, and CARRA reanalyses, and Sentinel-1. The Seklima data are plotted on the horizontal axis, all other sources are plotted along the respective vertical axes. The background heatmaps provide extra insight into the distribution of wind speed across the scatterplot. Accordingly, darker regions in the heatmap signify a higher density of data points. The scatterplots clearly demonstrate a very strong positive linear relationship between all pairs of datasets shown. The close clustering of data points suggests a strong similarity between the Seklima dataset and all three reanalyses, and between Seklima and the Sentinel-1 data. Linear regressions were fitted to each scatterplot, with the slope and intercept parameters varying only slightly between data sources.

4.1.2. Statistical Metrics

In order to quantitatively evaluate the statistical differences between the above-mentioned sources, a number of metrics, including RMSE, Pearson correlation coefficient, and standard deviation, were calculated for Sentinel-1, ERA5, NORA3, CARRA, each with respect to Seklima. Table 2 presents the results, while Figure 3 shows the Taylor diagram, visualizing all the above-mentioned metrics for all data sources in a single plot.

Table 2. Statistical metrics applied to wind speed time series from each data source compared to the reference source, Seklima, in situ observations.

Dataset	RMSE [m/s]	Correlation Coeff.	Std. Deviation [m/s]
Seklima	-	-	4.19
ERA5	1.65	0.92	3.82
NORA3	1.80	0.90	3.98
CARRA	2.05	0.88	4.03
Sentinel-1	2.00	0.88	3.95

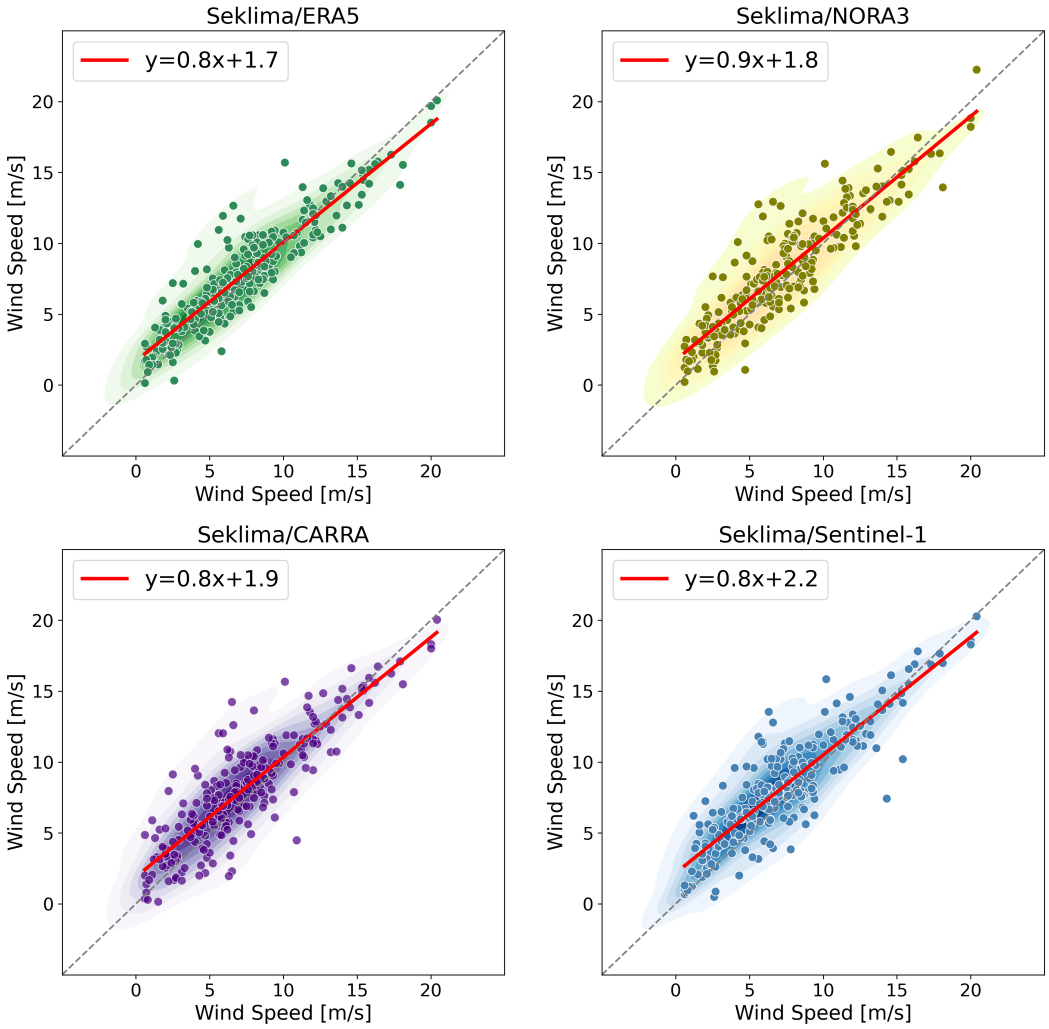


Figure 2. Scatterplots of wind speed data from Seklima (plotted on the horizontal axes) versus ERA5, NORA3, CARRA, and Sentinel-1 (plotted on the respective vertical axes).

The RMSEs of the four data sources ranged from 1.65 m/s to 2.05 m/s. ERA5 had the lowest RMSE with 1.65 m/s, indicating a slightly more accurate model compared to those with higher RMSEs, namely, CARRA with a maximum value of 2.05 m/s. Despite the range, it is worth noting that all four data sources had relatively close RMSE values, demonstrating similar levels of accuracy. Moreover, we can highlight that the RMSE of Sentinel-1 fell within the range of RMSE values obtained for the three reanalyses, displaying a similar closeness between the respective wind speed time series. The Pearson correlation coefficient ranged between 0.88 for both CARRA and Sentinel-1 and 0.92 for ERA5, again placing the SAR-based data within the values obtained for the three reanalyses, and indicating a high overall correlation with the reference data. While the scatterplots only visually demonstrated the relationship between the Seklima data with the Sentinel-1 and the reanalyses data, the correlation coefficient confirmed that there was a strong linear relationship between the compared sources and the reference data. Table 3 displays the cross-correlation between the wind speed data sources. This specifically demonstrates the close relationship between the global reanalysis ERA5 and the two regional reanalyses CARRA and NORA3. This likely results from CARRA and NORA3 employing ERA5 as their host model.

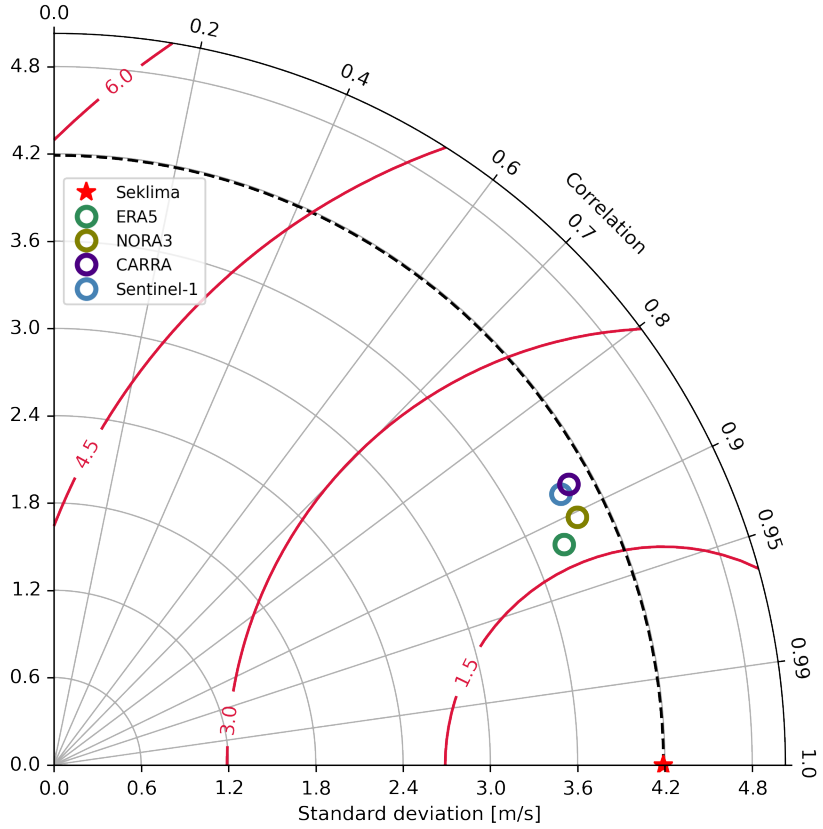


Figure 3. Taylor diagram for wind speed data from Seklima versus ERA5, NORA3, CARRA, and Sentinel-1. The red contours display the RMSE value of the data sources compared to the reference source.

Table 3. Cross-correlation matrix between the different wind speed data sources, i.e., Seklima, ERA5, NORA3, CARRA reanalysis, and Sentinel-1.

Variables	Seklima	ERA5	NORA3	CARRA	Sentinel-1
Seklima	1	0.92	0.90	0.88	0.88
ERA5		1	0.98	0.93	0.91
NORA3			1	0.94	0.91
CARRA				1	0.88
Sentinel-1					1

The reference data (Seklima—red star) had a higher standard deviation than all other sources, indicating the largest amount of wind speed variability. This is interesting as it is the only source that averages its output values over a longer time interval (10 min); all other sources provide instantaneous or pseudo-instantaneous values. Since shorter sampling periods lead to higher gusts being captured in the data as compared to long averaging periods, the difference in averaging periods present in the data sources investigated within this study could have led to the hypothesis that the reanalyses and the Sentinel-1 data display a higher wind speed variability, and thus, variance. The contrary seems to be the case here, as none of the other data sources were able to fully capture the wind speed variance displayed by the in situ observations; all other data sources had a slightly lower standard deviation than Seklima. ERA5 had the lowest standard deviation of 3.82, thus representing the least amount of variability in wind speed. Nevertheless, it should be noted that the standard deviation did not range significantly among the five data sources, which implies a similar variation of values among the datasets.

4.1.3. Wind Speed Occurrences

Figure 4 shows histograms of the differences in wind speed occurrences acquired by several sources between the reference source, Seklima, and the other data sources, namely, ERA5, NORA3, CARRA, and Sentinel-1. The horizontal axis represents the range of wind speeds in m/s, while the vertical axis displays the difference of wind speed occurrence in percent. Accordingly, when the occurrence frequency difference is 2, it shows that the corresponding wind speed occurred 2% more often in the reference data than the compared source. Moreover, each histogram displays the maximum, minimum, and variance of the differences.

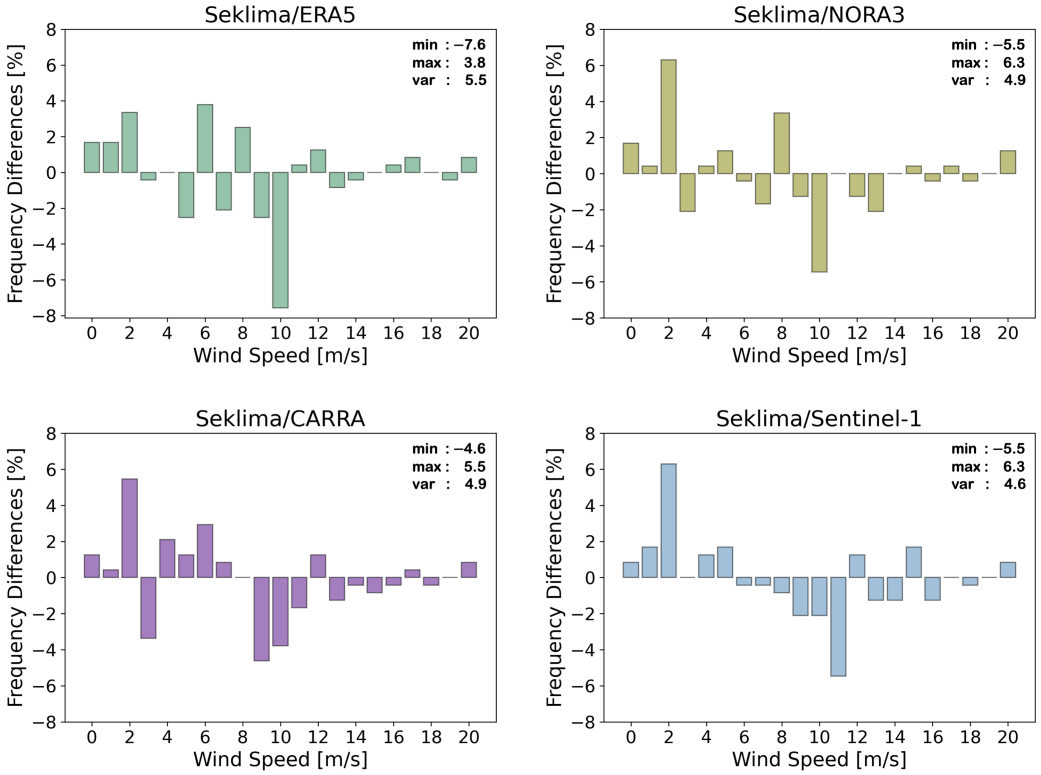


Figure 4. Histograms of the differences in wind speed occurrences between Seklima and the other data sources, i.e., ERA5, NORA3, CARRA reanalysis, and Sentinel-1.

Most of the wind speed differences fell within the range of −2 to +2%, suggesting a good level of agreement for those wind speed intervals. However, for all data sources, there were higher values of disagreement present, both on the positive and negative side. Generally, the outliers tended to be either consistently positive or negative throughout all combinations of data sources, indicating a similar bias present in all reanalyses and the Sentinel-1 data as compared to the in situ measurements. All sources displayed an under-representation of the 2 m/s wind speed, which might correspond to difficulties in capturing low-wind events. The most significant over-representation was detected for wind speed values around 9–11 m/s.

Another interesting point is that Sentinel-1 had the lowest variance in wind speed difference among all the sources, which means that the data points were relatively consistent and indicated a higher level of accuracy in the wind speed prediction than other sources. On the other hand, ERA5 provided the highest variance, which means that the spread of the values was larger and the wind speed precision might not have been very consistent. However, it should be mentioned that the variance differences were small between the sources, with the highest variance equal to 5.5 (ERA5) and the lowest to 4.6 (Sentinel-1).

The histograms provide valuable insight into the reliability and consistency of wind speed data obtained from different sources by highlighting intervals that are in need of

further calibration. Overall, they give rise to the impression that all reanalyses and the SAR-based product tended to underestimate low wind speeds and overestimate high wind speeds as compared to the in situ observations that were employed as reference data. It should be noted that these discrepancies do not necessarily indicate a bad performance of the reanalyses or the Sentinel-1 product; the fact that there were differences similar in tendency and magnitude that were consistent for all compared data sources suggests that there might have been a bias present in the in situ measurement data. As a matter of fact, a study investigating vertical wind profiles for nine oil platforms in the south of Norway concluded that the power law with a wind shear coefficient of 0.13 that is presently employed on Norwegian oil platforms results in an average 0.8 m/s underestimation of the 10 m wind speed [16]. As these observations are, in turn, used for the assimilation and verification of reanalysis products, these findings highlight the need for a more independent, yet accurate representation of the wind speed distribution.

4.2. Wind Direction

Another important parameter in the preliminary site assessment for wind energy purposes is wind direction. Along with wind speed, it plays a fundamental role in optimizing the wind energy generation of a given wind farm. In particular, information about the distribution of the wind direction significantly influences the wind park layout and, thus, aids in maximizing the energy output by avoiding wake losses within the wind farm. Furthermore, the complex topology of the fjords and mountains in the coastal areas of Northern Norway heavily affects the wind field over the ocean, as can clearly be seen in Figure 1. In order to choose the optimal location, a high spatial resolution capable of resolving the intricate wind patterns present in these areas is essential.

Figure 5 shows the wind roses of all data sources, displaying the occurrences of wind directions and a rough representation of the wind speed distribution for each wind direction bin, allowing for a visual comparison between sources. Overall, all diagrams display similar prevailing wind directions, namely, mainly from the East and West South West. Some differences are discernible, e.g., the wind direction distribution of ERA5 seemed to fluctuate less than all other data sources, which may result from its low spatial resolution and account for its inability to resolve the flow over the complex topology of the coastline nearby. Furthermore, while Sentinel-1 represented the West wind best, its representation of the wind coming from the East quartile was broader than that of Seklima and the two regional reanalyses. However, more detailed and quantitative analyses would be needed in order to properly assess the performance of the four data sources in capturing the distribution of wind direction.

4.3. Wind Speed Weibull Distribution Comparison

The importance of an accurate representation of the wind speed distribution has already been discussed in this study. In order to assess the ability of the SAR-based product to capture the stochastic properties of the wind speed despite its low temporal resolution, a comparison of the sparse match-up datasets of the reanalyses and in situ observations with their respective dense datasets, that is, with their full temporal resolution, is appropriate. The first row of Figure 6 shows the wind speed histograms of the dense and the sparse Seklima data, and the Sentinel-1 data, alongside their respective Weibull distribution. The histograms and Weibull distributions of the three reanalysis products, dense and sparse, are shown in the second and third row, respectively. Additionally, Table 4 summarizes the corresponding Weibull parameters obtained for the dense and sparse wind speed datasets from all sources. The Weibull distribution obtained from the dense Seklima data was reproduced in all other plots as a reference.

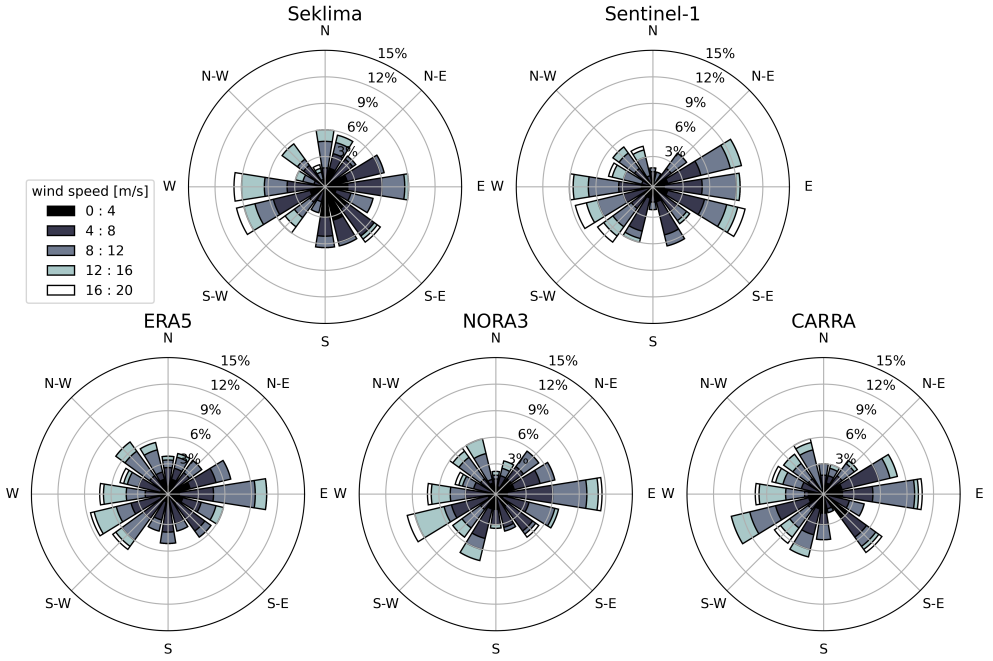


Figure 5. Wind roses compiling the wind direction distributions from the Seklima, ERA5, NORA3, CARRA, and Sentinel-1 datasets.

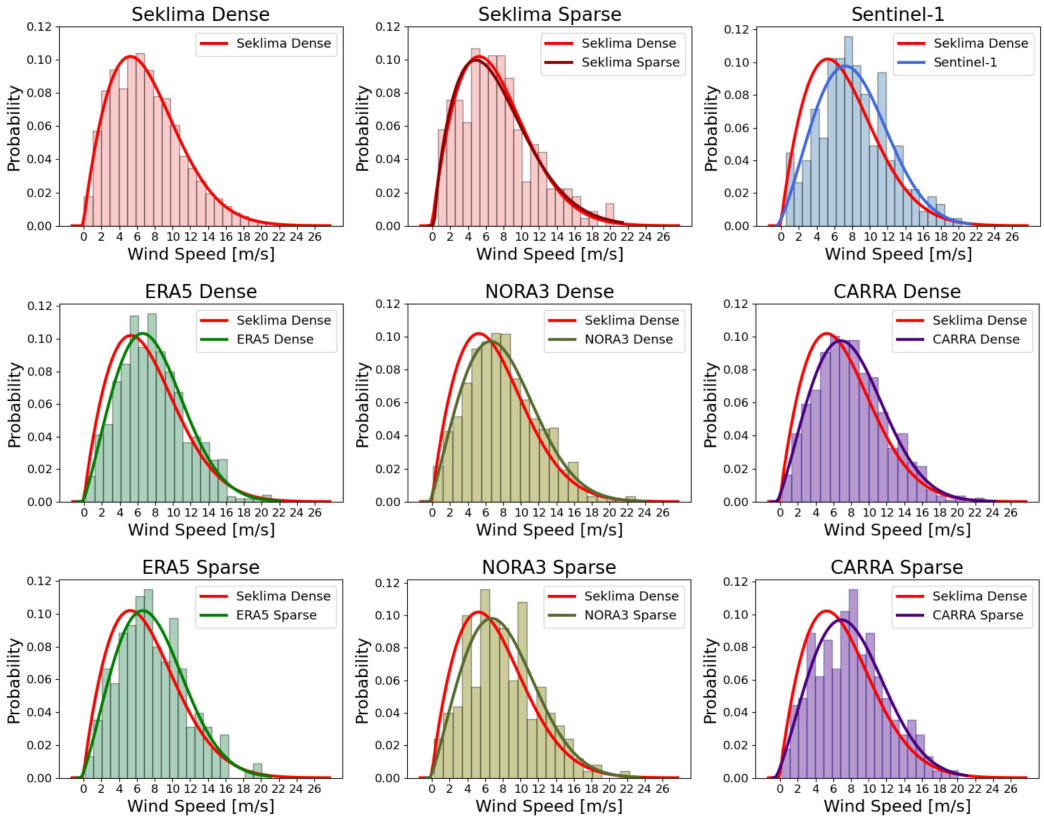


Figure 6. Weibull probability density function and observed wind speed histograms for several cases: dense and sparse datasets from Seklima, ERA5, NORA3, and CARRA, and Sentinel-1. *Dense* refers to the sample obtained for the Goliat station over the year 2022 with the respective full temporal resolution, that is, every 20 min for Seklima, and several times a day for the three reanalysis sources. *Sparse* refers to the subsets of the 238 data points overlapping temporally with the Sentinel-1 scenes.

Table 4. The Weibull shape and scale parameters k and λ obtained for the dense and sparse wind speed datasets from all sources.

Dataset		Weibull Parameters	
		k	λ
Seklima	dense	1.85	8.01
	sparse	1.70	7.83
ERA5	dense	2.21	8.90
	sparse	2.23	9.08
NORA3	dense	2.14	9.22
	sparse	2.17	9.23
CARRA	dense	2.22	9.43
	sparse	2.26	9.67
Sentinel-1	(sparse)	2.33	9.79

The Weibull parameters obtained for the Sentinel-1 dataset are $k = 2.33$ and $c = 9.79$, clearly illustrating an overestimation of the distributed wind speed compared to the reference data. This result strongly contrasts with those obtained in previous studies (see Section 2), concluding that, for other specific regions, Sentinel-1 and the reanalysis products generally tended to underestimate wind speed as compared to the in situ observations. However, as discussed previously, Olsen et al. found that wind speed data obtained from measurements at Norwegian oil platforms tended to underestimate wind speed due to a flawed vertical wind profile [16]. This provides a valid explanation for the discrepancy found here. This assumption is supported by the fact that the Weibull parameters of all three reanalysis products were not just close to one another, but also very close to those of the Sentinel-1 dataset.

Although there is a visible difference between the dense and the sparse Seklima histogram, their Weibull distributions were still considerably close. The change in shape and scale parameters from the dense to the sparse sample was $k = 1.85 \rightarrow 1.70$ (a decrease of $\sim$8.1%) and $\lambda = 8.01 \rightarrow 7.83$ (a decrease of $\sim$2.25%), respectively. The differences between parameters for the dense and sparse reanalysis datasets were even less for all three reanalyses (no more than a 2% change in k and 2.5% in λ). Pryor et al. determined that around 250 independent scenes are needed to characterize these parameters [24]. With these results achieved using 238 scenes, we are pleased to note that in this case, the temporal selectivity due to the sampling rate of Sentinel-1 did not seem to introduce a major bias into the distributions.

However, minor tendencies are observable. With decreasing temporal resolution, both k and λ decreased for Seklima but slightly increased for all three reanalysis products. For the latter, going from the dense to the sparse datasets, this resulted in lower and wider distributions with more pronounced tails and, thus, a higher mean wind speed. The opposite happened to the Seklima data; the distribution became slimmer and slightly less skewed, resulting in a lower mean wind speed. Thus, moving from the dense to the sparse datasets accentuated the discrepancy between the Seklima distribution and the four others. However, the relative changes of both k and λ, especially for the reanalysis products, were too small to be attributed to a potential bias resulting from temporal selectivity.

In the context of wind resource assessment, these results are valuable, but not enough. The average wind power of a site with a certain wind speed Weibull distribution is given by

$$P = \frac{1}{2}\rho\lambda^3\Gamma\left(1+\frac{3}{k}\right) \quad (2)$$

with the gamma function Γ [24]. The high order of the scale parameter λ explains why the accuracy achieved by the number of SAR scenes used here is insufficient to characterize

the annual average wind power. According to Barthelmie et al., almost 2000 independent scenes are required to achieve an uncertainty of ±10% at a confidence level of 90% [12]. For the Goliat site, this translates to almost 8.5 years of SAR data, which, for the Sentinel-1 OWI product evaluated here, are not yet available.

5. Conclusions

The presented analysis justifies the application of the Sentinel-1 OWI component for offshore wind resource evaluation. We found that the RMSE, correlation coefficient, and standard deviation were similar for all data sources, placing the performance of the SAR-based wind retrieval within the range of the three reanalysis products. An analysis of the wind speed differences displayed a clear bias between the in situ measurements and the other four data sources, which may result from an underestimation of the reduced wind speed as distributed from Norwegian oil platforms [16]. A qualitative review of the distribution of wind direction showed no distinct discrepancies between the data sources.

The analysis of the Weibull distributions fitted to each dataset led to a similar conclusion: while the Weibull parameters of the three reanalyses and the Sentinel-1 data were reasonably close to one another, those of the Seklima data differed considerably, resulting in a shift of the Weibull distribution towards lower wind speed values. The sparse reanalysis datasets, consisting of 238 points matching up with the SAR scenes, were compared with their respective dense datasets, containing their full temporal resolution. The change in Weibull parameters between the two variants was no more than 2.5% for both scale and shape parameters. This suggests the conclusion that the low temporal resolution of the Sentinel-1 data is not generally an obstacle in the accurate representation of the wind speed distribution. However, in order to perform a complete wind resource assessment with a reasonable accuracy, a longer time series would be required. Nonetheless, its unmatched spatial resolution renders it a valuable source of information for wind resource assessments in offshore areas affected by complex coastlines unresolved by other data sources. Furthermore, it offers the potential to resolve mesoscale features such as polar lows, which are often not captured by reanalysis products.

This study analyzed some of the similarities and dissimilarities present in the available data sources for assessing near-surface offshore wind in the Norwegian Arctic. However, it is far from exhaustive. Future studies could include both a spatial and temporal expansion of the data taken into consideration, enabling a more accurate representation of the stochastic characteristics of all sources. In this context, several questions could be addressed that arise naturally from this work: How does the Sentinel-1 OWI component perform under varying meteorological conditions, i.e., at the tails of the wind speed distribution? How severely is it affected by the coastline when moving closer to the shore or even into fjords? And how large are the errors introduced by atmospheric stratification that deviates from the assumed neutral stability? An in-depth analysis of spatial and temporal variability, such as diurnal, seasonal, and interannual variability, would also prove valuable to further characterize the wind climatology in the Norwegian Arctic and aid in the understanding of the limitations of the Sentinel-1 OWI component for offshore wind resource assessments. Lastly, a more holistic approach to wind resource evaluation procedures has to be developed in order to utilize the added value of SAR-based wind retrieval for wind farm planning and development.

Author Contributions: E.K.: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing—original draft, Writing—review and editing. P.A.: Conceptualization, Formal analysis, Investigation, Methodology, Writing—original draft, Writing—review and editing. P.A. and E.K. contributed equally to this work. L.Z.: Conceptualization, Data curation, Writing—review and editing. Y.B.: Conceptualization, Supervision, Writing—review and editing. I.E.: Conceptualization, Supervision, Writing—review and editing. B.R.: Conceptualization, Supervision, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by Equinor Akademiaavtalen with UiT—The Arctic University of Norway. Patricia Asemann is supported by a fellowship of the German Academic Exchange Service (DAAD). Igor Esau acknowledges a contribution from the ESA project MAXSS 4000132954/20/I-NB.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. Sentinel-1 SAR data is accessible on Copernicus Open Access Hub at https://dataspace.copernicus.eu/. ERA5 and CARRA data are available through the Climate Data Store https://cds.climate.copernicus.eu/, while NORA3 reanalysis is obtainable through https://marine.met.no/. Finally, Seklima data is available at the Norwegian Centre for Climate Services https://seklima.met.no/.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SAR	Synthetic Aperture Radar
OWI	Ocean Wind Field
NWP	Numerical Weather Prediction
ECMWF	European Centre for Medium-Range Weather Forecasts
SM	Stripmap Mode
EW	Extra-Wide Mode
WV	Wave Mode
IW	Interferometric-Wide Mode
NRCS	Normalized Radar Cross Section
RMSE	Root Mean Squared Error

References

1. International Energy Agency. *World Energy Outlook 2019*; International Energy Agency: Paris, France, 2019. [CrossRef]
2. Gualtieri, G. Reliability of ERA5 Reanalysis Data for Wind Resource Assessment: A Comparison against Tall Towers. *Energies* **2021**, *14*, 4169. [CrossRef]
3. Dörenkämper, M.; Olsen, B.T.; Witha, B.; Hahmann, A.N.; Davis, N.N.; Barcons, J.; Ezber, Y.; García-Bustamante, E.; González-Rouco, J.F.; Navarro, J.; et al. The Making of the New European Wind Atlas—Part 2: Production and evaluation. *Geosci. Model Dev.* **2020**, *13*, 5079–5102. [CrossRef]
4. Source Galileo. GoliatVIND. 2023. Available online: https://goliatvind.no (accessed on 17 October 2023).
5. Haakenstad, H.; Breivik, Ø.; Furevik, B.R. NORA3: A Nonhydrostatic High-Resolution Hindcast of the North Sea, the Norwegian Sea, and the Barents Sea. *Wind Energy Sci.* **2021**, *60*, 1443–1464. [CrossRef]
6. Moreno-Ibáñez, M.; Laprise, R.; Gachon, P. Recent advances in polar low research: Current knowledge, challenges and future perspectives. *Tellus A Dyn. Meteorol. Oceanogr.* **2021**, *73*, 1890412. [CrossRef]
7. Køltzow, M.; Schyberg, H.; Støylen, E.; Yang, X. Value of the Copernicus Arctic Regional Reanalysis (CARRA) in representing near-surface temperature and wind speed in the north-east European Arctic. *Polar Res.* **2022**, *41*, 8002. [CrossRef]
8. de Montera, L.; Remmers, T.; O'Connell, R.; Desmond, C. Validation of Sentinel-1 offshore winds and average wind power estimation around Ireland. *Wind Energy Sci.* **2020**, *5*, 1023–1036. [CrossRef]
9. Sandven, S.; Johannessen, O.M.; Kloster, K. Sea Ice Monitoring by Remote Sensing. *Encycl. Anal. Chem.* **2006**, *1993*, 1–43. [CrossRef]
10. Tollinger, M.; Graversen, R.; Johnsen, H. High-Resolution Polar Low Winds Obtained from Unsupervised SAR Wind Retrieval. *Remote Sens.* **2021**, *13*, 4655. [CrossRef]
11. ESA. Level 2 OCN Ocean Wind Field Component. Available online: https://sentinels.copernicus.eu/web/sentinel/ocean-wind-field-component (accessed on 31 August 2023).
12. Barthelmie, R.J.; Pryor, S.C. Can Satellite Sampling of Offshore Wind Speeds Realistically Represent Wind Speed Distributions? *J. Appl. Meteorol.* **2003**, *42*, 83–94. [CrossRef]
13. Hadjipetrou, S.; Liodakis, S.; Sykioti, A.; Katikas, L.; Park, N.W.; Kalogirou, S.; Akylas, E.; Kyriakidis, P. Evaluating the suitability of Sentinel-1 SAR data for offshore wind resource assessment around Cyprus. *Renew. Energy* **2022**, *182*, 1228–1239. [CrossRef]
14. Caligiuri, C.; Stendardi, L.; Renzi, M. The use of Sentinel-1 OCN products for preliminary deep offshore wind energy potential estimation: A case study on Ionian sea. *Eng. Sci. Technol. Int. J.* **2022**, *35*, 101117. [CrossRef]
15. Seklima. Norsk Klima Service Senter. Available online: https://seklima.met.no/ (accessed on 6 September 2023).

16. Olsen, A.M.; Øiestad, M.; Berge, E.; Køltzow, M.Ø.; Valkonen, T. Evaluation of Marine Wind Profiles in the North Sea and Norwegian Sea Based on Measurements and Satellite-Derived Wind Products. *Tellus A Dyn. Meteorol. Oceanogr.* **2022**, *74*, 1–16. [CrossRef]
17. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]
18. ECMWF. ERA5 Hourly Data on Single Levels from 1940 to Present. Available online: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form (accessed on 6 September 2023).
19. Solbrekke, I.M.; Sorteberg, A.; Haakenstad, H. The 3 km Norwegian reanalysis (NORA3)—A validation of offshore wind resources in the North Sea and the Norwegian Sea. *Wind Energy Sci.* **2021**, *6*, 1501–1519. [CrossRef]
20. MET.NO. NORA3. Available online: https://marine.met.no/node/19 (accessed on 18 September 2023).
21. ECMWF. Copernicus Arctic Regional Reanalysis. Available online: https://climate.copernicus.eu/copernicus-arctic-regional-reanalysis-service (accessed on 6 September 2023).
22. Theodoridis, S.; Koutroumbas, K. *Pattern Recognition*, 4th ed.; Academic Press, Inc.: Orlando, FL, USA, 2008.
23. Pavia, E.; O'Brien, J. Weibull Statistics of Wind Speed over the Ocean. *J. Appl. Meteorol.* **1986**, *25*, 1324–1332. [CrossRef]
24. Pryor, S.C.; Nielsen, M.; Barthelmie, R.J.; Mann, J. Can Satellite Sampling of Offshore Wind Speeds Realistically Represent Wind Speed Distributions? Part II: Quantifying Uncertainties Associated with Distribution Fitting Methods. *J. Appl. Meteorol.* **2004**, *43*, 739–750. [CrossRef]
25. Cook, N.J. The Deaves and Harris ABL model applied to heterogeneous terrain. *J. Wind Eng. Ind. Aerodyn.* **1997**, *66*, 197–214. [CrossRef]

3

Graph Neural Networks: Theoretical Background

3.1 Graph Theory

Graphs are mathematical structures used to model sets of objects where some are pair-wise connected or related. They provide a flexible framework for representing complex networks and interactions and are used to analyse a wide range of systems and processes in biological, chemical, social, technical, and physical sciences.

Definition: Graph A graph is a pair $G = (V, E)$, with the (generally finite) set V of *vertices* or *nodes* v_i and a set E of unordered pairs $\{v_i, v_j\}$ called *edges*.

If the edge $e_{ij} = \{v_i, v_j\}$ exists, the vertices v_i and v_j are called *adjacent*. The *neighbourhood* of a node v is defined as $\mathcal{N}(v) = \{u \in V | \{v, u\} \in E\}$. Let $n = |V|$ be the number of nodes in the graph. The *adjacency matrix* $\mathbf{A} \in \mathbb{R}^{n \times n}$ is defined by

$$\forall i, j \in \{1, \dots, n\} : \quad A_{ij} = \begin{cases} 1, & \{v_i, v_j\} \in E \\ 0, & \{v_i, v_j\} \notin E \end{cases} \tag{3.1}$$

A graph if *directed* if the set of edges contains ordered pairs instead of unordered pairs. A graph is undirected if and only if its adjacency matrix is symmetric.

A graph may have node *attributes* (or, in this context, *features*) $\mathbf{X} \in \mathbb{R}^{n \times f}$ with feature dimension f. In a *weighted* graph, a weight $w \in \mathbb{R}$ is assigned to each edge. Accordingly, a weighted graph can have edge attributes $\mathbf{X}^e \in \mathbb{R}^{m \times c}$, with m being the number of edges in the graph and c the number of edge attributes. This way, more than one weight can be assigned to each edge.

Definition: Bipartite graph A graph $G = (V, E)$ is *bipartite* if its set of vertices V can be divided into two disjoint and independent sets $V = U \cup W$. Two sets are disjoint if $U \cap W = \emptyset$; one set of vertices U is independent if no two elements within it are adjacent. This means that every edge of a bipartite graph $G = (U, W, E)$ connects one node in U to one in W.

3.2 Graph Neural Networks

Machine learning is the field of study concerned with developing algorithms that can learn to solve tasks from data and generalize this ability to previously unseen data. It includes a variety of models to classify, predict, cluster or recognize patterns in a given dataset. The most successful type of model from this field is the artificial *neural network*, loosely inspired by the function of neurons and synapses in human or animal brains. The "neurons" or *nodes* are typically arranged in layers, and layers are connected by "synapses" or edges, performing different transformations on the input data stored in the nodes. Each edge has a weight $w \in \mathbb{R}$ assigned to it that changes during the training process[1]. During training, the neural network processes input data through its multiple layers (one input layer, one output layer, and any number of "hidden" layers inbetween). In each node, a non-linear *activation function* is applied to the weighted sum of the outputs of the previous layer. This allows the network to capture complex, non-linear relationships in the data.
The network's performance is measured using a *loss function*, which quantifies the difference between the predicted output and the actual target, the *ground truth*. This means that in order to solve a task with a neural network, we need a training dataset for which the task has already been solved; it has to be *labeled*. This is the case for any *supervised* learning method. The choice of loss function depends on the task; for instance, mean squared error is usually used for regression tasks, while cross-entropy loss is used for classification tasks. To minimize the loss function, the network undergoes an optimization process called *backpropagation*. During backpropagation, the network calculates the gradient of the loss function with respect to each weight by applying the

1. We have already heard of many of these terms in the previous section on graph theory. In fact, any neural network can be described as a dynamic, directed, and weighted graph. This does not make it a graph neural network though.

chain rule backwards through the network—hence the name. These gradients indicate how much each weight should be adjusted to reduce the loss. Lastly, an *optimizer*, such as stochastic gradient descent or more advanced variants like Adam, updates the weights iteratively based on the calculated gradients. The optimization process continues until the network converges to a set of weights that minimize the loss function, resulting in a trained model. More detailed explanations of these concepts can be found in [1, 11, 44].

Deep learning, a subset of machine learning, refers to the use of sophisticated, hierarchical neural network architectures for complex task-solving. The term "deep" refers to the increased number of hidden layers in deep learning neural networks. The main difference between machine learning and deep learning models lies in their approach to data representation. While machine learning models typically require manual feature extraction, deep learning models automatically discover the representations needed for classification or detection through multiple layers of abstraction. Deep learning models are therefore particularly adept at handling large volumes of unstructured data, such as images, audio, or text.

Within deep learning, graph neural networks form the class of neural networks that processes graph-structured data. They operate by recursively aggregating and transforming information from each node's neighbourhood to compute new representations of all nodes. This process, known as *message passing*, allows GNNs to capture the dependencies and interactions between nodes in a graph. For a node feature vector $\mathbf{x}_i^{(k)} \in \mathbb{R}^f$ with f features in layer k of a network, and the edge attributes $\mathbf{e}_{j,i} \in \mathbb{R}^c$ from node v_j to node v_i ($\mathbf{e}_{j,i} = 1$ for unweighted graphs), a general message passing operation can be expressed as

$$\mathbf{x}_i^{(k)} = \gamma^{(k)} \left(\mathbf{x}_i^{(k-1)}, \bigoplus_{j \in \mathcal{N}(i)} \phi^{(k)} \left(\mathbf{x}_i^{(k-1)}, \mathbf{x}_j^{(k-1)}, \mathbf{e}_{j,i} \right) \right). \tag{3.2}$$

Here, $\bigoplus$ denotes the aggregation scheme, a differentiable, permutation invariant function, e.g. a sum, and γ and ϕ denote layer-specific, differentiable functions, e.g. linear transformations. Depending on the task at hand and the characteristics of a given graph, the optimal choices for those functions may differ, creating a variety of possible message passing schemes. Wu et al. [47] provide a comprehensive overview of GNNs, thoroughly explaining the different "flavours" of message passing.
GNNs have been proven to be highly effective for modeling complex physical systems and systems on non-Euclidean domains [35, 41]. Global weather data, which captures the state of the atmosphere, roughly conforms to the shape of a hollow sphere—a non-Euclidean space—and is governed by thermodynamic laws. Therefore, GNNs should be particularly well-suited for simulating weather systems.

Convolution is the core mathematical operation that underpins an entire category of neural networks known as Convolutional Neural Networks (CNNs). It is roughly defined as the integral of the product of two functions it is given as arguments[2]. In contrast to fully connected neural networks, convolutional filters aim to simulate the behaviour of a visual cortex, allowing for the recognition of spatial relationships in image data, for example. On a regional scale, the domain of weather data could be approximated as Euclidean and handled with conventional CNNs. According to Bruna et al. [4], however, employing graph-based versions of convolutional architectures can substantially lower the number of parameters in a network without degrading performance, often improving test error and accelerating forward propagation.
Convolutional GNNs generalize the convolution operation from regularly gridded data to graph data. Instead of passing a filter over the pixels of an image, the learned, localized, spectral filter is convolved with the node representations, aggregating the information of each node's neighbourhood. Applying graph convolutional layers iteratively increases the receptive field of the network by one further connection per layer. The connectivity of the graph will thus determine how quickly messages are "passed" within the graph. There are different variants of graph covolutional operators, two of which are specified hereafter.

Definition: GraphConv Operator The node-wise formulation of the GraphConv operator as defined by Morris et al. [29] is given by

$$\mathbf{x}_i^{(k)} = \mathbf{W}_1 \mathbf{x}_i^{(k-1)} + \mathbf{W}_2 \sum_{j \in \mathcal{N}(i)} e_{j,i} \cdot \mathbf{x}_j^{(k-1)}, \tag{3.3}$$

where $e_{j,i}$ denotes the edge weight from source node v_j to target node v_i. Comparing this expression to Eq. 3.2, a sum has been chosen as aggregation scheme, and simple linear transformations (without added bias) for the functions γ and ϕ. Note that, due to the separate transformation of the target node features $\mathbf{x}_i^{(k)}$ and the source node features $\mathbf{x}_j^{(k-1)}$, this definition allows for the handling of bipartite graphs where the nodes of the two sets U and W may have different feature dimensions.

Definition: GCNConv Operator The Graph Convolutional Network (GCN) layer as introduced by Kipf and Welling in [18] applied to the feature vector $\mathbf{x}_i$ of node v_i in layer k is mathematically defined as

$$\mathbf{x}_i^{(k)} = \sum_{j \in \mathcal{N}(i) \cup \{i\}} \frac{1}{\sqrt{\deg(i)}\sqrt{\deg(j)}} \cdot \left(\mathbf{W}^\top \cdot \mathbf{x}_j^{(k-1)} \right) + \mathbf{b}. \tag{3.4}$$

2. Technically, the exact mathematical definition requires one of the functions to be reflected about the y-axis and shifted before computing the product or the integral, but in the context of CNNs, this step is unnecessary and thus usually skipped.

The neighbouring node features including its own are transformed by a weight matrix $\mathbf{W}$, normalized by their degree, and summed up. Lastly, a bias vector $\mathbf{b}$ is added to the output. If the graph is weighted, this formulation generalizes to

$$\mathbf{x}_i^{(k)} = \sum_{j \in \mathcal{N}(i) \cup \{i\}} \frac{e_{j,i}}{\sqrt{\hat{d}_j}\sqrt{\hat{d}_i}} \cdot \left(\mathbf{W}^\top \cdot \mathbf{x}_j^{(k-1)}\right) + \mathbf{b}, \tag{3.5}$$

where $e_{j,i}$ denotes the edge weight from source node v_j to target node v_i ($e_{j,i} = 1$ for all edges in unweighted graphs) and $\hat{d}_i = 1 + \sum_{j \in \mathcal{N}(i)} e_{j,i}$, the weighted equivalent to the degree of node v_i, including the self-connection[3] [36]. This transformation can also be expressed as a spectral-based method:

$$\mathbf{X}^{(k)} = \hat{\mathbf{D}}^{-\frac{1}{2}}\hat{\mathbf{A}}\hat{\mathbf{D}}^{-\frac{1}{2}}\mathbf{X}^{(k-1)}\mathbf{W}^{(k-1)}, \tag{3.6}$$

where $\hat{\mathbf{A}} = \mathbf{A}+\mathbf{I}_n$ is the adjacency matrix with added self-loops ($\mathbf{I}_n$ is the identity matrix of dimension n), $\hat{\mathbf{D}}$ is a diagonal matrix of node degrees including the self-connections, so $\hat{D}_{ii} = \sum_j \hat{A}_{ij}$, and $\mathbf{W}^{(k-1)}$ is the trainable weight matrix of the previous layer [18]. This particular formulation ensures that the eigenvalues of the operator applied to $\mathbf{X}$ lie within the range $[0, 1]$, thereby avoiding numerical instabilities, i.e. vanishing or exploding gradients [47]. This modification of the adjacency matrix is called *renormalization trick*[4].

Note that the activation function has not been included in any of the above formulations and will have to be applied to their output before passing it on to the next layer. With this foundation laid, the next section will dive into a more detailed discussion of the possibilities and limitations of deep learning in meteorological research.

3.3 Possibilities and Limitations of Deep Learning in Meteorological Research

In recent years, machine and deep learning have emerged as powerful tools across various domains of research and everyday life, especially the latter often hailed for its transformative potential. Machine learning has first been applied in atmospheric science in 1990, performing cloud classification with neural

3. Note that both of these formulations for graph convolutional operators do not allow multidimensional edge features. The original formulation of the GCN layer does not include any edge features, but the implementation by PyTorch Geometric (see Section 4.2.1) allows for (scalar) edge weights for both GraphConv and GCN layer.
4. More recent works like [52] and [24] make further improvements to convolutional GNNs that could be explored in future works, taking long-range dependencies within graphs into account and introducing more flexibility.

networks [21]. While certainly applicable to some tasks, machine learning approaches are inherently limited by their use of domain-specific, sometimes hand-crafted features rather than a more generalized approach. In contrast, deep learning promises improved techniques for understanding temporal and spatial dependencies in data. Although deep learning undoubtedly offers potent opportunities, it is important to moderate the enthusiasm with a critical understanding of its boundaries. This section aims to briefly examine the possibilities, some prior applications, and limitations of deep learning within meteorological research, particularly wind prediction, to better appreciate where exactly deep learning might effectively contribute.

A fundamental aspect of learning theory is a model's ability to generalize from a finite set of training data. To logically infer rules applicable to an entire set, one would need information about every member of that set. Machine learning mitigates this by building on a probabilistic framework rather than definitive rules, aiming to find rules that are *likely* correct for most members of the set. However, weather systems are among the most complex physical phenomena on Earth, influenced by numerous variables. This means that available training data will likely never capture all possible scenarios, leading to models that may perform well on training and validation data but fail in real-world applications outside their training domain [40].
Another critical issue is the assumption that the training data are identically and independently distributed (i.i.d.). Many derivations of machine learning techniques rely on the i.i.d. assumption; without it, key statistical laws such as the law of large numbers and the central limit theorem no longer apply. This assumption may be violated in meteorological data, which typically exhibits spatial and temporal dependencies; possibly affecting parameter convergence, performance metrics, and optimization processes, increasing the risk of overfitting and reducing the model's generalization ability [11, 40].

Deep learning for meteorological problems faces the paradox of having both an overabundance and a scarcity of data. The vast amount of meteorological data collected by a wide variety of sensors, both terrestrial and extraterrestrial, as well as outputs from NWP models, is almost insurmountable. However, merely collecting data does not equate to improving our understanding of the system being measured—data alone does not constitute knowledge. Many deep learning methods are (semi-)supervised and rely on labeled data for training. In areas where we lack physical knowledge and could benefit from deep learning, the labeled datasets available are often small and require manual labeling, making them insufficient for the needs of many deep learning models that demand large amounts of training data. However, this task essentially equates to complex classification based on spatial and temporal clues, e.g. for detecting specific weather events in high-dimensional observational or model data. This is an area that deep learning can certainly contribute to, efficiently expanding

the small, existing labeled datasets [15, 26, 46].
The increase in available observational and model data calls for the sensible integration of data-informed approaches with traditional theory-driven research. This could be done by improving parameterizations or replacing physical sub-models with deep learning models, reducing biases where our physical understanding is incomplete or where resolved models are computationally expensive [10, 13]. Deep learning can also aid in spatial downscaling of wind fields with reduced computational demands, as demonstrated with CNNs and generative adversarial networks (GANs) [6, 25]. Generative models and recurrent neural networks have also been used for interval wind power prediction [34]. It should be noted that, while improving the predictive accuracy of NWP models is important, it does not necessarily enhance our scientific understanding of weather systems. Interpretability is a major weakness of deep neural networks due to their black box nature, and we are still far from recovering causal relationships from deep learning models trained on observational data.

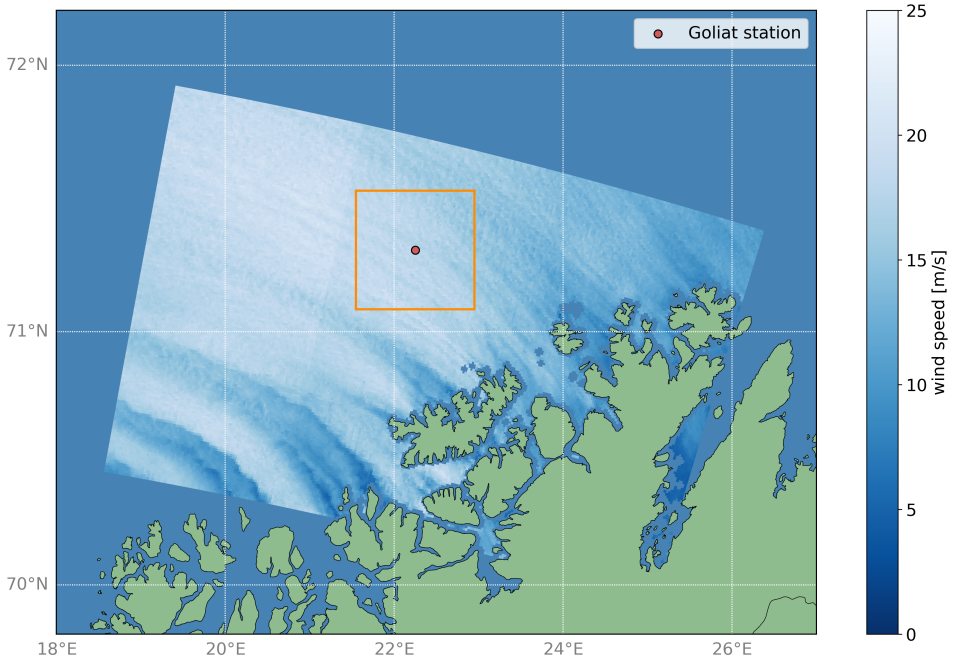
However, by learning from historical weather data, deep learning models offer a promising approach to improving nowcasting and forecasting abilities[5]. Especially GNNs have already proven to be more effective in this domain than traditional approaches, as seen in works by Ryan Keisler [17] and later by Google with GraphCast [20]. GraphCast claims to "significantly outperform the most accurate operational deterministic systems on 90% of 1380 verification targets, supporting better severe event prediction, including tropical cyclones, atmospheric rivers, and extreme temperatures". It is trained directly on reanalysis data, integrating historical weather observations and NWP model outputs. It leverages GNNs' ability to handle complex data structures, making it well-suited for capturing the spatial and temporal dynamics of atmospheric processes. The model predicts hundreds of weather variables over a 10-day period at a 0.25-degree resolution globally in under one minute. This promising result suggests that GNNs may also perform well on smaller-scale wind prediction.

5. It is worthy to note that, because weather is inherently chaotic and thus sensitively dependent on initial conditions, the time frame for which even a perfect weather model, physical or deep learning-based, can provide accurate forecasts is limited by the reliability and completeness of available observational data [32].

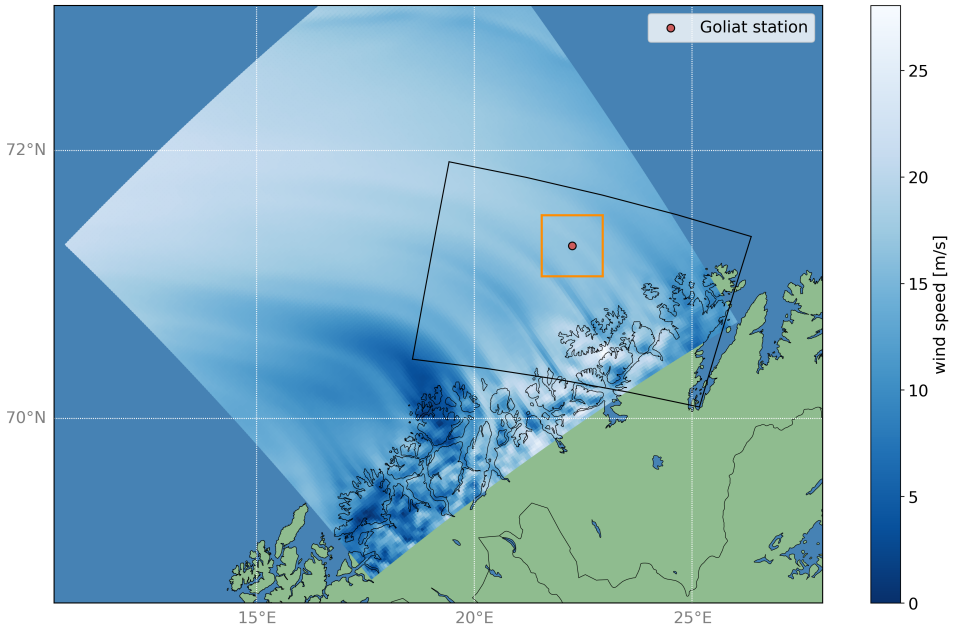
4

Graph Neural Networks: Experimental Approach

The last section has already highlighted the most prominent areas in which GNNs might benefit our cause of offshore wind prediction. Schultz et al. [42] concluded in their discussion that deep learning networks have particularly high chances of outperforming and replacing traditional approaches if they manage to leverage the small-scale patterns in observational data that NWP models fail to resolve. In our publication (see Chapter 2) we found that SAR-based wind retrieval can indeed compete with reanalysis products in terms of long-term wind resource representation while providing wind fields with a higher spatial resolution. Figure 4.1 shows two wind fields provided as provided by CARRA and Sentinel-1 OWI for (almost) the same point in time. It is apparent that not only is the spatial resolution of the SAR-based product much better, but there are also details and physical properties of the wind field discernible that are not or not well resolved by the regional reanalysis.
Previously proposed GNN models for weather predictions have used reanalysis datasets for training, which is convenient due to their spatial and temporal regularity, but it limits the models' learning potential. While these models can improve predictions by learning from historical data as opposed to NWP models, their performance is restricted by the data, which is informed by our imperfect physical understanding and numerical implementation of the weather system. There have been initial attempts at training deep learning networks with remote sensing data, such as precipitation nowcasting with radar data [38, 49],



(a) Sentinel-1 OWI wind field from January 3rd, 2022 at 5:00 UTC. Grid points over land are masked out as the retrieval method is only defined on open ocean ($\geq$ 1 km from shore). Note the high level of detail.



(b) CARRA wind field from January 3rd, 2022 at 6:00 UTC, the output time step closest in time to the SAR scene shown above. The outline of the corresponding Sentinel-1 SAR scene is shown in black.

Figure 4.1: Comparison of Sentinel-1 OWI and CARRA wind fields in our area of interest. The location of the Goliat station is indicated in both pictures (red dot) as well as the area of model prediction (orange frame, see Section 4.1.1). The differences in resolution and detail between the two data sources are clearly visible.

and studies employing deep learning networks, even GNNs, for a variety of tasks around the processing and classification of raw SAR data [27, 30]. Zhu et al. [51] provide an extensive review of existing studies and limitations of deep learning with SAR data. To our knowledge, no GNN or any other deep learning network has been trained using SAR data for wind applications.

The central problem presents as follows: CARRA has a high temporal resolution but too low spatial resolution; the SAR-based Sentinel-1 OWI component has a high spatial resolution and level of detail, but too low temporal resolution. Therefore, our goal is to fill these gaps by producing high-resolution and high-detail wind fields for times where CARRA output is available but no Sentinel-1 SAR scene. Our approach involves providing the GNN with CARRA data as input, one time step at a time, to predict wind speed values at the same spatial resolution as the Sentinel-1 OWI component ($1\,\text{km} \times 1\,\text{km}$). By leveraging the strengths of both data sources, our model aims to generate accurate and detailed wind fields, contributing to more reliable and complete offshore wind resource assessments in the Norwegian Arctic region.

4.1 Graph Setup

4.1.1 Datasets

Sentinel-1 OWI The Sentinel-1 constellation consists of two satellites (Sentinel-1A and Sentinel-1B), which together provide global coverage with relatively short revisit times. The European Space Agency (ESA) operates Sentinel-1 as an open data mission, which means that the acquired data are publicly available for non-commercial use and can be accessed through the Copernicus Data Space Ecosystem [8].
Sentinel-1 delivers radar imagery products distributed on three levels of processing. Level-0 consists of raw data. Level-1 provides baseline products for multiple sensing modes. Level-2 consists of geophysical products derived from Level-1. The comparatively new Ocean Wind Fields (OWI) component used in this thesis is one of three Level-2 Ocean (OCN) products for wind, wave and currents applications. Wind speed values in the range of $[0\,\text{m/s}, 25\,\text{m/s}]$ and wind direction values in degrees are provided alongside their respective grid points of given latitude and longitude at a spatial resolution of $1\,\text{km} \times 1\,\text{km}$. The wind direction values are informed by reanalysis data.
We collected all available Sentinel-1 OWI wind data overlapping with our area of interest—the Goliat station in the Barents Sea—in the period between February 2021 and December 2023, acquiring a total of 832 SAR scenes. The amount of data points varies from scene to scene depending on the area it covers, and lies anywhere between roughly $20,000$ and $45,000$.

CARRA The essential characteristics of the Arctic-specific, regional reanalysis CARRA are presented in Section 3.1.2 of the publication in Chapter 2. From the three reanalysis products investigated in the publication, CARRA has the highest spatial resolution of 2.5 km and a substantially higher temporal resolution than Sentinel-1 (3-hourly vs. 1-2 times in 2 days). That and its enhanced performance in the Arctic regions made it a solid choice for the purposes of this project.
We use output data from the CARRA reanalysis for the same period of time as Sentinel-1 OWI, February 2021 until December 2023, at its full temporal resolution (output every 3 hours), including the variables 10 m wind speed, 10 m wind direction, surface pressure, and 2 m air temperature. The downloaded data has a spatial resolution of 2.5 km $\times$ 2.5 km and each file covers the same area of roughly 400 km $\times$ 400 km including the Goliat station, meaning each file contains 160 $\times$ 160 grid points.

Ground truth data To enable the model to learn from observational rather than model output data, the ground truth data was constructed using the Sentinel-1 OWI component. Due to the satellites' varying orientation as they pass over our area of interest, the geographical locations of the grid points vary with each SAR scene. Since we are trying to fill the temporal gaps in-between SAR scenes, the model will be expected to predict a high-resolution wind field for a particular geographic area without a corresponding SAR scene for comparison. Consequently, the ground truth data was fitted to an *output grid* of 51 $\times$ 51 grid points around the Goliat station with fixed geographical locations, maintaining Sentinel-1 OWI's spatial resolution of 1 km $\times$ 1 km.
For each available SAR scene, a linear interpolation was performed using the four closest neighbours of each output node within the SAR grid, interpolating the given wind speed values. Before conducting the nearest-neighbour search, all parts of the SAR scene that were more than 2.5 km away from the outer edges of the output grid were masked out to reduce runtime. Finally, the wind speed values were normalized by dividing by 25 m/s, the maximum output value of the Sentinel-1 OWI component.

One might question, why not use a larger output grid? After all, 51 km $\times$ 51 km does not represent a very large offshore area; and the SAR scenes cover areas of up to around 250 km in width, with over 40,000 grid points instead of 2,601. For meaningful interpolation of a ground truth wind field, the Sentinel-1 scene must fully cover the output grid. However, as the coverage of Sentinel-1 scenes varies, increasing the grid size would result in fewer scenes overlapping entirely with the output grid, rendering them useless for training. From an initial collection of 832 Sentinel-1 SAR scenes between February 2021 and December 2023 that overlap with the Goliat station, only 548 scenes overlap with at least 49 out of the 51 rows or columns of the output grid. This constraint reduces the already small dataset available for model training by approximately

35%. Luckily, this issue will simply be solved over time as the amount of Sentinel-1 OWI data increases daily, gradually expanding the available dataset for training.

Lastly, although the wind direction is provided alongside the wind speed in the Sentinel-1 OWI component, we decided to only predict the spatially distributed wind speed. As of today, SAR-based wind retrieval methods are insufficiently capable of retrieving an accurate estimation of the wind direction and thus rely on wind direction information from NWP models. Despite the higher spatial resolution of the Sentinel-1 data, this causal relationship between the two makes the prediction of wind directions based on the SAR data superfluous.

4.1.2 Preprocessing

Given that meteorological variables like wind speed, temperature, and pressure are inherently positive, it seems appropriate to normalize them to the $[0, 1]$ range, as this best reflects their original values.
Within the CARRA files, all values for surface pressure and 2 m temperature were normalized using their global minimum and maximum values. All wind speed values were capped at 25 m/s, as Sentinel-1 OWI values cannot exceed that value anyway, and then wind speed was normalized in both Sentinel-1 and CARRA files by dividing by 25, resulting in wind speed values between 0 and 1. Although wind directions are of course circular in nature, this cannot be represented properly by any interval in $\mathbb{R}$. The wind direction values in the CARRA files were therefore normalized by dividing by 360, again resulting in values between 0 and 1.

4.1.3 Graph Structure

The model operates on a graph $G = (V^{\text{CARRA}}, V^{\text{out}}, E^{\text{CARRA}}, E^{\text{out}}, E^{\text{CARRA}\to\text{out}})$. The features and edge weights are time step dependent, meaning a graph $G = G(t)$ is constructed for each time step t. Graphs of different time steps are not connected. The following paragraphs describe the individual components in detail and Figure 4.2 illustrates the structure of a graph $G(t)$ schematically.

CARRA nodes CARRA nodes $v_i^{\text{CARRA}} \in V_{\text{CARRA}}$ represent the larger grid of $160 \times 160 = 25,600$ nodes as spanned by the CARRA reanalysis, with a spatial resolution of 2.5 km × 2.5 km. CARRA nodes contain the four features wind speed, wind direction, surface pressure, and 2 m air temperature, as provided by the reanalysis for a given time step t.

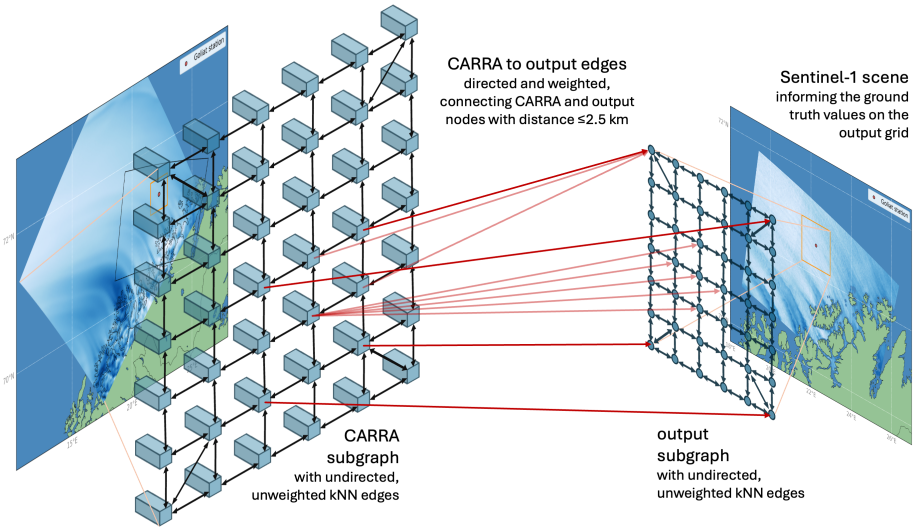


Figure 4.2: Illustration of the input graph structure. The two node types are shown on the left (CARRA) and right (output). The longer cuboids of the CARRA nodes represent feature vectors vs. the scalar output nodes shown as circles. The homogeneous edge types created via kNN search are schematically shown in black for both subgraphs. The heterogeneous edges going from CARRA to output nodes are schematically shown in orange. Note that one CARRA node may be connected to several output nodes, and vice versa, as illustrated by the light orange edges.

Output nodes Output nodes $v_i^{\text{out}} \in V^{\text{out}}$ represent the small grid of $51 \times 51 = 2,601$ points centered around the Goliat station, with a spatial resolution of $1\,\text{km} \times 1\,\text{km}$. They have no features associated with them, but for implementation purposes of the message passing have been initialized with one feature that is set constant.

In contrast to many other graphs commonly used in the context of GNNs, the graph nodes in this case have geophysical meaning and are thus tied to geographical locations. In addition to the node type and edge type dictionary items, the latitudes and longitudes of the individual subgraphs' nodes are stored as a separate dictionary item so as to not confuse them for features.

CARRA edges E^{CARRA} contains the undirected edges on the CARRA subgraph, i.e. edges that connect two CARRA nodes v_i^{CARRA} and v_j^{CARRA}. For each CARRA file, edges are created using a k-nearest-neighbour (kNN) search with $k = 4$ (or $k = 8$), resulting in $76,800$[1] CARRA edges for each time step. As the CARRA nodes are arranged in a regular grid, the four nearest neighbours of most nodes will be those directly "above, below, left, and right" of it, so the

1. This number is much smaller than $4 \cdot 25,600 = 102,400$ since in an undirected graph, the edges $\{v_i, v_j\}$ and $\{v_j, v_i\}$ are not distinct and thus only count for one edge.

distances spanned by most edges will be roughly the same; 2.5 km. The CARRA edges are therefore unweighted.

Output edges The edges on the output subgraph, E^{out}, share most of those characteristics. They are also undirected and unweighted, and each output node is connected to its four nearest neighbours, resulting in 10, 404 output edges.

CARRA-to-output edges Lastly, $E^{\text{CARRA}\to\text{out}}$ contains the directed, weighted edges pointing from source nodes within the CARRA subgraph to target nodes within the output subgraph. An edge is added if the distance between a CARRA node and an output node is less than or equal to 2.5 km. Due to the regular distances within both grids, this procedure ensures that each output node is adjacent to at least one CARRA node, allowing for sufficient message passing from one to the other subgraph. As the geographical distances spanned by this edge type may be anything between 0 and 2.5 km, edge weights are stored with them, containing the inversed, normalized distances. Depending on the time step, $E^{\text{CARRA}\to\text{out}}$ contains 8317 or 8310 edges[2]. On average, each output node is therefore connected to roughly 3.2 CARRA nodes.

4.2 Model Setup

4.2.1 Implementation

All code for this thesis project was written in Python. Several Python libraries are available that allow the implementation of sophisticated deep learning networks with only a few dozen lines of code; here, PyTorch was used [31]. PyTorch Geometric (PyG) provides a convenient framework for the implementation of GNNs [9]. Some components of the graphs and ground truth data were computed with scikit-learn [33]. All trainings were run on a MacBook Pro 2021 (M1), so the computational power available was very limited, presupposing efficient implementation and coding practices.

4.2.2 Architecture Overview

For forecasting purposes, GNNs are commonly built in an "encode-process-decode" configuration: first, the input graph G_{in} is transformed into a latent representation G_0 by an encoder, another transformation is then applied to

2. This difference is the result of a slight shift of the latitudes and longitudes of the CARRA grid points after June 30th, 2022.

G_0 M times to return G_M, which is then decoded into an output graph G_{out} [2, 17, 20].

Here, this approach is modified slightly as our interest is not to forecast wind fields but to downscale them from the coarser CARRA grid onto a different, finer grid. We therefore skip the "decode" part of the process and remain with a "(process)-encode-process" architecture. Depending on the experiment, one or more message passing layers first operate on the input graph $G_{\text{in}} = (V^{\text{CARRA}}, E^{\text{CARRA}})$. Then, a single message passing step over the CARRA-to-output bipartite subgraph $(V^{\text{CARRA}}, V^{\text{out}}, E^{\text{CARRA}\to\text{out}})$ is performed, encoding the (representation of the) CARRA input subgraph and creating the latent representation on the output subgraph, $G_0 = (V^{\text{out}}, E^{\text{out}})$. This representation is then processed by a number M of message passing layers, varying the number of channels, to produce the final output graph $G_M = G_{\text{out}}$, with one feature dimension representing the wind speed.

The message passing schemes chosen for this purpose are of convolutional nature. The operator `conv.GCNConv` from `torch_geometric.nn` [36] was used for neighbourhood aggregation on the homogeneous subgraphs G_{CARRA} and G_{out}. Although it provides more numerical stability (see Section 3.2) it does not support the handling of bipartite graphs[3], rendering it unsuitable for the message passing from the CARRA to the output subgraph. Therefore, the `conv.GraphConv` layer from `torch_geometric.nn` [37] was used as an encoder. Lastly, a linear transformation (`torch_geometric.nn.Linear`) was applied to reduce the number of features to one for the final output. After every layer except the last one, the rectified linear unit (ReLU) activation function was applied to introduce non-linearity to the function modeled by the network and to ensure the output values of hidden layers are between 0 and 1.

4.2.3 Training Details

Regardless of their exact architecture, the training objective of all GNNs is defined as the mean squared error (MSE) between the ground truth values and the predicted values for the normalized wind speed on the output grid:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{2601} \sum_{i=1}^{2601} (y_i - \hat{y}_i)^2, \tag{4.1}$$

3. Looking at Eq. 3.6 we see that, if the adjacency matrix is no longer square (due to different numbers of nodes in V^{CARRA} and V^{out}), this expression is no longer valid, as the square root and inversion of matrices is only defined for square matrices. Furthermore, self-loops cannot be added when the sets of source and target nodes are not the same.

where y_i and $\hat{y}_i$ are the ground truth and the predicted wind speed value for node v_i, respectively. The model was then trained to minimize this loss using the Adam optimizer for backpropagation of the gradients and update the network weights. Adam was chosen for its adaptive learning rate capabilities and computational efficiency.
Prior to training, the weights were initialized using the He Kaiming method, due to its compatibility with ReLU activation functions. The learning rate was set between 0.001 and 0.0002, depending on the specific experiment. This range was chosen to balance the convergence speed and stability of the training process. The dataset was randomly split into training, validation, and test sets with a ratio of 80%/10%/10%. Random splitting was used to prevent potential temporal biases from using consecutive time steps and to ensure a fair evaluation of the models' performances across different data subsets.

5

Results and Discussion

This chapter reports on the results of the experiments with GNNs. The results from the publication on the comparison of Sentinel-1 OWI versus in situ observations and reanalyses can be found in Chapter 2 and will not be repeated here. There is no strict separation between results and discussion in this chapter as they go hand in hand in this rather conceptual work.

Training a neural network is mathematically equivalent to a non-convex optimization problem. The objective of the network is to minimize the loss, which depends on a large number of parameters. One could imagine the training process as a marble moving on the *loss surface* in the high-dimensional parameter space, pushed down-slope by the optimizer; its trajectory often navigating around local maxima and minima while aiming to find the global minimum. Whether a neural network can be trained for a given task highly depends, among other factors, on a variety of architecture design choices. The number and order of layers and channels, choice of activation function, loss function, optimizer, weight initialization method, and learning rate can substantially impact the structure of the loss surface and the learning process. As a neural network grows deeper or wider and thus, more parameters are added to the model, the loss surface generally becomes more complex and "sharp", increasing the chances of getting trapped in a local minimum. However, the exact nature of how the above-mentioned design choices affect the training performance are unfortunately unclear [22].
In that sense, finetuning a deep learning network for optimal performance on a given task becomes an empirical science, despite the mathematical nature

of the network. If the computational resources allow for it, the easiest way of finding the best possible choice for a design parameter may just be through trial and error, which is what was considered best in this context of a preliminary methodology evaluation.

Table 5.1 and Table 5.2 provide an overview of some of the tested architecture and training configurations, respectively. Some factors are not mentioned in the table because they were identical for all configurations: the MSE loss function, ReLU activation function, the Kaiming weight initialization method, the optimizer Adam, a batch size of 1 (necessary due to differences in graph dimensions), and the single GraphConv layer from CARRA to the output subgraph. In the following, we will evaluate pairs of network configurations by means of their loss curves, i.e. the development of training and validation losses over training epochs, and look at some examples of their respective predictions on the test data.
More network configurations were trained in the process of this project but were not deemed worth mentioning in this thesis, as their behaviour was as expected. The four configurations discussed here were selected for explicit comparison of individual parameters which will be discussed in the following.

version	**number of GCN layers**		**number of neighbours**	
	on CARRA	on output	CARRA	output
V1	3	6	8	8
V2	3	6	4	4
V3	4	2	4	4
V4	2	3	4	4

Table 5.1: Architecture design parameters for the presented selection of network configurations. "Number of neighbours" corresponds to the chosen values for k in the kNN graph constructions. The highlighted version **V4** is investigated in more depth in this chapter.

version	**hidden channels**	**learning rate**	**epochs**
V1	64	0.001	100
V2	64	0.0002	500
V3	64	0.0005	500
V4	64	0.0005	500

Table 5.2: Training parameters for the presented selection of network configurations. Other parameter choices were identical for all versions; loss function: MSE, activation function: ReLU, weight initialization: Kaiming method, optimizer: Adam, batch size = 1. The highlighted version **V4** is investigated in more depth in this chapter.

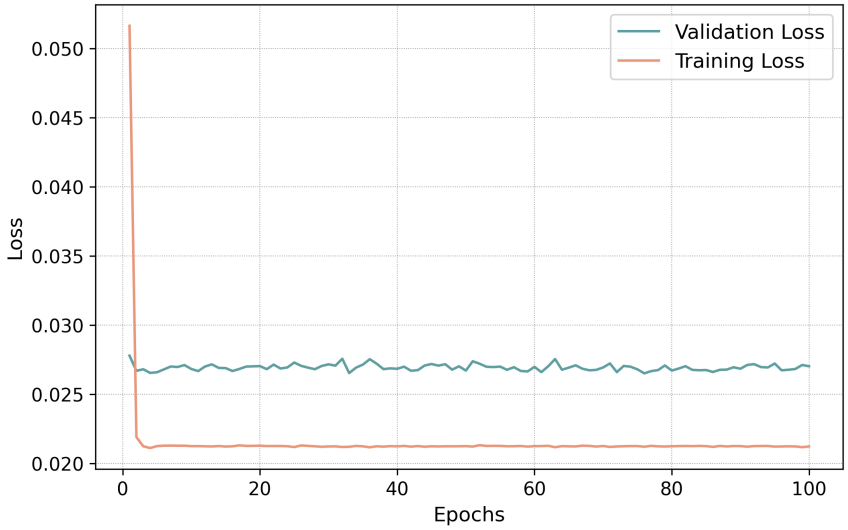
5.1 Over-Smoothing and Connectivity

A major difference between training performances and generalization ability (on the validation set) can be seen when the connectivity or the number of GCNConv layers operating on the CARRA and the output subgraphs changes. Figures 5.1 and 5.2 show the loss curves of the four selected network trainings. The most obvious deficit can be seen in Figure 5.1a, the loss curve of network **V1** which was trained on the 8NN-subgraphs[1] of CARRA and output as compared to the 4NN-subgraphs that were fed to the other three networks. This higher level of connectivity resulted in a degeneration of learning capability; although the loss decreased in the first epoch, it almost immediately stagnated around a much higher value than the other configurations trained on 4NN-subgraphs. The second difference between network **V1** and the others is the higher initial learning rate lr = 0.001, but the same value was used for the preliminary assessment of the learning behaviour of all network configurations and worked fine in all other cases. It is therefore unlikely that the failed training is a result of the initial learning rate being too high.

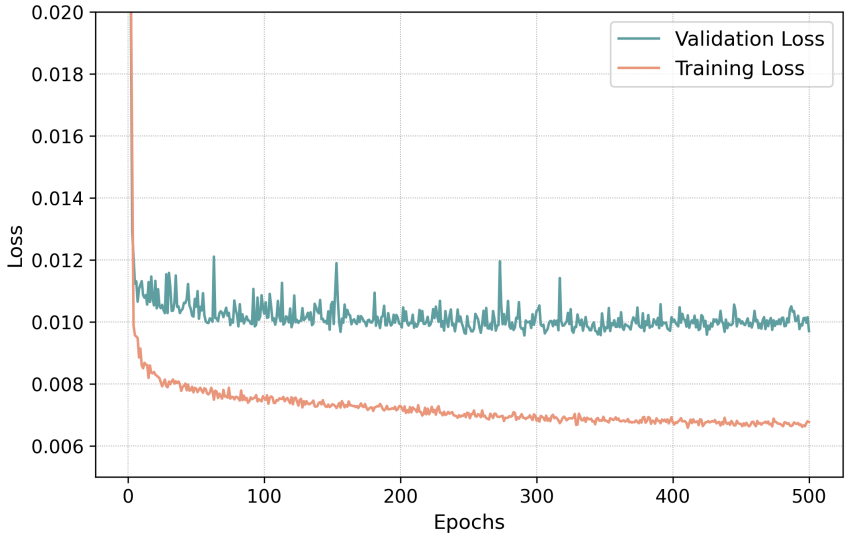
The remaining three configurations differed mainly in their number of GCNConv layers: network **V2** has 3/6 layers on the CARRA and output subgraphs, respectively, network **V3** has 4/2, and network **V4** has 2/3. All other parameters except a small change in the initial learning rate of network **V2** are identical. All three graphs clearly display learning of the networks, since both training and validation loss decrease for all configurations. The development of the training loss is indeed very similar for all networks, dropping dramatically after the first epoch and continuing to decrease with an increasingly shallow slope down to a value of $0.0065 - 0.0068$ after 500 epochs. The validation loss on the other hand displays distinctly different behaviour: For configuration **V2**, it decreases only until epoch 60 and then remains relatively constant, at a much higher loss of almost 0.010. For configuration **V3**, it decreases, but with a more shallow slope, and finishes after training around 0.008. For configuration **V4**, the validation loss decreases together with the training loss, and stagnates only after around 300 epochs. Its final loss value is roughly 0.007 and thus much closer to the final training loss, indicating better generalization. Table 5.3 summarizes the post-training losses of the four configurations.

This is no surprise, as the over-smoothing tendency of deep GNNs (so, GNNs with more stacked layers) is a well-known problem in literature and can be hard to overcome [5, 23, 50]. Although the configuration with fewer layers seems to perform better here, too, we should keep in mind that the number of layers also limits how far information can "travel" within the graph while being fed forward in the network; i.e. for two GCNConv layers on the CARRA 4NN-

1. 8NN (and 4NN) is short for kNN (k nearest neighbours) with $k = 8$ (or $k = 4$).

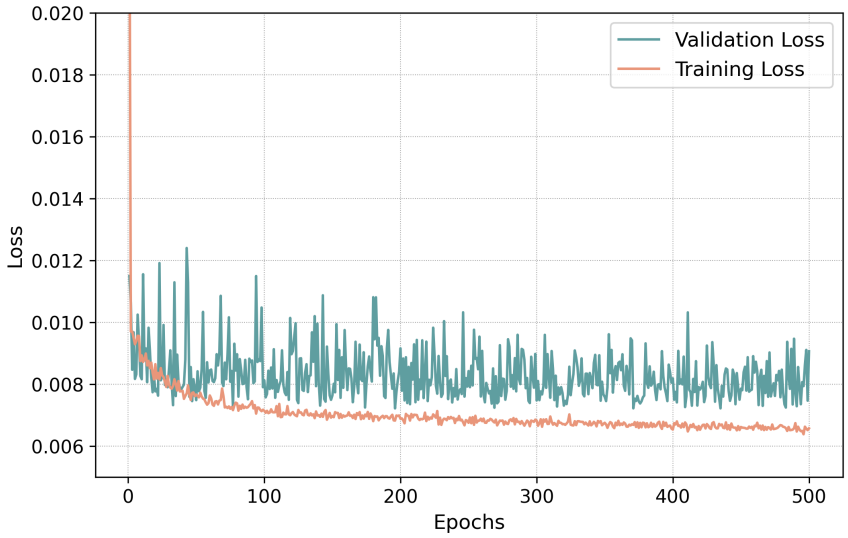


(a) Training and validation loss curves of network configuration **V1**, with 3 GCNConv layers on the CARRA and 6 GCNConv layers on the output subgraph, operating on 8NN-subgraphs instead of 4NN. Note the different scale of the vertical axis in this graph compared to the others. Since the learning did not progress after the first epochs, the training was aborted after 100 instead of 500 epochs.

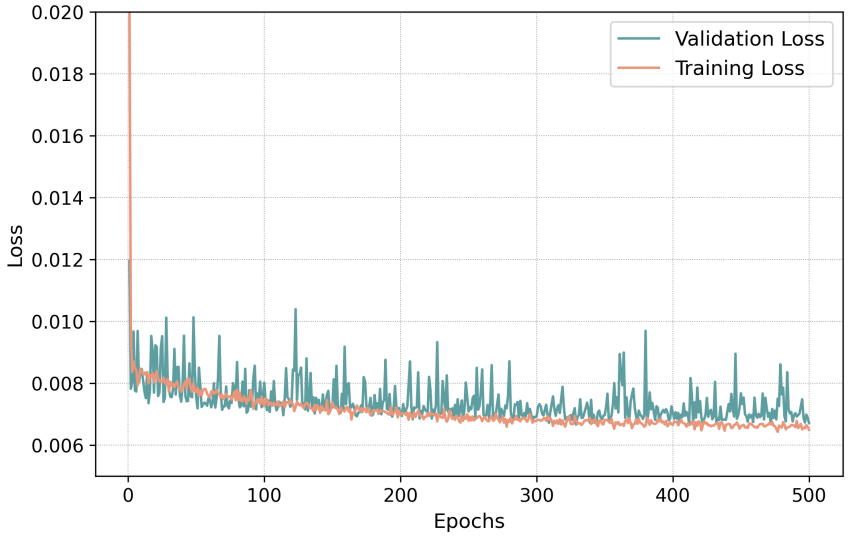


(b) Training and validation loss curves of network configuration **V2**, with 3 GCNConv layers on the CARRA and 6 GCNConv layers on the output subgraph.

Figure 5.1: Comparison of loss curves for the network configurations **V1** and **V2**. The main difference was in the graph setup; **V1** operated on 8NN-subgraphs of CARRA and output, **V2** on 4NN-subgraphs.



(a) Training and validation loss curves of network configuration **V3**, with 4 GCNConv layers on the CARRA and 2 GCNConv layers on the output subgraph.



(b) Training and validation loss curves of network configuration **V4**, with 2 GCNConv layers on the CARRA and 3 GCNConv layers on the output subgraph.

Figure 5.2: Comparison of loss curves for the network configurations **V3** and **V4** that differed in the number of GCNConv layers applied to the CARRA and output subgraphs. Although both graphs display learning, **V4** shows better learning and generalization behaviour.

version	train loss	val loss
V1	0.02124	0.02704
V2	0.00680	0.00970
V3	0.00658	0.00812
V4	0.00650	0.00672

Table 5.3: Training and validation losses of the four network configurations after training. The best-performing network **V4** is highlighted. The average MSE loss of a linear interpolation of the CARRA values is 0.01322.

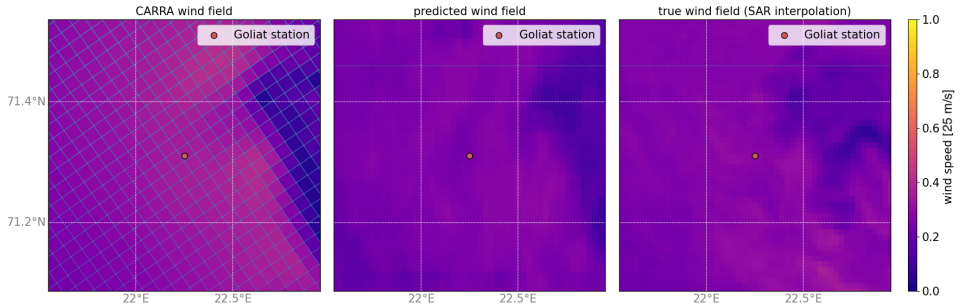
subgraph, one GraphConv layer from CARRA to output and three GCNConv layers on the output 4NN-subgraph, information from one CARRA node can be passed a maximum distance of $(2.5+2.5+2.5+1+1+1)$ km $= 10.5$ km across the geographical area covered by the graph nodes[2]. In this case, no CARRA node that is further than 3×2.5 km away from any output node will contribute at all to the final prediction, and the prediction from medium- to large-scale features of the wind field (relative to the size of the output grid) may not be possible unless they are already present in the CARRA wind field.

This illustrates that a bigger number of layers, or the ability to propagate information over larger areas within the graph, may be beneficial to our use case. A number of methods have been proposed to mitigate the problem of over-smoothing in deep GNNs or overcome the localized nature of shallow GNNs, e.g. co-training and self-training [23], topology-based regularization [5], or skip connections [50]. In this context, adding skip connections may allow for more effective networks and should be considered for future work. Nevertheless, the average MSE between linearly interpolated CARRA wind fields and the true wind fields is 0.01322, which compared to the validation losses in Table 5.3 shows that some of the GNN configurations evaluated here improve the small-scale representation, albeit not exhaustively.

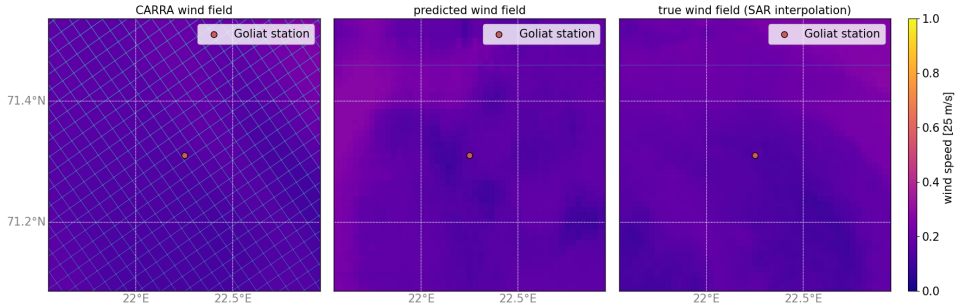
5.2 Spatial Variability

For better visualization of the problem, Figure 5.3 shows four predictions of network **V4** based on previously unseen data sampled from the test set. Due to the slightly better performance of this network compared to the others, the following results will be shown for this configuration only to stay concise.

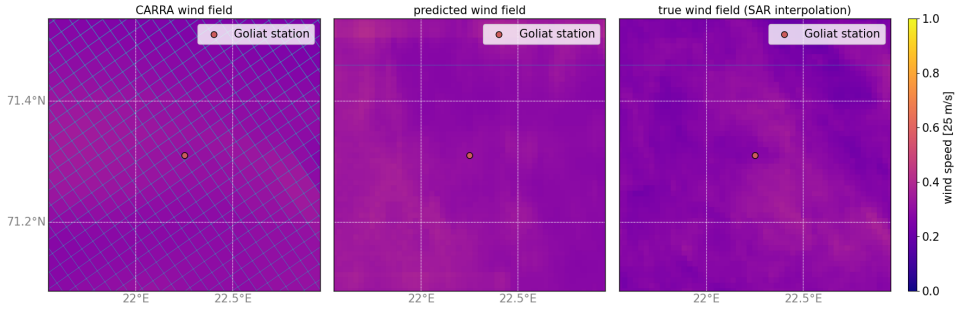
2. This is only true for *most* nodes: In the case of an output node that is part of the outer edge of the output grid and thus connected to "diagonal" neighbours, up to one connection within the output grid can span 2 km and up to two can span $\sqrt{2}$ km.



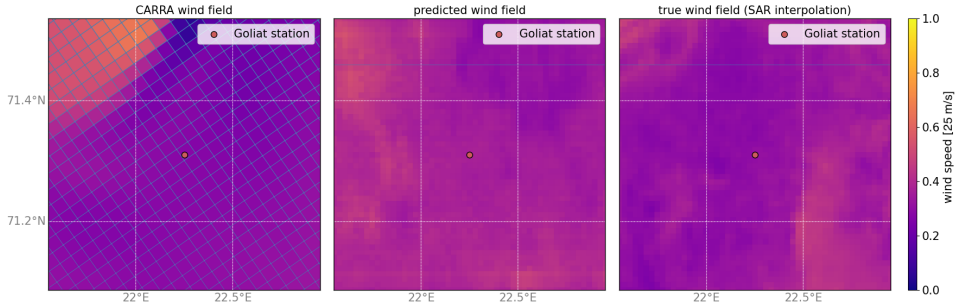
(a) May 1st, 2021 at 16:00 UTC



(b) September 16th, 2021 at 5:00 UTC



(c) October 23rd, 2021 at 16:00 UTC



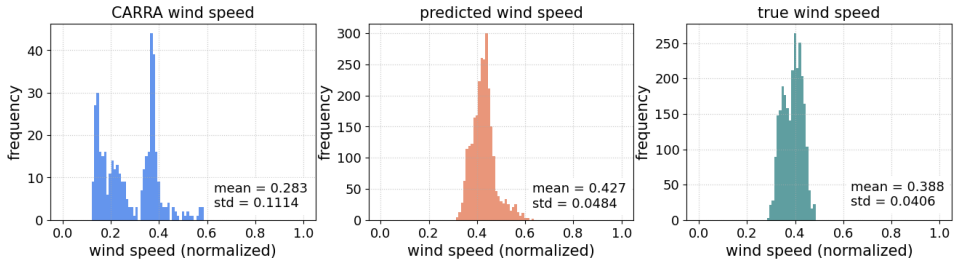
(d) November 11th, 2021 at 5:00 UTC

Figure 5.3: Wind fields drawn from CARRA (left), as predicted by the model **V4** (center), and the true wind field interpolated from Sentinel-1 SAR data (right), for a subset from the test data. All tiles display the same geographical area, 51 km × 51 km around the Goliat station. One value in CARRA corresponds to an area of 2.5 km × 2.5 km, whereas one value in the prediction or the truth corresponds to 1 km × 1 km, hence the difference in resolution. Note that the normalized wind speed is displayed, 1 on the colourmap corresponds to 25 m/s.

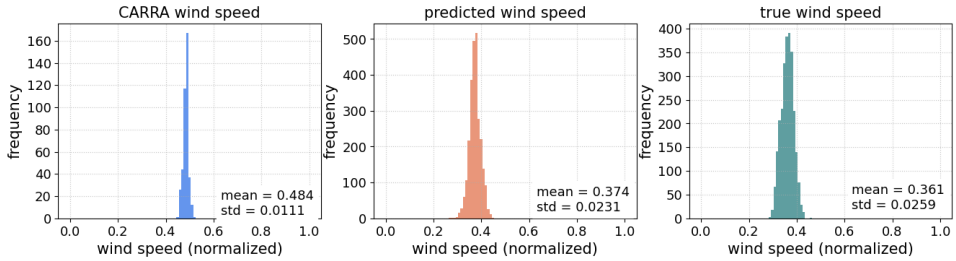
On the left of each subfigure, the CARRA input wind field is shown, the center tile shows the wind speed prediction of the model on the finer output grid and on the right, the ground truth wind field on the output grid as interpolated from the Sentinel-1 data is shown.
Independently from the quality of the prediction, the increased spatial resolution drastically improves the rendition of the wind field. The CARRA wind field tends to be quite smooth and often almost constant in this area. Only two out of the four shown CARRA scenes display any discernible features (top and bottom). The interpolations of the SAR scenes, on the other hand, although containing a similar range of wind speeds per time step as the CARRA scenes, display much more spatial variability and features in the wind field that sometimes only span a couple of "pixels" or, in this context, nodes. The predictions by the GNN configuration **V4** fail to replicate the precise locations and arrangements of those structures. But, at least to the human eye, they clearly show much more similarities to the ground truth data than a simple interpolation of the CARRA wind field would, for example.

The improved spatial variability and presence of small-scale details is exactly what we aim to reproduce with the GNN, but quantifying their presence is not straightforward. The easiest way to at least compare the range of wind speeds present in individual predictions is to compute the mean and standard deviation. Figure 5.4 shows examples of the wind speed histograms of CARRA, the model prediction, and the ground truth, for four time steps sampled from the test data. The plots only accumulate the wind speeds within the geographical area of the output grid. Due to its lower spatial resolution, the total number of points in the CARRA data is therefore lower than those of prediction and ground truth. The normalized values for mean wind speed and standard deviation are shown individually for each plot.
Although this is just a subset of the data and not representative, the figures indicate that in these cases, the standard deviations of the predicted wind speeds is much closer to the true value than to the CARRA value. If there are distinct differences in the appearance of the wind speed distributions of CARRA and ground truth, like in Figures 5.4a and 5.4b, the predicted distributions have more similarity in shape to those of the true distributions vs. those of CARRA.

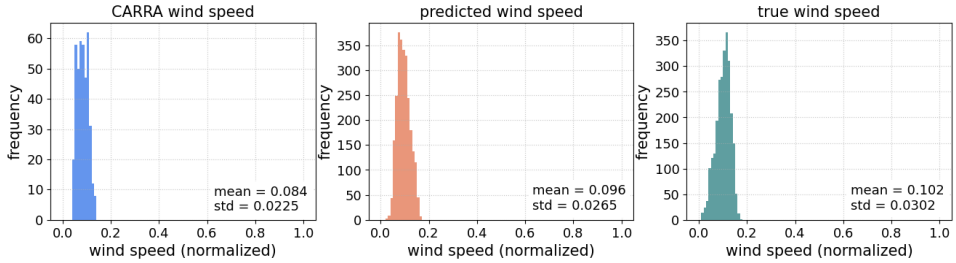
As discussed in more depth in the publication in Chapter 2, distributions representing long time series of wind speed measurements tend to have distinct shapes and can therefore readily be quantified and compared. Unfortunately, this is not the case here. As any of these scenes only capture a short moment in time and a very limited area of the Earth's surface, their wind speed histograms cannot be considered representative of a probabilistic distribution. The calculation of Weibull parameters would therefore be meaningless, and the values for mean and standard deviation have to be interpreted carefully.



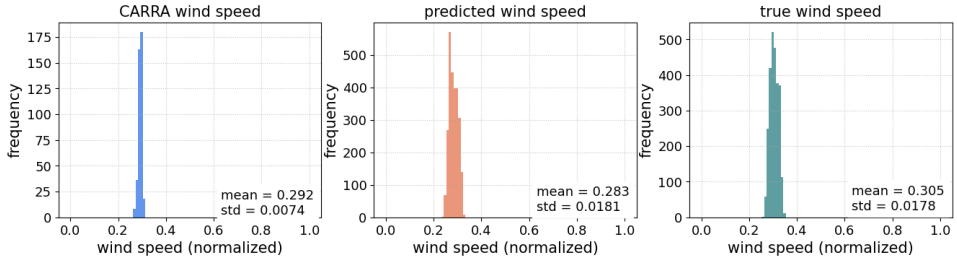
(a) March 28th, 2021 at 5:00 UTC



(b) December 5th, 2021 at 5:00 UTC

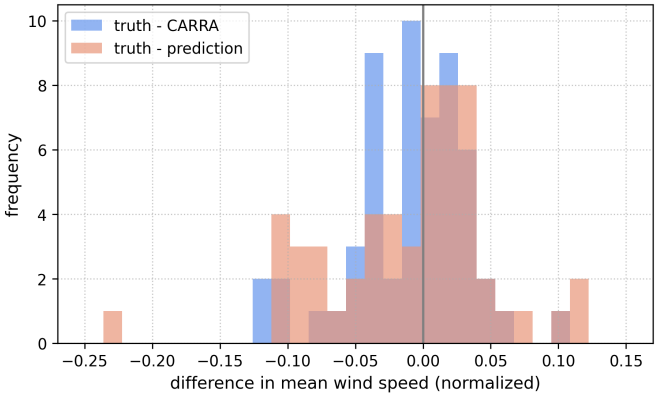


(c) August 25th, 2022 at 16:00 UTC

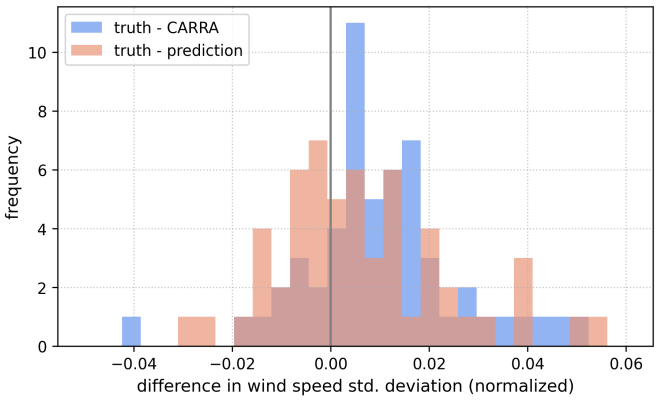


(d) November 30th, 2022 at 5:00 UTC

Figure 5.4: Histograms of the wind speeds of individual time steps within the geographical area of the output grid. From left to right: the wind speeds from the corresponding CARRA time step, the wind speeds as predicted from network **V4**, and the wind speeds as present in the Sentinel-1 interpolation. Note that CARRA's spatial resolution and therefore its total number of points is lower than that of model prediction and ground truth.



(a) The differences in mean wind speed between the true wind field and the CARRA wind field (blue) and the predicted wind field (orange) for individual time steps. Computed on the test data with network **V4**.



(b) The differences in wind speed standard deviation present in the true wind field minus CARRA wind field (blue) and minus the predicted wind field (orange) for individual time steps. Computed on the test data with network **V4**.

Figure 5.5: Histograms comparing the mean wind speed and standard deviation of CARRA and the model prediction to those of the ground truth. Computed for all time steps of the test data, on the area of the output grid.

Figure 5.4 shows the histograms of the differences in mean wind speed (top) and standard deviation (bottom) between the truth and CARRA (blue), and between truth and prediction (orange), for all data from the test set. We should note that, containing only 56 (approx. 10% of 548) samples, this set is still not representative. Nevertheless, we can observe that both histograms are in the same range of values, indicating that on average, the GNN configuration **V4** may not lead to an improvement. In Figure 5.5, the differences in standard deviation between truth and CARRA seem to have their maximum on the positive side. The differences in standard deviation between truth and prediction on the other hand seem to be more centered, which would indicate that the widths of the predicted wind speed distributions do not have a clear tendency to be smaller than the ground truth. However, this conclusion cannot confidently be drawn from this sample; more data would be needed.

5.3 Data Availability and Generalization Issue

When neither the dataset nor the network design have been previously established, it is difficult to infer whether the failure of network training is due to the dataset, the network design, or the task itself. By changing one network parameter at a time and comparing the models' relative performances, we can draw some conclusions about which parameter choice might be better. But if the dataset available for training is simply not representative of the collection of possible states, generalization becomes logically impossible (see Section 3.3). That is almost certainly the case here: A total of 548 CARRA and SAR scenes cannot possibly capture the whole picture of how any weather situation will influence the wind field.

Consequently, none of these results give a clear answer to the most essential question of whether our problem statement is well-defined in the first place. Does a preferably invertible function exist that maps the information contained in our (single-level, four meteorological variables) CARRA data to the corresponding high-resolution SAR wind field? If there is no such function and the problem is ill-defined, then no network finetuning or data accumulation will help the case.

6 Conclusions

The following paragraphs will conclude the findings presented in this thesis with regard to the two project phases and their respective research questions as stated in Chapter 1, followed by a brief evaluation and outlook.

6.1 Data Exploration

SAR-based offshore wind retrieval methods demonstrated a high spatial resolution and level of detail that may prove particularly advantageous for offshore wind resource assessment. The results from the publication indicated that SAR data can indeed compete with reanalysis products and in situ observations over the long term in the representation of wind resources. However, the temporal resolution of SAR data can remain a limitation e.g. in areas with pronounced diurnal variability, as it does not provide continuous coverage, which is crucial for comprehensive and accurate wind resource assessments.

The currently available SAR wind data, while providing high spatial resolution, was found to be insufficiently comprehensive as a standalone training source for GNNs. As an emerging technique for wind retrieval, products like the Sentinel-1 OWI component are only available for the last few years, aggravating its limited temporal coverage. But the datasets are growing, and it is therefore still worth taking the first steps now down the avenue of SAR-data-informed deep learning for wind resource evaluation.

Translating SAR wind data into a graph structure that facilitates physically consistent message passing in a GNN is a challenging task. Reducing a three-dimensional, dynamic fluid system to a two-dimensional, small-area, discrete grid inevitably results in some loss of information. While the graph structure used here enabled localized message passing, it did not incorporate long-range connections.

6.2 Graph Neural Networks

While GNNs are theoretically well-suited for modeling complex systems like weather on non-Euclidean domains, their practical application is not straightforward. The results suggest that simple GNN architectures may not be sufficient for this task. More sophisticated architectures might be necessary to capture both spatial and temporal dependencies effectively.

The experiments with simple GNN architectures showed that while they can generate reasonable two-dimensional near-surface wind fields, the level of detail did not consistently match that of SAR observations. The models struggled with accurately capturing the fine-scale variations present in the SAR data, indicating that more sophisticated architectures or additional data inputs are necessary to achieve higher accuracy.

It was demonstrated that network parameters and architecture choices significantly impact the performance and accuracy of GNNs for wind field prediction. Variations in the number of layers, along with the setup of the graph structure, influenced the model's ability to learn from the data and generalize to unseen conditions. These results underscore the importance of careful design and tuning of GNN configurations to optimize their performance for this specific application.

6.3 Epilogue

This research marks an early step in the application of GNNs for offshore wind prediction. As part of an interdisciplinary project in its initial stage—aiming to combine renewable energy and boundary layer meteorology with deep learning—, significant time was dedicated to literature research and discussion in both disciplines to develop comprehensive understanding of the possibilities of this integration. While this foundational work provided valuable insights and a strong knowledge base, it naturally limited the scope for extensive original research.

Nevertheless, this work addressed the research questions and made preliminary conclusions, which was the primary aim. Moving forward, the project group could benefit from the findings presented here by exploring any of the following directions:

- Considerations toward the inclusion of long-range spatial dependencies within the input graph without risking over-smoothing, by modifying the graph setup and refining the GNN architecture design. Existing research proposes a variety of methods for this purpose, e.g. by implementing skip connections [50] or an improved type of GCN layer [52].

- Improved GNN performance evaluation tailored to the characteristics of our requirements. Especially the reproduction of physical details present in the SAR data should be evaluated in more detail, i.e. through spatial statistics, and the suitability of the chosen loss function for this task should be determined.

- Expanding the training dataset, e.g. by learning wind fields for more than one location, thus having access to more SAR scenes, or by incorporating additional data, either observations for ground truth or adding more variables or levels to the CARRA input data.

Lastly, the fundamental purpose of this work as a Master's thesis was to determine the depth of my understanding, knowledge, and competence in my field of studies. In that sense, this project proved invaluable to me as a young researcher. I had the chance to conduct both collaborative and independent research, contribute to original findings, and navigate the ups and downs of peer-reviewed publishing. This experience has not only strengthened my expertise but also prepared me for future challenges and opportunities in the field of atmospheric sciences and beyond.

Bibliography

[1] Ethem Alpaydin. *Introduction to Machine Learning – Third Edition*. The MIT Press, Cambridge, Massachusetts, 2014.

[2] Peter W. Battaglia, Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinícius Flores Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, Çaglar Gülçehre, H. Francis Song, Andrew J. Ballard, Justin Gilmer, George E. Dahl, Ashish Vaswani, Kelsey R. Allen, Charles Nash, Victoria Langston, ..., and Razvan Pascanu. Relational inductive biases, deep learning, and graph networks. *CoRR*, abs/1806.01261, 2018. arXiv:1806.01261.

[3] Jonathan Bosch, Iain Staffell, and Adam D. Hawkes. Temporally explicit and spatially resolved global offshore wind energy potentials. *Energy*, 163:766–781, 2018. doi:10.1016/j.energy.2018.08.153.

[4] Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann Lecun. Spectral networks and locally connected networks on graphs. In *International Conference on Learning Representations (ICLR2014), CBLS, April 2014*, 2014. arXiv:1312.6203.

[5] Deli Chen, Yankai Lin, Wei Li, Peng Li, Jie Zhou, and Xu Sun. Measuring and Relieving the Over-smoothing Problem for Graph Neural Networks from the Topological View, 2019. arXiv:1909.03211.

[6] Jérôme Dujardin and Michael Lehning. Wind-Topo: Downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning. *Quarterly Journal of the Royal Meteorological Society*, 148(744):1368–1388, 2022. doi:10.1002/qj.4265.

[7] Charles Elachi and Jakob van Zyl. *Introduction to the Physics and Techniques of Remote Sensing*, chapter 7, Ocean Surface Sensing, pages 334–376. John Wiley Sons, Ltd, 2021. doi:10.1002/9781119523048.ch7.

[8] ESA. The Copernicus Data Space Ecosystem. https://dataspace.

copernicus.eu. [Online; accessed 26.01.2024].

[9] Matthias Fey and Jan E. Lenssen. Fast graph representation learning with PyTorch Geometric. In *ICLR 2019 Workshop on Representation Learning on Graphs and Manifolds*, 2019. URL: https://arxiv.org/abs/1903.02428.

[10] P. Gentine, M. Pritchard, S. Rasp, G. Reinaudi, and G. Yacalis. Could Machine Learning Break the Convection Parameterization Deadlock? *Geophysical Research Letters*, 45(11):5742–5751, 2018. doi:10.1029/2018GL078202.

[11] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[12] Giovanni Gualtieri. Reliability of ERA5 Reanalysis Data for Wind Resource Assessment: A Comparison against Tall Towers. *Energies*, 14(14), 2021. doi:10.3390/en14144169.

[13] C. Hallgren, J. A. Aird, S. Ivanell, H. Körnich, V. Vakkari, R. J. Barthelmie, S. C. Pryor, and E. Sahlée. Machine learning methods to improve spatial predictions of coastal wind speed profiles and low-level jets using single-level ERA5 data. *Wind Energy Science*, 9(4):821–840, 2024. doi:10.5194/wes-9-821-2024.

[14] International Energy Agency. *World Energy Outlook 2019*. 2019. doi:10.1787/caf32f3b-en.

[15] Valérian Jacques-Dumas, Francesco Ragone, Pierre Borgnat, Patrice Abry, and Freddy Bouchet. Deep Learning-Based Extreme Heatwave Forecast. *Frontiers in Climate*, 4, 2022. doi:10.3389/fclim.2022.789641.

[16] P. C. Kalverla, J. B. Duncan Jr., G.-J. Steeneveld, and A. A. M. Holtslag. Low-level jets over the North Sea based on ERA5 and observations: together they do better. *Wind Energy Science*, 4(2):193–209, 2019. doi:10.5194/wes-4-193-2019.

[17] Ryan Keisler. Forecasting Global Weather with Graph Neural Networks, 2022. arXiv:2202.07575.

[18] Thomas N. Kipf and Max Welling. Semi-Supervised Classification with Graph Convolutional Networks. *arXiv e-prints*, 2016. doi:10.48550/arXiv.1609.02907.

[19] Morten Køltzow, Harald Schyberg, Eivind Støylen, and Xiaohua Yang.

Value of the Copernicus Arctic Regional Reanalysis (CARRA) in representing near-surface temperature and wind speed in the north-east European Arctic. *Polar Research*, 41, March 2022. doi:10.33265/polar.v41.8002.

[20] Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, Ferran Alet, Suman Ravuri, Timo Ewalds, Zach Eaton-Rosen, Weihua Hu, Alexander Merose, Stephan Hoyer, George Holland, Oriol Vinyals, Jacklynn Stott, Alexander Pritzel, Shakir Mohamed, and Peter Battaglia. Learning skillful medium-range global weather forecasting. *Science*, 382(6677):1416–1421, 2023. doi:10.1126/science.adi2336.

[21] J. Lee, R.C. Weger, S.K. Sengupta, and R.M. Welch. A neural network approach to cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, 28(5):846–855, 1990. doi:10.1109/36.58972.

[22] Hao Li, Zheng Xu, Gavin Taylor, Christoph Studer, and Tom Goldstein. Visualizing the Loss Landscape of Neural Nets, 2018. arXiv:1712.09913.

[23] Qimai Li, Zhichao Han, and Xiao-ming Wu. Deeper Insights Into Graph Convolutional Networks for Semi-Supervised Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), Apr. 2018. doi:10.1609/aaai.v32i1.11604.

[24] Ruoyu Li, Sheng Wang, Feiyun Zhu, and Junzhou Huang. Adaptive Graph Convolutional Neural Networks. *CoRR*, abs/1801.03226, 2018. arXiv:1801.03226.

[25] Jia Liu, Yongjian Sun, Kaijun Ren, Yanlai Zhao, Kefeng Deng, and Lizhe Wang. A Spatial Downscaling Approach for WindSat Satellite Sea Surface Wind Based on Generative Adversarial Networks and Dual Learning Scheme. *Remote Sensing*, 14(3), 2022. doi:10.3390/rs14030769.

[26] Yunjie Liu, Evan Racah, Prabhat, Joaquin Correa, Amir Khosrowshahi, David Lavers, Kenneth Kunkel, Michael F. Wehner, and William D. Collins. Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets. *CoRR*, abs/1605.01156, 2016. arXiv:1605.01156.

[27] Fei Ma, Fei Gao, Jinping Sun, Huiyu Zhou, and Amir Hussain. Attention Graph Convolution Network for Image Segmentation in Big SAR Imagery Data. *Remote Sensing*, 11(21), 2019. doi:10.3390/rs11212586.

[28] Marta Moreno-Ibáñez, René Laprise, and Philippe Gachon. Recent advances in polar low research: current knowledge, challenges and future

perspectives. *Tellus A: Dynamic Meteorology and Oceanography*, 73(1):1–31, 2021. doi:10.1080/16000870.2021.1890412.

[29] Christopher Morris, Martin Ritzert, Matthias Fey, William L. Hamilton, Jan Eric Lenssen, Gaurav Rattan, and Martin Grohe. Weisfeiler and Leman Go Neural: Higher-order Graph Neural Networks. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*. AAAI Press, 2019. doi:10.1609/aaai.v33i01.33014602.

[30] Hemani Parikh, Samir Patel, and Vibha Patel. Classification of sar and polsar images using deep learning: a review. *International Journal of Image and Data Fusion*, 11(1):1–32, 2020. doi:10.1080/19479832.2019.1655489.

[31] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. *PyTorch: an imperative style, high-performance deep learning library*. Curran Associates Inc., Red Hook, NY, USA, 2019. arXiv:1912.01703.

[32] Jaideep Pathak, Alexander Wikner, Rebeckah Fussell, Sarthak Chandra, Brian R. Hunt, Michelle Girvan, and Edward Ott. Hybrid forecasting of chaotic processes: Using machine learning in conjunction with a knowledge-based model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(4):041101, 04 2018. doi:10.1063/1.5028373.

[33] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. arXiv:1201.0490.

[34] Xiaosheng Peng, Hongyu Wang, Jianxun Lang, Wenze Li, Qiyou Xu, Zuowei Zhang, Tao Cai, Shanxu Duan, Fangjie Liu, and Chaoshun Li. EALSTM-QR: Interval wind-power prediction model based on numerical weather prediction and deep learning. *Energy*, 220:119692, 2021. doi:10.1016/j.energy.2020.119692.

[35] Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, and Peter W. Battaglia. Learning Mesh-Based Simulation with Graph Networks. *Inter-*

national Conference on Learning Representations, 2020.

[36] PyG Team. conv.GCNConv. https://pytorch-geometric.readthedocs.io/en/latest/generated/torch_geometric.nn.conv.GCNConv.html. [Online; accessed 26.01.2024].

[37] PyG Team. conv.GraphConv. https://pytorch-geometric.readthedocs.io/en/latest/generated/torch_geometric.nn.conv.GraphConv.html#torch_geometric.nn.conv.GraphConv. [Online; accessed 26.01.2024].

[38] Suman Ravuri, Karel Lenc, Matthew Willson, Dmitry Kangin, Remi Lam, Piotr Mirowski, Megan Fitzsimons, Maria Athanassiadou, Sheleem Kashem, Sam Madge, Rachel Prudden, Amol Mandhane, Aidan Clark, Andrew Brock, Karen Simonyan, Raia Hadsell, Niall Robinson, Ellen Clancy, Alberto Arribas, and Shakir Mohamed. Skilful precipitation nowcasting using deep generative models of radar. *Nature*, 597:672–677, 2021. doi:10.1038/s41586-021-03854-z.

[39] Regjeringen.no. Ambitious offshore wind initiative. https://www.regjeringen.no/en/aktuelt/ambitious-offshore-wind-power-initiative/id2912297/. [Online; accessed 18.10.2023].

[40] Markus Reichstein, Gustau Camps-Valls, Bjorn Stevens, Martin Jung, Joachim Denzler, Nuno Carvalhais, and Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature*, 566:195–204, 2019. doi:10.1038/s41586-019-0912-1.

[41] Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. Learning to Simulate Complex Physics with Graph Networks. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 8459–8468. PMLR, 2020. URL: https://proceedings.mlr.press/v119/sanchez-gonzalez20a.html.

[42] M. G. Schultz, C. Betancourt, B. Gong, F. Kleinert, M. Langguth, L. H. Leufen, A. Mozaffari, and S. Stadtler. Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194):20200097, 2021. doi:10.1098/rsta.2020.0097.

[43] I. M. Solbrekke, A. Sorteberg, and H. Haakenstad. The 3 km Norwegian reanalysis (NORA3) – a validation of offshore wind resources in the North Sea and the Norwegian Sea. *Wind Energy Science*, 6(6):1501–1519, 2021. doi:10.5194/wes-6-1501-2021.

[44] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition – Fourth Edition*. Elsevier Academic Press, 2009.

[45] Mathias Tollinger, Rune Graversen, and Harald Johnsen. High-Resolution Polar Low Winds Obtained from Unsupervised SAR Wind Retrieval. *Remote Sensing*, 13(22), 2021. doi:10.3390/rs13224655.

[46] M. Vega-Bayo, J. Pérez-Aracil, L. Prieto-Godino, and S. Salcedo-Sanz. Improving the prediction of extreme wind speed events with generative data augmentation techniques. *Renewable Energy*, 221:119769, 2024. doi:10.1016/j.renene.2023.119769.

[47] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. A Comprehensive Survey on Graph Neural Networks. *CoRR*, abs/1901.00596, 2019. arXiv:1901.00596.

[48] Jie Zhang, Caroline Draxl, Thomas Hopson, Luca Delle Monache, Emilie Vanvyve, and Bri-Mathias Hodge. Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods. *Applied Energy*, 156:528–541, 2015. doi:10.1016/j.apenergy.2015.07.059.

[49] Yuchen Zhang, Mingsheng Long, Kaiyuan Chen, Lanxiang Xing, Ronghua Jin, Michael I. Jordan, and Jianmin Wang. Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*, 619:526–532, 2023. doi:10.1038/s41586-023-06184-4.

[50] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020. doi:10.1016/j.aiopen.2021.01.001.

[51] Xiao Xiang Zhu, Sina Montazeri, Mohsin Ali, Yuansheng Hua, Yuanyuan Wang, Lichao Mou, Yilei Shi, Feng Xu, and Richard Bamler. Deep Learning Meets SAR: Concepts, models, pitfalls, and perspectives. *IEEE Geoscience and Remote Sensing Magazine*, 9(4):143–172, 2021. doi:10.1109/MGRS.2020.3046356.

[52] Chenyi Zhuang and Qiang Ma. Dual Graph Convolutional Networks for Graph-Based Semi-Supervised Classification. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, page 499–508. International World Wide Web Conferences Steering Committee, 2018. doi:10.1145/3178876.3186116.