Norwegian hindcast archive (NORA3) - A validation of offshore wind resources in the North Sea and Norwegian Sea

Ida Marie Solbrekke[1], Asgeir Sorteberg[2], and Hilde Haakenstad[3]

[1]Geophysical Institute, Bergen Offshore Wind Centre (BOW), University of Bergen, Allegaten 70, 5020 Bergen, Norway
[2]Geophysical Institute, Bjerknes Centre for Climate Research (BCCR), Bergen Offshore Wind Centre (BOW), University of Bergen, Norway
[3]Geophysical Institute, University of Bergen, Norwegian Meteorological Institute, Norway

Correspondence: Ida Marie Solbrekke (ida.solbrekke@uib.no)

Abstract. A new high-resolution (3 km) numerical mesoscale weather simulation spanning the period 2004-2018 is validated for offshore wind power purposes for the North Sea and Norwegian Sea. The NORwegian hindcast Archive (NORA3) was created by dynamical downscaling, forced with state-of-the-art hourly atmospheric reanalysis as boundary conditions. A validation of the simulated wind climatology has been carried out to determine the ability of NORA3 to act as a tool for planning
5 future offshore wind power installations. Special emphasis is placed on evaluating offshore wind power-related metrics and the impact of simulated wind speed deviations on the estimated wind power and the related variability. The general conclusion of the validation is that the NORA3 data is rather well suited for wind power estimates, but gives slightly conservative estimates on the offshore wind metrics. Wind speeds are typically 5 % (0.5 ms^{-1}) lower than observed wind speeds, giving an underestimation of offshore wind power of 10 %-20 % (equivalent to an underestimation of 3 percentage point in the capacity factor),
10 for a selected turbine type and hub height. The model is biased towards lower wind power estimates because of overestimation of the frequency of low-speed wind events (<10 ms^{-1}) and underestimation of high-speed wind events (>10 ms^{-1}). The hourly wind speed and wind power variability are slightly underestimated in NORA3. However, the number of hours with zero power production (around 12 % of the time) is fairly well captured, while the duration of each of these events is slightly overestimated, leading to 25-year return values for zero-power duration being too high for four of the six sites. The model is relatively good at
15 capturing spatial co-variability in hourly wind power production among the sites. However, the observed decorrelation length was estimated to be 432 km, whereas the model-based length was 19 % longer.

1 Introduction

Exploiting the Norwegian continental shelf for offshore wind power purposes is highly relevant due to the excellent wind climate (Zheng et al., 2016) and the recent increase in political engagement. In June 2020 the Norwegian government decided
20 to open the country's first two offshore areas at the Norwegian continental shelf, "Utsira Nord" and "Sørlige Nordsjøen II", for concessions to build and operate large wind power installations (Regjeringen, 2020). In this context, the ability to map the spatial and temporal wind power potential is crucial for selecting the best areas for wind power production.

Observational sites in the North sea and the Norwegian sea are sparse and their numbers are insufficient to map the regional wind power potential. The lack of observational data makes it challenging for stakeholders and decision makers to choose new

25   sites to open for offshore wind power concessions. Apart from using satellite data on surface winds, the only way to map the total wind power potential for a large offshore area is to use data from high-resolution numerical weather prediction models that provides data at a typical hub height.

In this study we validate a new and freely available high-resolution data set the NORwegian hindcast Archive (NORA3), to be used to estimate wind power potential on the Norwegian continental shelf. NORA3 is made by the Norwegian Meteorologi-

30   cal institute (Haakenstad et al., 2021), and is a high-resolution atmospheric dynamic downscaling of the ERA-5 reanalysis data set from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hersbach et al., 2020). In contrast to older and coarser reanalysis data sets NORA3 contains detailed information on high-resolution wind features. Therefore, the data set brings new aspects to the exploitation of the offshore wind power potential of the North Sea and the Norwegian Sea.

Several studies have mapped the wind energy potential of the North Sea and/or the Norwegian Sea (Berge et al., 2009;

35   Byrkjedal and Åkervik, 2009; Byrkjedal et al., 2010; Skeie et al., 2012; Hasager et al., 2020). Berge et al. (2009) used the mesoscale Weather Research and Forecasting Model (WRF). The model setup composed of three nested domains (18 km, 6 km and 2 km horizontal resolution). The model-integration time was 4 years (2004-2007), and the temporal resolution of the output-data was 1 hour. The model was verified using observations from offshore sites in the North Sea. Based on the comparison with these observations, the authors concluded that the WRF model is a reliable tool for characterizing the average

40   wind conditions in the region in question.

Byrkjedal and Åkervik (2009) simulated the wind power potential of the Norwegian economic zone. The WRF model produced the wind field basis they used in their wind power calculations. The model horizontal resolution was 1 km, and the temporal resolution was 1 hour. The model integration period was one year (2005), but the data was corrected for climatic variations through a separate model run spanning the period from 2000-2008 using two nested domains (16 km, 4 km).

45   Byrkjedal et al. (2010) used the WRF model to simulate the offshore wind power potential in the North Sea. The model domain had a horizontal resolution of 6 km and the model integration period spanned from 2000 to 2009 with hourly model output. Based on their 10-year WRF-simulation they estimated wind power and identified areas with the greatest wind power potential, in addition to the dependency between separation distance and the correlation between two wind power production sites.

50   The New European Wind Atlas (NEWA) has been created as a cooperation between several European wind energy-related companies. The NEWA aims to provide a high-resolution, freely available data set on wind energy resources in Europe (NEWA, 2020). The NEWA was created using the numerical weather prediction (NWP) WRF model running 10 separate model runs for 10 independent regions (see Fig 1. in Dörenkämper et al. (2020)). Each region is composed of three one-way nested domains (27 km, 9 km, and 3 km horizontal resolution) giving model output every 30 min.

55   A peer-reviewed validation of wind model simulations before using the data for offshore wind power purposes is very important. Using either onshore observations or satellite data of surface winds the studies mentioned above to some degree

conducted peer-reviewed validation their simulated offshore wind results. The lack of detailed in situ validation of previous
data sets and the introduction of a new numerical model increases the relevance of the present study.

We validate the NORA3 data set for wind power purposes using observational wind data from six offshore sites. In Chapter
60   3 we investigate how well NORA3 captures the statistical wind speed measures and the related distributions. We also study the
model performance in terms of the wind speed ramp-rates (Section 3.1), spatial wind speed gradient (Section 3.2), and wind
direction (Section 3.3). In addition, uncertainties related to observations sampled at large structures are discussed in (Section
3.4). After converting wind speed data to hourly wind power data, we examine the performance of NORA3 related to wind
power climatology (Chapter 4), including wind power variables such as median production and capacity factor (CF). Revealing
65   the wind power potential in an area also requires mapping the wind power intermittency and variability at different spatial and
temporal scales. The ability of the model to capture wind power variability and intermittency are investigated using hourly
wind power ramp-rates (Section 4.1) and long-term variability in CF (Ch. 4.2). In addition to temporal variability, we also
consider the ability of NORA3 to capture the spatial co-variability between production sites (Section 4.3). It is crucial for a
data set to reveal the length, duration, and total number of hours of zero wind power production, and NORA3 performance
70   against these measures is discussed in Section 4.4. Moreover, we calculate and validate the maximum expected length of a
zero-event occurring at least once during the turbine lifetime (Section 4.5). In the last chapter (Chapter 5) we summarize the
validation results.

2   Data and method

2.1   Model data

75   NORA3 is obtained by high-resolution atmospheric dynamic downscaling of the state-of-the-art ERA-5 reanalysis data set from
the ECMWF (Hersbach et al., 2020). ERA-5 covers the Earth in an approximately 31 x 31 km horizontal grid, providing hourly
information in 137 vertical layers. The model used in the downscaling process is the nonhydrostatic, convection-permitting
NWP model HARMONIE-AROME (Cycle 40h1.2). Boundary values from ERA-5 are provided to the model every 6 h.
Hourly[1] NORA3 output data is stored (some outputs are stored every third hour). The model domain in NORA3 encloses
80   almost the entire northern part of the Atlantic Ocean (see Fig. 1), and the model runs with a horizontal resolution of 31 x 31
km, with the atmosphere divided into 65 vertical layers.

The model is a high-resolution NWP model solving the fully compressible Euler equations using forward time integration,
and it is used in short-range operational forecasting and research by many European countries (Seity et al., 2011; Bengtsson
et al., 2017). The NORA3 data set is a hybrid between a hindcast and a reanalysis data set because of to the way the observations
85   are treated in the model. The HARMONIE-AROME model performs data assimilation of 2-m temperature and 2-m relative
humidity.

---

[1]instantaneous values
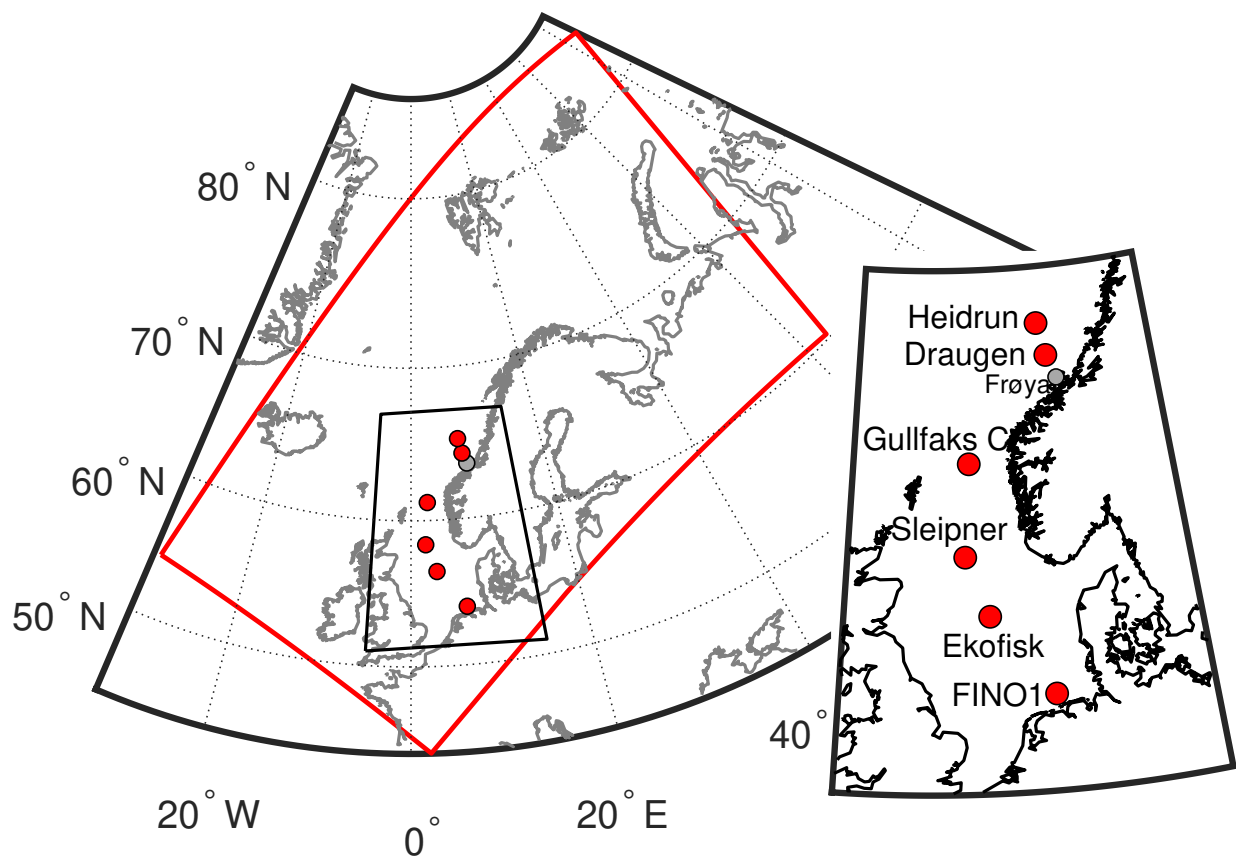
Figure 1. The domain (red rectangle) covered by the HARMONIE-AROME simulation and the locations of the six sites used in verifying the NORA3 data set (red dots), in addition to the met mast located at Frøya. A zoom-plot of the positions and the names of the stations is also shown. Details for the sites are given in Table 2.

The NORA3 data is so far covering the period 2004-2018 and will later be extended back to 1979 and also updated continuously. We will therefore focus on the period 2004-2018 in this study. For further details on the model set-up, the NORA3 generation process, and a comparison between NORA3 and ERA-5 see Haakenstad et al. (2021).

2.1.1 Offshore wind data sets

Table 1 compares five data sets covering the offshore areas surrounding Norway. NWP models are continuously updated and improved. The first three data sets ("BYR2009","BYR2010", and "BER2009") are approximately 10 years old, thus the output from the newer data sets NEWA and NORA3 has higher quality given the improvements in NWP models over the last 8-10 years.

In contrast to the older data sets, NEWA and NORA3 are created using improved reanalysis products as boundary information. At 31 km ERA-5 is able to resolve smaller scale atmospheric features than the data set from the National Centers for Environmental Protection (NCEP) (1° resolution), and the downscaling process done by WRF and HARMONIE-AROME benefits from the increased resolution in the reanalysis. In addition to the improved boundary conditions, the longer time periods covered in NEWA and NORA3 contain longer-term variability and statistics, which are beneficial when dealing with fluctuating geophysical variables and making decisions about wind power installations.

Not all of the model data sets listed in Table 1 underwent a peer-review process (see column "Val"). NEWA uses met masts on land to validate the onshore model data, while the offshore data is validated at 10 meters above sea level (m.a.s.l.) using satellite data. In addition, a validation at 100 m.a.s.l. is conducted by extrapolating the 10-m data. In addition to the validation conducted by Haakenstad et al. (2021), the offshore data of NORA3 is validated using in situ observations near a typical hub height (100 m.a.s.l.). The near-hub-height validation of the NORA3 data minimizes the uncertainties introduced by power-law extrapolation of the wind from near surface to hub-height.

2.2 The observational data

The observations used in the verification of NORA3 are hourly wind observations [2] from five oil- and gas platforms (Ekofisk, Sleipner, Gullfaks C, Draugen, and Heidrun) and one met-mast (Fino1) (see Fig. 1 for the location of the sites and Table 2 for further site information), retrieved from the Norwegian Meteorological Institute. The observational data was quality checked prior to the validation of NORA3. For a detailed description of this quality check process see Solbrekke et al. (2020). In addition to the routine Solbrekke et al. (2020) describe we also exclude all records of zero-wind conditions (u=0) that are likely to be erroneous according to the following:

$$u_{obs}(i) = 0 \quad \wedge \quad u_{n3}(i) \geq \frac{5}{m}\frac{1}{n}\sum_{j=1}^{m}\sum_{i=1}^{n}|u_{obs}(i,j) - u_{n3}(i,j)| = 5\overline{MAD}, \tag{1}$$

where $u_{obs}(i,j)$ and $u_{n3}(i,j)$ are the observed and modeled wind speeds, respectively, at hour i for site j. n is the total number of hours; m is the total number of sites; and $\overline{MAD}$ is the mean absolute deviation between the observed and modeled wind speeds

---

[2]10-min average values: the average of 5 min before and 5 min after every hour

Table 1. Information for five offshore wind data sets covering the Norwegian continental shelf and the surrounding offshore areas. "BYR2009" is the data set created by Byrkjedal and Åkervik (2009); "BYR2010" was produced by Byrkjedal et al. (2010); "BER2009" comes from Berge et al. (2009); NEWA is the New Europen Wind Atlas NEWA (2020); and NORA3 was created by Haakenstad et al. (2021). "Model" indicates the model name used in the creation of each data set, with the version of the model given in parentheseis. Information regarding the name of the reanalysis product (model), the corresponding horizontal resolution in kilometers (res) and the forcing update of lateral boundaries (int) are found under "Boundary". "Period" is time period covered by the data set. "Temp res" is the number of outputs stored per hour per variable, and "Horiz res" is the horizontal resolution, and "Vert res" is the number of vertical layers in the model. "Val" tells whether the model data set is validated, and "PA" tells if the data set is publicly available for downloading. Abbreviations: WRF: Weather Research and Forecasting model; NCEP: National Centers for Environmental Protection; H-A: HARMONIE-AROME; ERA-5: the name of the newest reanalysis product from the European Centre for Medium Range Weather Forecasts.

Offshore wind data sets									
Name	**Model** (version)	**Boundary** (model,res,int)			**Period**	**Temp res**	**hor res**	**vert res**	**val**	**PA**
BYR2009	WRF (v 3.0.1)	NCEP	1°	6	2005 (2000-2008)	1	1 (16, 4)	-	yes	no
BYR2010	WRF (v 3.0.1)	NCEP	1°	6	2000-2009	1	20, 5 (18, 6)	-	no	no
BER2010	WRF (v -)	NCEP	1°	6	2004-2007	1	18, 6, 2	32	no	no
NEWA	WRF (v 3.8.1)	ERA-5	31 km	6	1989-2018	2	27, 9, 3	61	yes	yes
NORA3	H-A (Cy 40h1.2)	ERA-5	31 km	6	2004-2018	1	3	65	yes	yes

averaged over all sites. In other words, whenever the observed wind speed at hour i and site j is zero and the corresponding modeled wind speed exceeds $5\overline{MAD} = 7.2 \text{ ms}^{-1}$, the observed value at hour i is excluded from the time series for site j. This additional quality control leads to the exclusion of up to 5 h of observations per site, except at Heidrun which excludes 58 h of observations. For Heidrun, the removal of these erroneous records of zero-wind conditions (u=0) corresponds to an exclusion of approximately 0.035 % of the total data.

2.3 Wind speed interpolation

To avoid introducing additional uncertainties into the observational data set, we verify the wind variables from NORA3 at the wind sensor heights, ranging from 68-140 m.a.s.l., for each site (see "WSH" in Table 2 for the sensor heights). By contrast, the wind power verification is performed at a typical hub-height, at 100 m.a.s.l., to ensure the production estimate are comparable between sites.

The interpolation of wind speed data to sensor height or hub-height is done using the power law relation (Emeis, 2018). The interpolated wind speed is sensitive to the choice of the power law exponent α. Usually, α is assigned based on assumptions about atmospheric stability and surface roughness, both of which can introduce erroneous results. However, the data from

NORA3 allows us to calculate α for each time step (i). Rearranging the power law relation, we get the following expression for the power law exponent α:

$$\alpha(i) = \frac{ln\frac{u(i)_{z2}}{u(i)_{z1}}}{ln\frac{z_2}{z_1}}, \tag{2}$$

where the height-subscripts $_1$ and $_2$ corresponds to the two layers within which the wind shear is calculated. The heights used to calculate α depend on the wind-sensor height (WSH) at the site in question: if WSH < 100 m.a.s.l. then α is calculated using NORA3 wind-shear between the two model layers $z_1 = 50$ m.a.s.l. and $z_2 = 100$ m.a.s.l.. If WSH >100 m.a.s.l. then alpha is calculated using the wind shear between $z_1 = 100$ m.a.s.l. and $z_2 = 250$ m.a.s.l.. The mean α for the whole time period for the six stations ranges from 0.05-0.08 between 50 and 100 m.a.s.l., and 0.03-0.06 between 100 and 250 m.s.a.l..

2.4 Normalized wind power

To ensure our validation results are as general as possible, and since the turbine park at each site is only imaginary and of unknown capacity, we use normalized power calculations $P_w(i) = \frac{P_w^T(i)}{P_w^{max}}$ to validate the wind power potential at each site (Solbrekke et al., 2020). $P_w^T(i)$ is the produced wind power at each time step (i) for a given site, and P_w^{max} is the nameplate capacity. Hence, the normalized wind power $P_w(i)$ is defined as follows:

$$P_w(i) = \begin{cases} 0 & u(i) < u_{ci}, \\ \frac{u(i)^3 - u_{ci}^3}{u_r^3 - u_{ci}^3}, & u_{ci} \le u(i) < u_r, \\ 1, & u_r \le u(i) < u_{co}, \\ 0, & u_{co} \le u(i), \end{cases} \tag{3}$$

where $u(i)$ is the wind speed at hour i, $u_{ci} = 4$ ms^{-1} is the cut-in wind speed, $u_r = 13$ ms^{-1} is the rated wind speed, and $u_{co} = 25$ ms^{-1} is the cut-out wind speed. These numbers were retrieved from the SWT-6.0-154 turbines used in Hywind, Scotland - the first floating wind park in the world (AG, 2011).

2.5 Ramp rates

To validate the ability of NORA3 to capture the wind speed and wind power variability we calculate the ramp rates (R), defined as how much the wind speed (u) or wind power (P_w) changes during a time-increment τ (Milan et al., 2014):

$$R_{P_w}(i) = P_w(i) - P_w(i+\tau), \tag{4a}$$

$$R_u(i) = u(i) - u(i+\tau), \tag{4b}$$

Table 2. Relevant information for the sites used in the validation of NORA3. "Abb" lists the site-name abbreviations. "Lat" and "Lon" are the latitude and longitude for the site locations, respectively. "WSH" (in meters above sea level) corresponds to the wind sensor height at each site. The sensor type is listed under "Sensor", and the data period for the available observations for each site is listed under "Data period". In addition, the percentage of valid observations is also shown under "Valid obs (%)".

					Site information		
Site	**Abb**	**Lat**	**Lon**	**WSH**	**Sensor**	**Data period**	**Valid obs (%)**
Fino1	f1	54.02	06.59	100	A100LK Cup-anemometer	01.01.2004-31.12.2018	75.1
Ekofisk	ek	56.52	03.22	68/102	Vaisala WMT703	01.01.2000-31.09.2016	85.3
Sleipner	sl	58.37	01.91	136	Gill Ultrasonic	01.01.2000-31.09.2016	83.4
Gullfaks C	gf	61.22	02.27	140	Gill Ultrasonic	01.01.2000-31.09.2016	80.2
Draugen	dr	64.35	07.78	78	Gill Ultrasonic	01.01.2002-31.09.2016	66.6
Heidrun	he	65.33	07.78	131	Gill Ultrasonic	01.01.2000-31.09.2016	84.6

setting τ to be 1, we validate the model performance on hourly ramp rates. To gain a general picture of the model performance in terms of how much the wind speed or wind power changes from one hour to the next we calculate the mean absolute ramp rate (MAR) for each site, for both the observational data and the modeled data. MAR is defined as follows:

155
$$MAR = \frac{1}{n}\sum_{i=1}^{n}|R(i)|, \tag{5}$$

where $R(i)$ is the ramp rate at hour i, and n is the total number of hours.

2.6 Zero-event duration using extreme value theory

A wind turbine has an expected life time of approximately 20 years. If the right steps are taken, the lifetime can be extended
15% -25 % depending on whether the structure is bottom-fixed or floating (Wiser et al., 2016). This means that the lifetime is
160 expected to increase to 23-25 years. Therefore, determining the duration of long-lasting shutdowns expected to happen during the lifetime of a turbine is important for estimating the levelized cost of energy (LCOE). The 25-year return value of the duration of a zero-event (a period of zero wind power production), the corresponding confidence interval, and the p-values are calculated from the observations and the model data using two statistical methods, "block maxima" (BM) in which the data is fitted to a generalized extreme value (GEV) distribution using yearly values of maximum zero-event duration, and "peak over threshold"
165 (POT) in which the data is fitted to a generalized Pareto distribution (for more information see Smith (2002)) using the 99th-percentile of zero-event duration (the highest 1 % of zero-event in terms of duration) as the selected threshold. We calculate the Kolmogorov-Smirnov p-value (KS_p) to test the null hypothesis. The Kolmogorov–Smirnov statistic quantifies the distance between the empirical and the theoretical cumulative distribution functions. Hence, the cumulative distribution function from the BM data (POT data) is compared to the cumulative distribution function from the GEV (Pareto) distribution. Thus, given
170 a significance level of p=0.025, if the KS_p value is small ($KS_p < p$), the distance between the cumulative distributions is too

large, and we can conclude that the empirical data (BM or POT) was sampled from a different population than the theoretical GEV or Pareto distribution with a probability of 1-p.

3 Validation of wind speed

The NORA3 wind estimates are extensively validated against observations and compared against the ERA-5 reanalysis in Haakenstad et al. (2021). For offshore observations (six oil platforms, many the same as used in this study) the NORA3 wind estimates were shown to be better than the wind estimates from ERA-5 for all months and for all investigated percentiles of wind speed. Monthly wind speed biases were typically reduced from 6-8 % to 3-5 %. The improvement was particularly pronounced for strong winds, where the bias was reduced from 10-20 % to 2-4 %, while the bias reduction for median winds typically was reduced from 7-8 % to 3-4 %. In addition, improvements in coastal winds influenced by topography were shown to be significantly larger than for the offshore stations.

Prior to exploiting NORA3 as a planning tool for future offshore wind power installations the model performance in terms of variables related to wind power also has to be validated and verified. We start with the validation of mean quantities and distributions of wind speed. The most relevant wind speed measures can be seen in Table. 3. Arithmetic mean (μ) is used as a measure of the average wind speed. The corresponding wind speed variability is given by the Weibull standard deviation (std). The Weibull std is used instead of the Gaussian std because of the shape of the wind speed distribution (see Fig. 2a for an example wind speed distribution). Mean wind speeds (μ) for the six sites lie within the interval $10 - 12 \ \mathrm{ms}^{-1}$. For all the sites except Fino1, the observed mean wind speeds are higher than the wind speeds from NORA3, indicating that the model underestimates the mean wind speed. The largest difference can be seen for Sleipner where the observed mean wind speed is 8.9 % higher than the simulated wind speed. The wind speed at each site is highly variable, with the Weibull std (σ) for the observations varying from $4.7 - 6.5 \ \mathrm{ms}^{-1}$, where the model wind speed being slightly less variable (0-20 %). Hence, the observed wind speed is somewhat more intermittent and variable than the modeled wind speed, indicating that HARMONIE-AROME is missing some of the high-frequency variability embedded in the wind field.

The Weibull scale parameter ("a" in Table 3) indicates height and width of the distribution. A larger scale parameter indicates a wider and lower probability distribution. In all but one site the observed scale parameter is slightly higher than the modeled; the modeled scale parameters are on average 3.74 % lower than the observed. In other words, the observations contain more wind speed events at the tails of the Weibull distributions resulting in a larger scale parameter.

As all observed and modeled Weibull shape parameters ("b" in Table 3) are less than 2.6 the distributions are positively skewed, with a long tail to the right of the mean. The observed shape parameter is lower than the modeled (on average 7.6 % lower) indicating that the observed data is more positively skewed with a longer right tail, meaning that the observed data contain more high wind speed events than the NORA3 wind speed data.

According to Table 3 the model underestimates the wind speed at five of the six sites. Since the wind power production is a function of the wind speed cubed the wind power is highly sensitive to systematic deviations between the observed and simulated wind speeds. However, the sensitivity varies with wind speed and is especially strong within the interval between
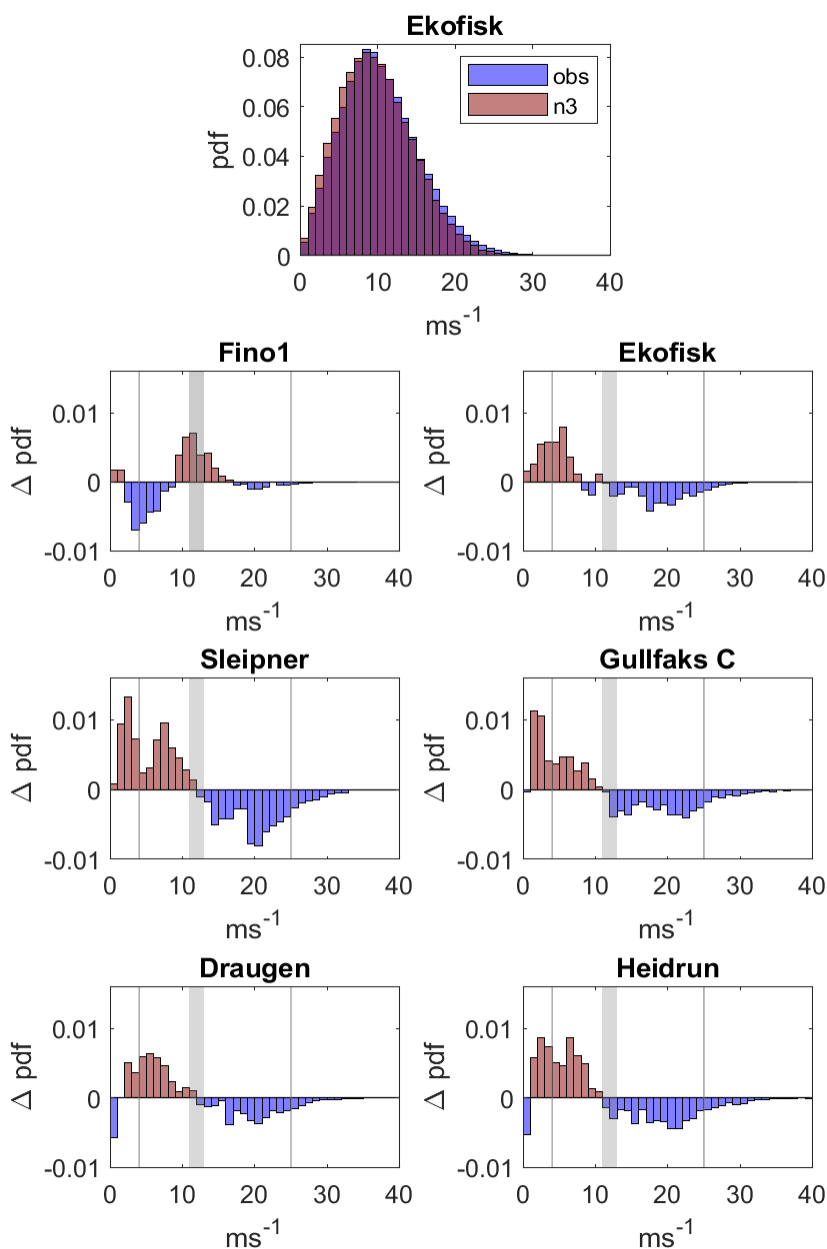
Figure 2. a) Example wind speed probability density function (pdf) (Ekofisk) for NORA3 (n3) in red and observations (obs) in blue. b)-h) Differences between NORA3 and observational wind speed probability density functions ($\Delta pdf = pdf_{mod} - pdf_{obs}$) for the six sites. When $\Delta pdf = 0.01$ the probability that the given wind speed will occur is 1 % higher in the model output. The large gray area corresponds to the range within which the rated wind speed usually falls. The gray vertical lines at the left and right mark the cut-in and cut-out wind speed limits used in this study, respectively.

Table 3. Statistical measures of the wind speed for the observations (Obs) and the model (n3). μ is the arithmetic mean while σ is the Weibull standard deviation. "a" and "b" are the Weibull scale and shape parameters, respectively. The wind speed validation is performed at the sensor height to avoid uncertainties related to power-law extrapolation (see Table 2 for information on heights).

		Wind speed								
		μ			σ			a		b
Site	Obs	n3	%	Obs	n3	%	Obs	n3	Obs	n3
FINO1	9.62	9.72	**1.04**	4.66	4.56	**-2.15**	10.86	10.96	2.18	2.25
Ekofisk	10.35	9.85	**-4.83**	5.09	4.76	**-6.48**	11.61	11.12	2.12	2.18
Sleipner	11.70	10.66	**-8.89**	6.19	5.44	**-12.12**	12.99	12.02	1.94	2.05
Gullfaks C	11.45	10.70	**-6.55**	6.39	5.67	**-11.27**	12.64	12.06	1.82	1.97
Draugen	9.87	9.44	**-4.36**	5.76	5.06	**-12.15**	10.95	10.65	1.75	1.94
Heidrun	10.56	9.87	**6.53**	6.45	5.22	**-19.07**	11.50	11.13	1.64	1.97
Average	**10.59**	**10.04**	**-5.19**	**5.76**	**5.12**	**-11.11**	**11.76**	**11.32**	**1.91**	**2.06**

cut-in and rated wind speeds. Fig. 2b-h shows the differences in the observed and modeled wind speed probability density functions (Δ pdf = pdf mod - pdf obs) for the six sites, in addition to the wind speed distribution for Ekofisk (Fig. 2a). The main finding is that the model underestimates the number of events with high wind speed, and overestimates the number of events with low wind speed for all sites, except Fino1. The model is biased towards too few high-wind events and too many low-wind events than the observations, and the transition occur near the typical rated wind speed ($11 - 13 \text{ ms}^{-1}$) for state-of-the-art wind turbines (the widest gray area in Fig. 2, panels b-h). This model-bias will have a large impact on the difference between the observed and modeled wind power.

3.1 Wind speed ramp rates

The wind speed ramp rate is a measure (ms^{-1}) of the hourly variability in the data set. In other words, the ramp rate quantifies how much the wind speed changes during 1 h. Figure 3 shows the distributions of observed and modeled hourly wind speed ramp rates for Ekofisk (the other sites have similar distributions). The distribution is wider for the observations than for the modeled data, illustrating that the observed wind speed change from one hour to the next is greater than that in the modeled wind speed data.

The mean absolute ramp rate (MAR) for the observed wind speed (u) is shown in Table 4. Typically observed MAR is around 1 ms^{-1} and the difference between modeled and observed ramp rates (Table 3) indicates that the model underestimates the variability in hourly wind speed by 30 %-36 %.

3.2 Far offshore to coastal wind speed gradient

An important feature of a model wind data set is the ability to properly estimate the horizontal wind speed gradient from far offshore to coastal areas. There are limited possibilities to investigate this using the available observational data. However, we
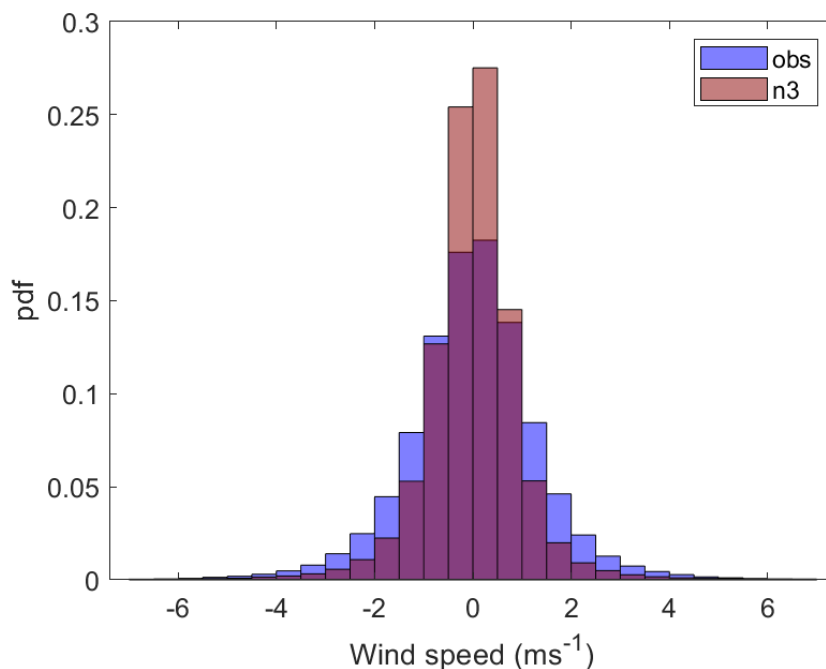
Figure 3. The probability density distribution (pdf) of the modeled (n3) and observed (obs) hourly wind speed ramp rates.

Table 4. Mean absolute ramp rate (MAR) for the observed and modeled wind speed (u). The difference between the modeled and observed MAR_u divided by the observed MAR_u, is given as a percentage (%).

	MAR_u		
Site	Obs	n3	%
Fino1	1.00	0.68	**-32.00**
Ekofisk	1.04	0.67	**-35.58**
Sleipner	1.15	0.75	**-34.78**
Gullfaks C	1.15	0.81	**-29.57**
Draugen	1.31	0.85	**-35.11**
Heidrun	1.22	0.80	**-34.43**

made use of data from an observational met-mast situated on the coastal island of Frøya (see Fig. 1) to present some indicative results. Generally, using wind speed data at sensor height for the three sites Heidrun (far offshore), Draugen (near coastal) and Frøya (coastal) as depicted in Fig. 1, shows that there is no clear bias in the model (see Table 5). NORA3 underestimates the local far-offshore to near-coastal wind speed gradient, but slightly overestimates the near-coastal to coastal gradient.

Table 5. The average change in wind speed at sensor height per 100 km kilometer for Heidrun-Draugen and Draugen-Frøya.

	Wind speed gradient		
Site	obs	n3	diff
Heidrun-Draugen	0.66	0.43	**-0.23**
Draugen-Frøya	1.52	1.68	**0.16**

3.3 Wind direction

Another important factor for planning a wind park using modeled data is the quality of the modeled wind direction. State-of-the-art wind turbine technology allows the wind turbines to yaw to face the main wind direction. Mapping the wind direction

230 climatology is important for the wind park configuration and internal positioning of the turbines. Wind-rose plots (see Chapter A Fig. A1) demonstrate that the modeled and observed data in general show the same wind direction distributions, with only small differences, except for Fino1. Fino1 is excluded from the verification of wind direction because the wind rose for that site shows a clear directional disturbance, as the wind is affected by the observation mast (see Fig. A1, lower most row). Fig. 4 graphs the differences between the modeled and observed data (%) in the number of wind direction events ($30°$-intervals) for

235 four wind speed categories ($u < u_{ci}$, $u_{ci} \leq u < u_r$, $u_r \leq u < u_{co}$, and $u_{co} \leq u$). There is no systematic bias in wind direction that can be seen across the sites, and the biases in frequency are less than 5% for all directional intervals and all sites. The wind speed interval with the greatest difference between the model and the observations features wind events corresponding to $u \geq u_{co}$. The wind speed interval with the smallest difference between the model and the observation are the "too low" wind events ($u < u_{ci}$). Hence, the model is better at capturing the wind direction when the wind speed is low.

240 Sleipner is the site with the greatest difference between model and observations for almost all wind direction intervals. The mismatch between the observed and modeled wind direction events for Sleipner is probably tied to the model performance. But, we cannot rule out that the platform design at Sleipner affects the flow field more than the design of the other platforms.

3.4 Uncertainties in observed wind speed

Working with observational data and numerical weather prediction models involves dealing with data that contains uncertainties

245 and errors of known or unknown character. The majority of the observational sites used in this study (five of six sites) are oil- and gas platforms. The platforms are large structures that may influence the upcoming flow. On the other side, an observational mast may also influence the flow when the upcoming wind is guided to pass through the mast before being recorded by the sensor.

To what extent these large offshore structures influence the ambient flow field is unclear (Berge et al., 2009; Vasilyev et al.,

250 2015; Furevik and Haakenstad, 2012). Nevertheless, using observations from oil- and gas platforms enable us to validate NORA3 over ocean areas where observational data is sparse. A potential alteration of the wind would be a function of the platform layout, the atmospheric stability, the upcoming wind direction, and the ambient wind speed. To investigate the distortion
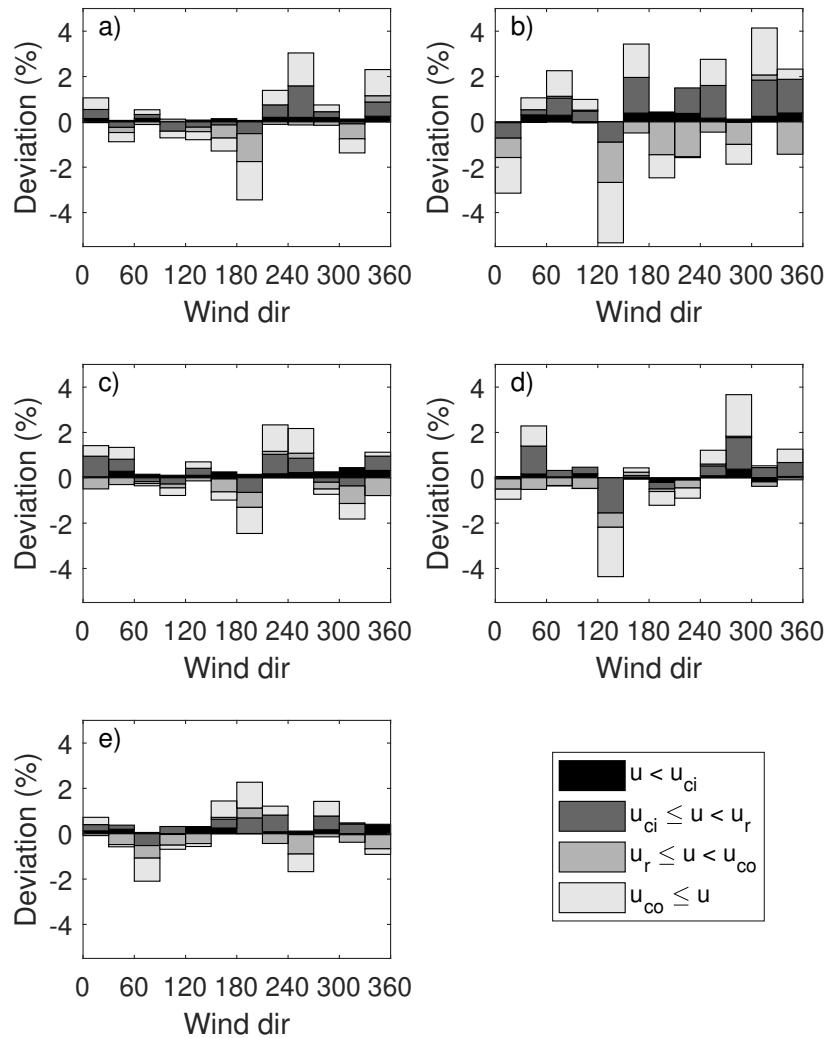
Figure 4. Difference in the occurrence (%) of wind events categorized in different wind direction-intervals (30°-intervals) between NORA3 and observations (model-obs) for a) Ekofisk; b) Sleipner; c) Gullfaks C; d) Draugen; e) Heidrun. For each wind direction interval the wind events are divided into four different wind speed categories, the first one corresponds to u less than cut-in wind speed ($u < u_{ci}$), the second is the wind speed interval where the the wind power is a function of the wind speed cubed ($u_{ci} \leq u < u_r$), the third interval contain the wind speeds corresponding to rated wind power production ($u_r \leq u < u_{co}$), and the last interval is where the wind speeds are too strong resulting in a terminated wind power production ($u_{co} \leq u$).

caused by these large structures, we compared wind speed data from the platforms with data from Fino1, and from the met
mast at the Frøya field station.

255   The main results are illustrated and discussed in Appendix B (Figs. B1 and B2) and indicates that flow-disturbance by large
oil- and gas platforms is to some extent visible in the wind speed data for some of the platforms. However, results illustrated
in Figs. B2 and Fig. 4 show no clear signal indicating whether the large structures distort the flow field.

4   Comparison of estimated wind power from observed and modeled wind speed

Because the conversion from wind speed to wind power is non-linear (see Eq. 3), the wind power distribution differs greatly
260   from the wind speed distribution. The statistical measures for the wind power are shown in Table 6. Median (q_{50}) and inter-
quartile range (IQR) are independent of data distribution and are therefore good representations of the average wind power
production and the intermittency, respectively. All wind power estimates are calculated at a hub height of 100 m.a.s.l. using the
wind extrapolation given in Section 2.3 and the power curve given in Section 2.4.

Both the observation-based and model-based median wind power production estimates reveal very good wind power poten-
265   tial for the six sites (Table 6). Nevertheless, since the model underestimates those wind speeds exceeding the rated wind speed,
this partly counteracts the model's overestimation of low wind speeds, making the modeled total power production slightly
underestimated. Therefore, the observation-based estimates of the median hourly power production q_{50} span from 0.3-0.5 (i.e.
the median power production for a given hour would typically be 30 %-50 % of installed capacity), compared to 0.3-0.4 for the
model-based estimates. IQR, a measure of the variability, is the range between the first and third quartiles ($q_{75} - q_{25}$). Since the
270   range of the normalized wind power P_w is 0-1, IQR values close to 1 correspond to high variability, since almost the entire data
range is present between the first and third quartiles. Hourly IQRs range from 0.84-0.95 for the observation-based estimates
and from 0.84-0.94 for the model. However, there is no systematic difference between the IQRs of the model-based estimates
and the observation-based estimates.

The capacity factor (CF) is another statistical measure quantifying the wind power potential. CF is here defined as the average
275   wind power potential divided by the installed capacity. The observation-based estimates of CF vary between 0.4 and 0.6, and
the CF values from the model-based estimates are slightly smaller (except for Fino1). The observation-based CF values exceed
the modeled values for five of the six sites, by an average of 3 percentage point.

4.1   Wind power ramp rates

Figure 5 shows the distribution of observation-based and model-based hourly normalized wind power ramp rates for Ekofisk
280   (the other sites have similar distributions). As for the distribution of hourly wind speed ramp rates, the distributions of hourly
wind power ramp rates are wider for the observation-based ramp rates than for the model-based, illustrating that the hourly
estimated wind power variability based on observations is greater than the estimated variability based on NORA3 data. The
difference in MARs indicate an hour-to-hour variability typically 7 %-9 % (Table 7) of the installed capacity based on obser-
vations. In contrast, the variability for model-based estimates is 5 %-6 % and is underestimated at all sites.

Table 6. Statistical measures of the observation-based (Obs) and model-based (n3) wind power production. q_{50} is the hourly median production, IQR is the inter-quartile range of the hourly production, and CF is the wind power capacity factor. The wind power measures and estimates are performed at a typical hub-height of 100 m.a.s.l. using the interpolation of observed wind speeds as outlined in Section 2.3 and the power curve given in Section 2.4 for all the sites.

	Wind power								
	q_{50}			**CF**			**IQR**		
Site	Obs	n3	%	Obs	n3	n3-obs	Obs	n3	%
FINO1	0.33	0.36	**9.10**	0.44	0.46	**0.02**	0.84	0.84	**0**
Ekofisk	0.45	0.40	**-11.1**	0.51	0.48	**-0.03**	0.88	0.91	**3.41**
Sleipner	0.54	0.42	**-22.2**	0.55	0.50	**-0.05**	0.88	0.91	**3.41**
Gullfaks C	0.51	0.42	**-17.7**	0.53	0.50	**-0.03**	0.90	0.92	**2.22**
Draugen	0.33	0.30	**-9.10**	0.46	0.43	**-0.03**	0.95	0.91	**-4.21**
Heidrun	0.39	0.32	**-17.95**	0.48	0.45	**-0.03**	0.93	0.94	**1.08**
Average	**0.43**	**0.37**	**-13.95**	**0.50**	**0.47**	**-0.03**	**0.90**	**0.91**	**1.11**

Table 7. Mean absolute ramp rate (MAR_{P_w}) for the observation-based and model-based estimates of the wind power output. The difference between the modeled and observed MAR_{P_w} divided by the observed MAR_{P_w} is given in percentage (%).

	MAR_{P_w}		
Site	Obs	n3	%
Fino1	0.075	0.051	**-32.00**
Ekofisk	0.079	0.049	**-37.97**
Sleipner	0.077	0.051	**-33.77**
Gullfaks C	0.079	0.054	**-31.65**
Draugen	0.092	0.060	**-34.78**
Heidrun	0.084	0.055	**-34.52**

4.2 Inter-annual and seasonal capacity factor

In addition to encompass short-term variations in wind speed and estimated power production it is essential for a model data set to contain the correct long-term variations. In this section we evaluate NORA3's ability to capture the longer-term climatic variability of the wind power potential for a given site. The inter-annual and seasonal variations in CF provide a good indication of how NORA3 performs in terms of long-term wind power fluctuations.

Fig. 6a and b illustrate the inter-annual and seasonal CF, respectively, from the observation-based estimates. In addition, the CF deviations (Δ CF) between the model-based estimates and the observation-based estimates are illustrated in Fig. 6c and d. The observed year-to-year variation in CF is substantial, varying up to 0.12 (12% of installed capacity) from one year to the
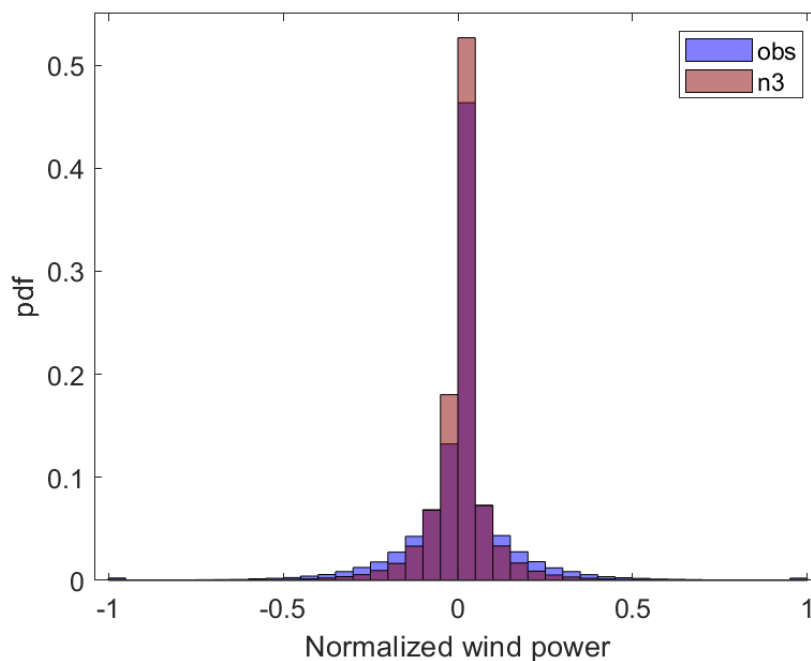
Figure 5. The probability density functions (pdfs) of the ramp rates for observation-based (obs) and model-based (n3) hourly normalized wind power.

next. Fig. 6c shows that the yearly CF values from the model are systematically lower than the observed CF values. This result is most pronounced for Sleipner, where the difference in Δ CF ≈ -0.05, meaning that the model-based CF is on average 5 percentage point lower than the observation-based CF. The Fino1-results are at odds with the results from the other sites; Δ CF for Fino1 demonstrates a positive trend, going from initial negative values to positive values after 2010. This trend is not seen for the other sites.

The model's underestimation of CF can also be seen in the seasonal CF-values. Fig. 6d shows that Δ CF < 0 for all the sites except Fino1. The underestimation of the seasonal CF values is largest during the summer months (May-September), meaning that the relative importance of the summer months in wind power production will be slightly underestimated in the model-based estimates.

4.3 Spatial wind power co-variability

Many studies have shown that interconnection of wind power production sites mitigates wind power intermittency (Kempton et al., 2010; Reichenberg et al., 2014; St. Martin et al., 2015; Reichenberg et al., 2017; Solbrekke et al., 2020). Therefore, modeling tools for use in making decisions about future wind power installations should be able to represent spatial and temporal co-variability between wind power sites.
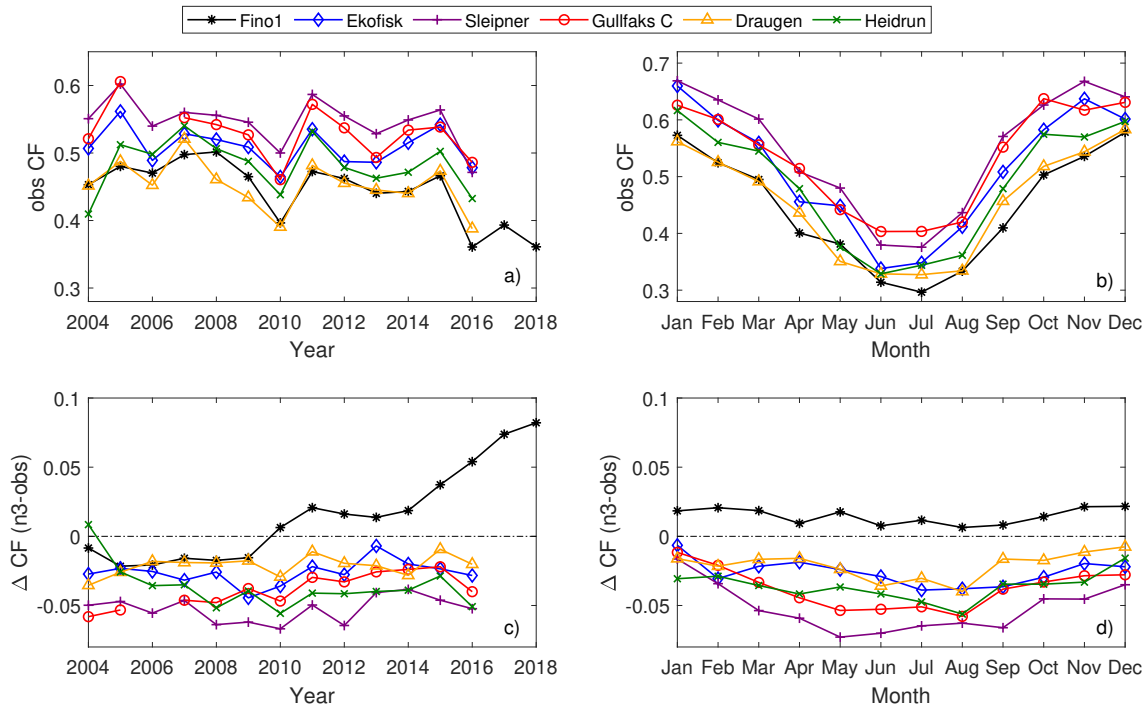
Figure 6. a) The observed inter-annual variation in the capacity factor (obs CF). b) Seasonal variation in CF for the observations. c) The difference in the inter-annual CF (Δ CF) between the model and the observations (n3-obs). d) The difference in the seasonal CF between the model and the observations (n3-obs). A specific year was excluded from the plot if more than one-half of the data for that year was missing.

Figure 7 illustrates the ability of NORA3 to capture the spatial co-variance in estimated hourly wind power production between the six sites. The figure demonstrates how the correlation between two sites changes as a function of the separation distance, both for the observation-based estimates (blue) and the model-based estimates (red). For almost all separation distances the model overestimates the correlation between two connected sites. The overestimation is generally small, but is greatest for small separation distances. This result indicates that NORA3 is better at capturing the large-scale spatial variability than variance on smaller scales.

A general description of the dependency between correlation and separation distance can give us information on the de-correlation length for the sites used in this study. Using the station-pair correlations we identify a best fitting exponential curve and a de-correlation length L (in kilometers). Connecting sites separated by a distance greater than the de-correlation length ensures that the collective wind power intermittency from the two sites is substantially reduced compared to the intermittency from one of the sites. We use the e-folding distance[3] as a measure of the offshore de-correlation length L. The exponential curves and the corresponding de-correlation lengths for both the observations and NORA3 are presented in Fig. 7. The observation-based L is 432 km compared to a 513-km L based on NORA3. The model-based estimates indicate that to ensure relatively

---

[3]The distance where the correlation has dropped to $\frac{1}{e} = 0.37$.
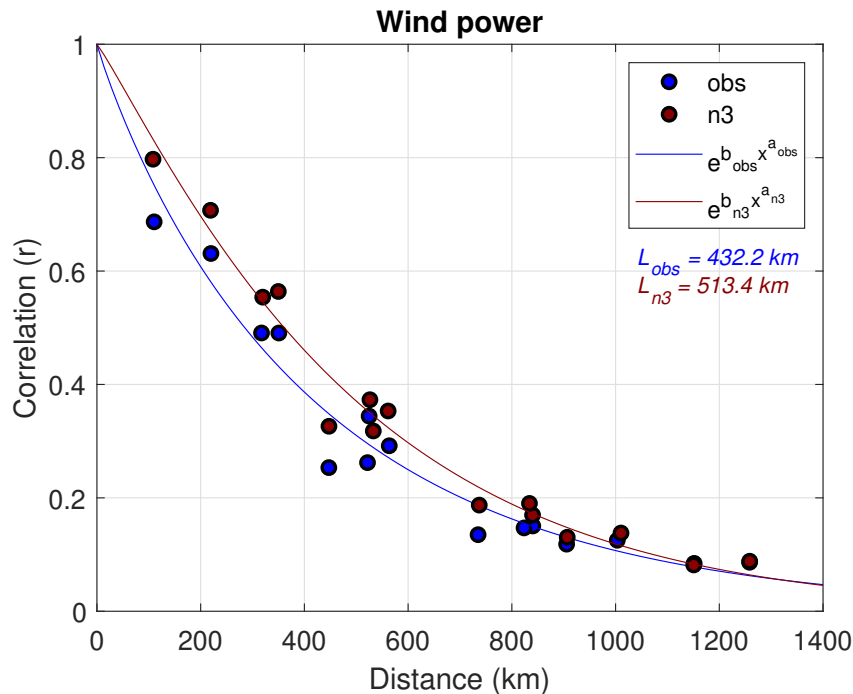
Figure 7. Correlation of wind power time series as a function of the distance between the connected site-pairs for the observations (obs, blue) and NORA3 data (n3, red). An exponential fit is also shown (e^{bx^a}) for both the data sets with the corresponding de-correlation lengths, L.

320   independent hourly power production, a greater interconnection distance is needed than that indicated by the observation-based estimates.

4.4   Zero wind power events

Knowing about the risk, duration, and frequency of zero-events (periodes of zero wind power production) is important for decision-making and also in turbine maintenance planning, as these measures influence the levelized cost of energy and hence

325   the decision-making process (Cory and Schwabe, 2009). A zero-event is caused by a wind speed that is too low ($u < u_{ci}$) or too high ($u \geq u_{co}$), and these events depend to some extent on the technical specifications of a wind turbine, but also, and more significantly, on the ambient wind climate in the area of interest. Table 8 shows the percentages of all hourly wind speed data values that fall into each wind power category ($u < u_{ci}$, $u_{ci} \leq u < u_r$, $u_r \leq u < u_{co}$, and $u_{co} \leq u$) for each site. In addition, the table lists the total risk of having zero wind power production ($P_w = 0$). The percentage of hours when the wind is too weak to

330   produce wind energy ($u < u_{ci}$) ranges from 8 % to 14% in the observation-based estimates and is overestimated by the model by an average of 1.4 percentage point for all sites except Fino1. On the other hand, the observation-based estimates indicate that the fraction of hours in which the wind speed is too high ($u \geq u_{co}$) is about 1 %-2 %, and the model underestimates this by approximately 0.6 percentage points. The model's overestimation of the number of hours with winds that are too weak to
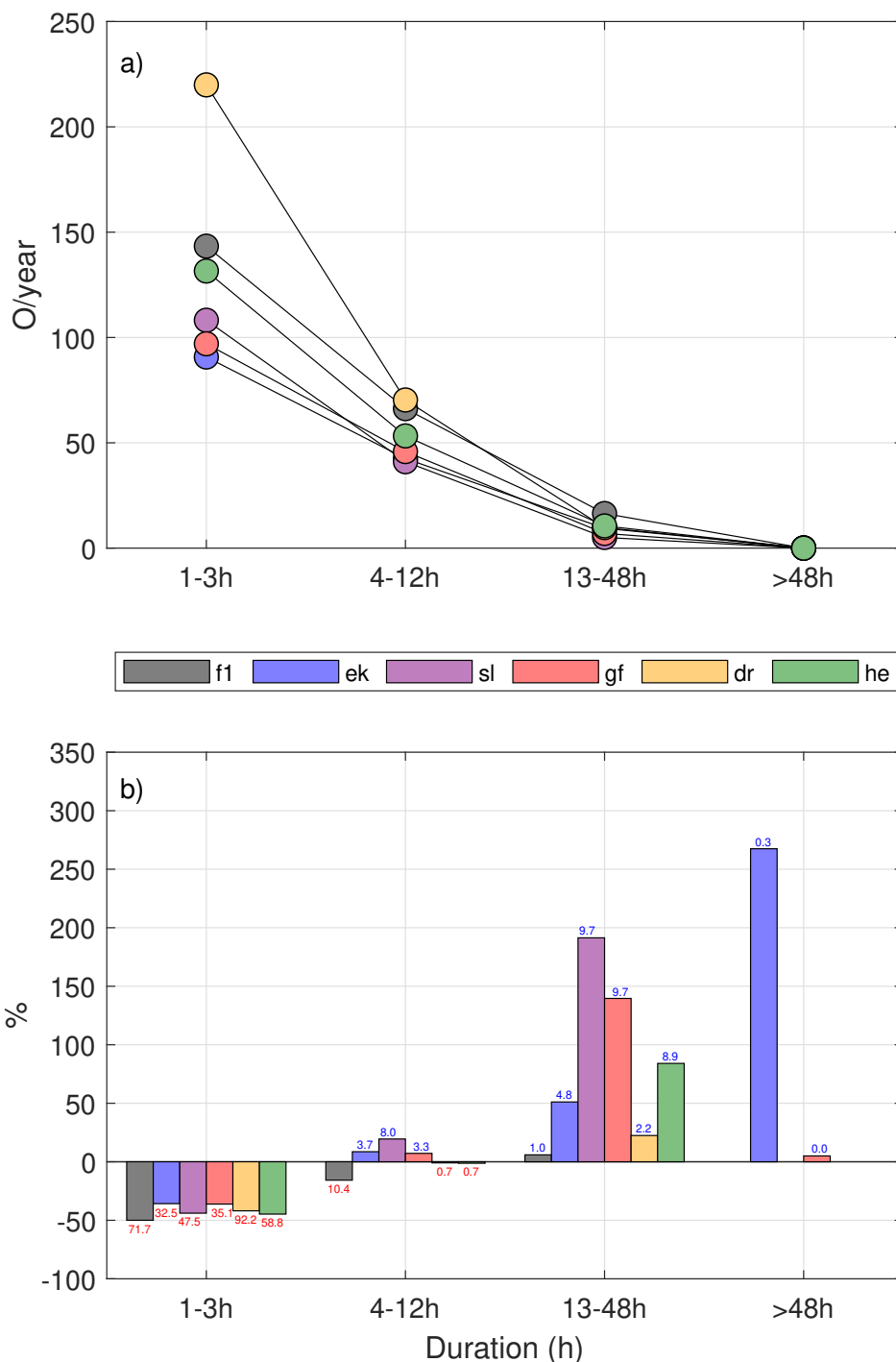
Figure 8. a) Yearly occurrence and corresponding duration of observation-based zero-events caused by to wind speeds lower than cut-in wind speed ($u < u_{ci}$). b) The differences between model-based (n3) and observation-based (obs) zero-event occurrences divided by the total number of observed occurrences of "too low" wind speeds. Values over or under each bar correspond to the differences (n3-obs) in the number of yearly occurrences between the model and observations. Abbreviations: f1: Fino1; ek: Ekofisk; sl: Sleipner; gf: Gullfaks C; dr: Draugen; he: Heidrun.
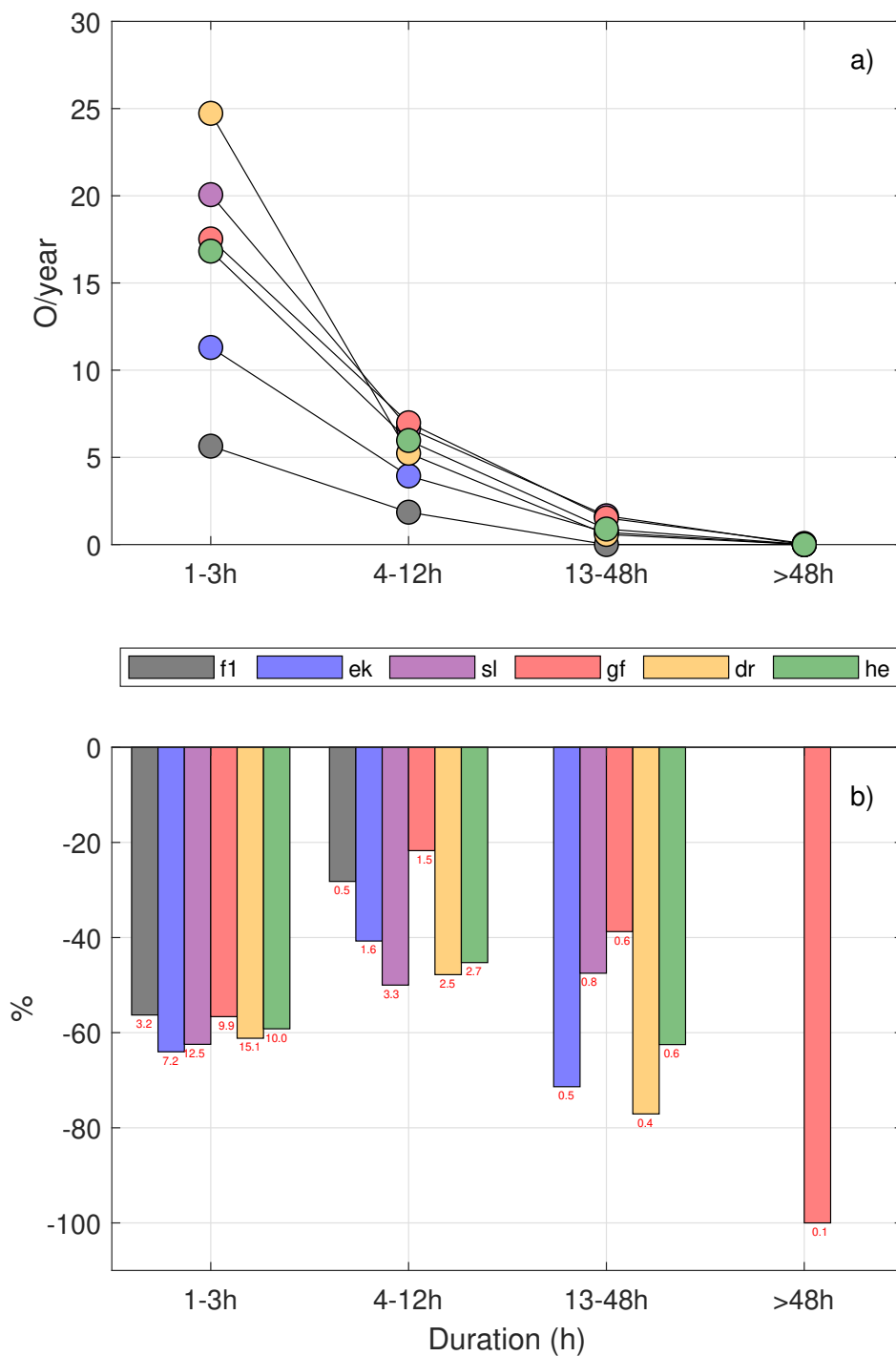
Figure 9. a) Yearly occurrence and corresponding duration of observation-based zero-events caused by to wind speeds higher than cut-out wind speed ($u \geq u_{co}$). b) The differences between model-based (n3) and observation-based (obs) zero-event occurrences divided by the total number of observed occurrences of "too high" wind speeds. Values over or under each bar correspond to the differences (n3-obs) in the number of yearly occurrences between the model and observations. Abbreviations: f1: Fino1; ek: Ekofisk; sl: Sleipner; gf: Gullfaks C; dr: Draugen; he: Heidrun.

produce wind power and its underestimation of the number of hours with winds that are too strong results in a well-captured

335 total numbers of hours of zero wind power production, which differs from the observed value by 0.8 percentage point.

The atmospheric conditions causing winds that are too weak for wind power production are very different from those causing winds that are too strong. Therefore, we split the zero-events accordingly. Figs. 8 and 9 illustrate the ability of the NORA3 to capture the observation-based estimates of zero-events of different durations. Fig. 8a shows the observation-based numbers of zero-events of varying durations caused by too weak winds. As expected, the number of zero-events decreases as the duration

340 of the events increases, ranging from around 90-150 yearly events lasting less than 3 h for most sites to close to zero such events lasting longer than 2 days. Figure 8b graphs the relative differences (in percentage) between the NORA3 and observation-based estimates of the numbers of zero-events by duration. The model-based estimates typically have 40 %-50 % too few zero-events of short duration (1-3 h) compared to the observations. For longer zero-events the model is biased towards too many events.

The model's underestimation of short zero-events caused by too low wind speeds and its overestimation of longer zero-events

345 occur as a result of the model having lower variability than the observations, as seen in the ramp-rate analysis (see Section 4.1). This lower variability means that when these zero-events occur in the model they tends to be of longer duration, but the frequency of such events is too low.

From Fig. 9a it is evident that the yearly average occurrence of zero-events caused by too strong winds is a factor of ten lower than the number of zero-events caused by winds that are too weak. Hence, one zero-event caused by too strong winds

350 happens for approximately every 10 zero-events caused by too weak winds. The model underestimates the number of zero-events caused by too strong winds for all sites (Fig. 9b); depending on the zero-event duration, NORA3 typically has 40 %-70 % too few zero-events caused by too strong winds.

Table 8. The percentages of observed wind speeds (Obs) and modeled wind speeds (n3) that fall into the following four categories: 1) the wind speed is less than the cut-in limit ($u < u_{ci}$), 2) the wind speed interval in which the wind power is a function of the cube of the wind speed ($u_{ci} \leq u < u_r$), 3) wind power production is rated ($u_r \leq u < u_{co}$), and 4) wind speed exceeds the cut-out limit ($u_{co} \leq u$). In addition, the total hours of zero wind power production ($P_w = 0$) divided by the total number of observations, are shows as a percentage.

	$u < u_{ci}$		$u_{ci} \leq u < u_r$		$u_r \leq u < u_{co}$		$u \geq u_{co}$		$P_w = 0$ (%)	
Site	Obs	n3	Obs	n3	Obs	n3	Obs	n3	Obs	n3
Fino1	10.98	10.33	66.03	66.50	22.78	23.04	0.21	0.14	**11.19**	**10.47**
Ekofisk	8.88	10.42	63.45	64.87	27.07	24.42	0.60	0.29	**9.48**	**10.71**
Sleipner	7.52	10.62	54.08	57.67	36.48	30.85	1.92	0.86	**9.45**	**11.48**
Gullfaks C	8.96	11.52	53.98	55.72	35.07	31.53	2.00	1.23	**10.96**	**12.75**
Draugen	13.81	14.12	59.67	62.45	25.59	22.97	0.93	0.46	**14.75**	**14.58**
Heidrun	11.05	12.69	58.12	60.85	29.41	25.83	1.42	0.63	**12.52**	**13.32**
Average	**10.20**	**11.62**	**59.22**	**61.34**	**29.40**	**26.44**	**1.18**	**0.60**	**11.39**	**12.22**
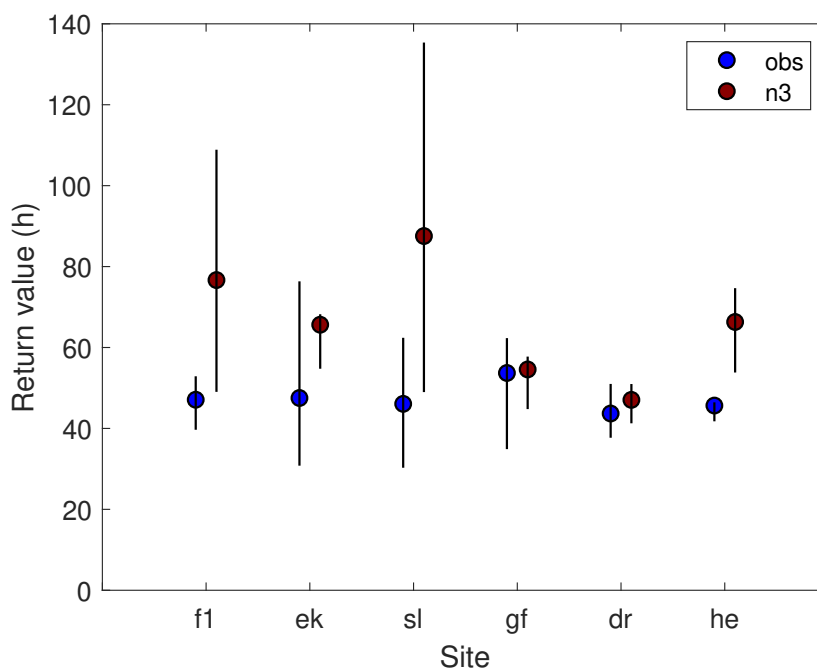
Figure 10. The 25-year return value with the corresponding confidence interval of the maximum duration of a zero-event generated by fitting a generalized Pareto distribution to the POT-data (peak over threshold) using both observations (obs) and modeled data (n3) for each of the sites. Abbreviations: f1: Fino1; ek: Ekofisk; sl: Sleipner; gf: Gullfaks C; Draugen: Dr; he: Heidrun.

4.5 Expected maximum zero-event duration over the turbine lifetime

In this section we attempt to validate the model's ability to provide reliable estimates of extremely long zero-events. This is
355 done by estimating the 25-year return value for the duration of a zero-event (the typical length of a zero-event that statistically would occur at least once over a 25-year period) using the method outlined in Section 2.6. Using the Kolmogorov-Smirnov test, we cannot exclude the possibility that the BM-data and POT-data are drawn from a GEV distribution and a Pareto distribution, respectively. Thus, it is reasonable to fit the observation-based and model-based extreme zero-event duration estimates to these distributions and find the 25-year maximum expected zero-event duration.

360   Figure 10 displays the results from fitting the Pareto distribution to the POT-data (the results fitting the BM-data to the GEV distribution are similar). From the observed data the typical length of the longest zero-event expected to occur at least once during the lifetime of a turbine is on the order of 40-60 h, but a zero-event of more than five days cannot be ruled out. The uncertainty in the estimations make it difficult to judge which sites have the shortest and longest maximum zero-even duration. Using the model data, the estimates are typically longer than the observation-based estimates (not significant at the 2.5 %
365 significance level for five of six sites); and are in line with the lower variability in the modeled hourly wind speed and wind power as seen in the ramp-rate analysis (see Sections 3.1 and 4.1). In conclusion, using NORA3 to estimate extreme zero-event

duration would lead to a conservative estimate of the return values, and the duration might be overestimated due to the lower variability in the model.

5   Summary

370   NORA3, a high-resolution (3-km) numerical mesoscale weather simulation data set from the Norwegian Meteorological Institute for the period 2004-2018 has been validated for offshore wind power purposes using observations of wind speed and direction from six offshore sites along the Norwegian continental shelf. Validation of mean quantities, hourly distributions, and variability have been carried out for wind speed and wind power. In addition, wind power estimates based on a selected hub height and power curve have been conducted using the observations and modeled wind speed. Comparisons between modeled
375   and observed data have been made for estimated wind power capacity factors; temporal variability; spatial co-variability between production sites; frequency, duration and total number of hours of wind power zero-events; and the maximum length of a zero-event expected to happen at least once during the turbine lifetime. The general picture is that the NORA3 data is rather well suited for wind power estimates in the absence of in situ data. There is a tendency for the modeled data to give slightly conservative estimates of the wind resources.

380   In five of the six offshore sites NORA3 seems to be biased towards lower mean wind speeds ($u_{obs} = 10.59 \text{ ms}^{-1}$, $u_{n3} = 10.04 \text{ ms}^{-1}$). The differences in wind speed distribution between the observations and the model output reveal that the model underestimates the number of events with wind speed exceeding the rated wind speed and overestimates the number of events with wind speeds below the rated wind speed (see Fig. 2). The transition between over- and underestimation by the model occurs near a typical rated wind speed ($11 - 13 \text{ ms}^{-1}$). As the model underestimates the wind episodes above the rated wind
385   speed, this partly counteracts the model's overestimation of low wind speeds, making the total modeled power production slightly underestimated.

NORA3 is also biased towards less variable wind speeds on hourly timescales. Analysis of hourly wind speed ramp rates show that the hour-to-hour variability is typically slightly above 1 ms^{-1} while the model-based ramp rates are slightly below 1 ms^{-1}, resulting in an underestimation of wind speed ramp rates on the order of 30% (see Table 4).

390   Generally, estimates of wind power from NORA3 are biased towards too low median values ($P_{w,obs} = 0.43, p_{w,n3} = 0.37$) and wind power CF ($CF_{obs} = 0.50, CF_{n3} = 0.47$). The negative bias is a consistent feature seen in all years and for all months for five of the six sites.

The wind power ramp-rate analysis shows that the hourly wind power variability of the model-based estimates are too low. The observation-based wind speed variability leads to a corresponding wind power ramp rate that is typically 0.08 (8% of
395   installed capacity), while the model-based ramp rate estimated is typically 0.05.

By interconnection of site-pairs we demonstrate that the spatial co-variability in estimated hourly wind power production between sites is slightly higher for the modeled data than for the observational data. Hence, the decorrelation length is estimated to be 19 % longer in the model-based estimates.

The estimation of the occurrence and duration of zero-events show a fairly well-captured total risk of hourly zero-events (n3
400 = 12.22 %, obs = 11.39 % of the time). We split the zero-events into episodes of no wind power production caused by either too
low ($u < u_{ci}$) or too high ($u \geq u_{co}$) wind speeds. For zero-events caused by winds that are too strong NORA3 underestimates
the occurrence of zero-events for all durations. For winds that are too weak, NORA3 underestimates the number of short zero-
events (1-3 h) but is biased towards an excess of zero-events with longer duration. As a result, when a zero-event occurs in the
model, it tends to be of longer duration, but the frequency of such events is too low. This deviation from the observation-based
405 zero-events is in line with the lower variability in hourly wind speeds seen in the ramp rate analysis (Section 3.1 and Section
4.1).

In the extreme-value analysis we found that at least once during the lifetime of a turbine (25 years) a zero-power event is
expected to last for 1 to 3 days, depending on the site in question (see Fig. 10). However, a zero-event lasting longer than 5 days
cannot be ruled out for some sites. Overall, the model is somewhat conservative, with a tendency towards longer maximum
410 zero-event duration than seen in the observation-based return values.

To a large degree the NORA3 wind speeds and directions resemble the characteristics seen in observations. This results in
rather good estimates of wind power and wind power variability. However, the model slightly underestimates the wind power
potential, and the hourly variability in the model output is lower than in the observations. These characteristics should be kept
in mind when using the NORA3 data set as an offshore wind power planning tool.

415 *Author contributions.* IMS and AS conceptualized the overarching research goals of the study, in addition to conducting formal analysis
regarding statistical and mathematical techniques and methods. HH is the creator of NORA3 data set and is responsible for the data resources.
IMS was responsible for the data curation and creation of software code, validation, visualization of data, and preparation of the original
draft, with contribution from AS. In addition, AS was responsible for the supervision. All the authors contributed to the review and editing
process of the paper.

Competing interests. No competing interests to declare.

Acknowledgements. Thanks to the Meteorological Institute, and a special thank to Magnar Reistad, Øyvind Breivik, and Ole Johan Aarnes
for providing all the observational data used in this study and for interesting discussions.
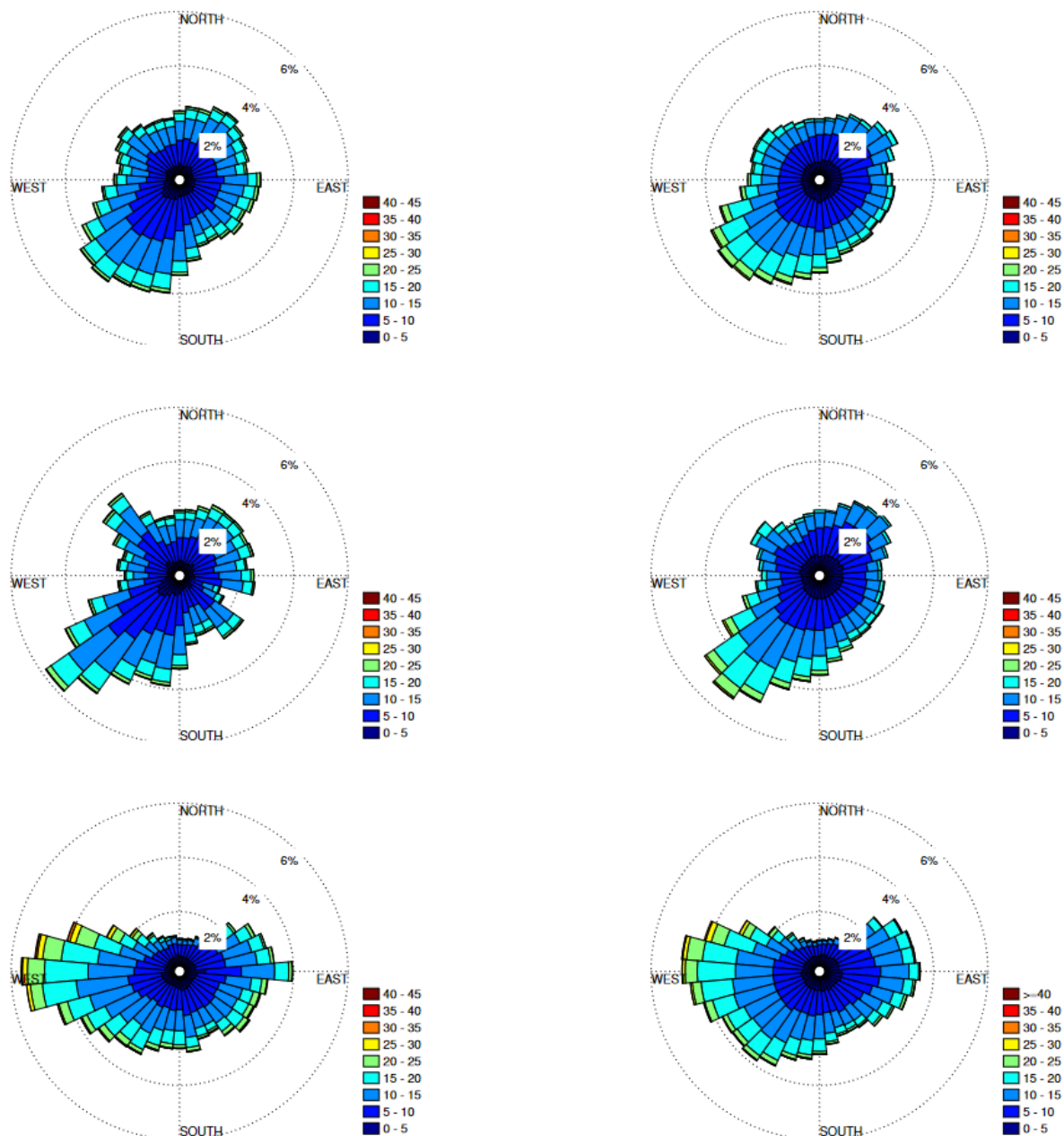
Figure A1.

Appendix A: Wind direction

See Fig. A1 for the observed and modelled wind rose plot.
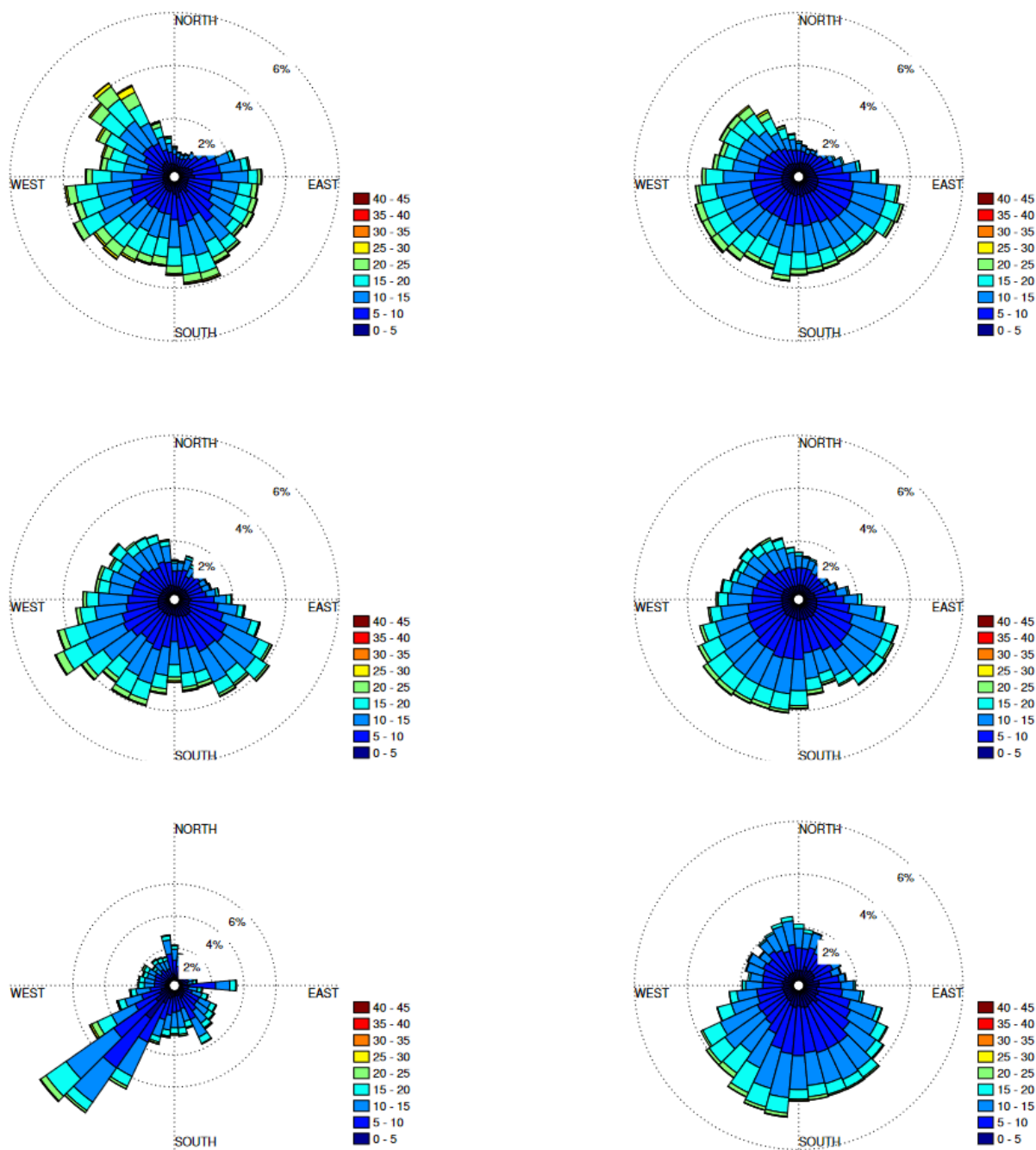
Figure A1. Observed (left column) and modeled (right column) wind roses for the six sites. The colorbars show the wind speed intervals in meters per second. The upper row corresponds to the northern most site, Heidrun, followed by Draugen, and the last row is southern-most site Fino1.

Appendix B: Flow distortion by large structures
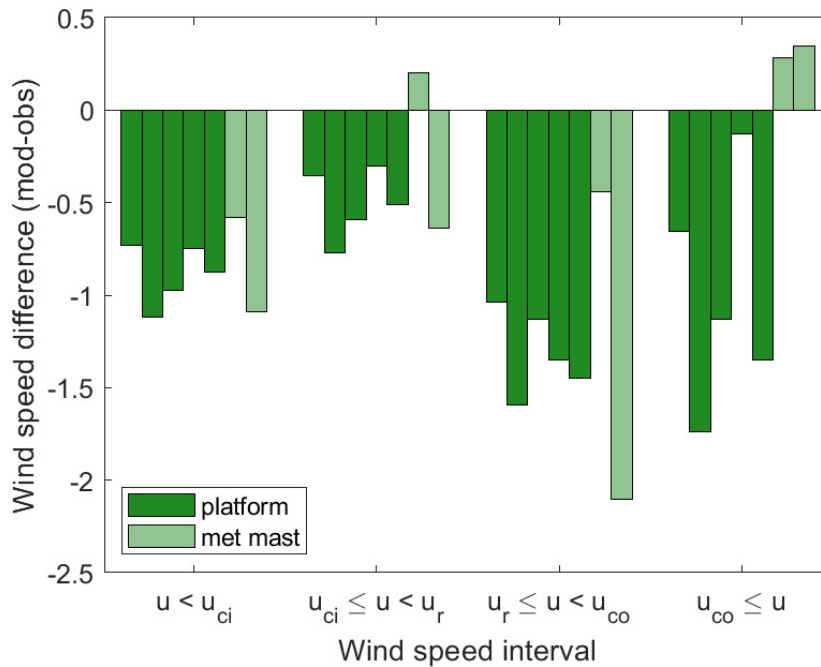
Figure B1. The wind speed difference (ms^{-1}) between the model and the observations (mod-obs) for the six sites. Dark green bars correspond to platform-sites (Ekofisk, Sleipner, Gullfaks C, Draugen and Heidrun) and light green to the met masts (Fino1 and Frøya).

To investigate a potential flow distortion effect in the observations caused by offshore platforms we sort the modeled wind speed data into categories according to the four wind power regimes: simulated wind speeds lower than cut-in wind speed ($u < u_{ci}$), simulated wind speeds falling between the cut-in and rated wind speed ($u_{ci} \leq u < u_r$), simulated wind speeds corresponding to nameplate capacity ($u_r \leq u < u_{co}$), and simulated wind speeds exceeding the cut-out limit ($u_{co} \leq u$). We then extract the observations for each hour that falls into each of the four categories, and determine the mean wind speeds for each category and for each site. The results are shown in Fig. B1. For the platform-sites, the observed wind speeds are greater than the modeled wind speeds in all four categories, with the largest difference seen in the rated wind regime ($u_r \leq u < u_{co}$). The finding that the observed wind speeds are greater than the modeled wind speeds for all four categories suggests a speed-up effect may be present for both weak and strong winds. For the two met masts, at Fino1 and Frøya, the model wind speed is weaker than the observed wind speed for the three first categories, but for the "too strong" wind regime ($u_{co} \leq u$) the modeled wind speed is greater than the observed wind speed.

The five platforms have different layouts and therefore influence the upcoming flow field differently. Therefore, we calculate the difference in mean wind speed (n3-obs) for each wind direction interval ($30°$-interval) (see Fig B2) to investigate a potential flow distortion caused by the platforms. For almost all wind direction intervals the observed mean wind speeds are higher than
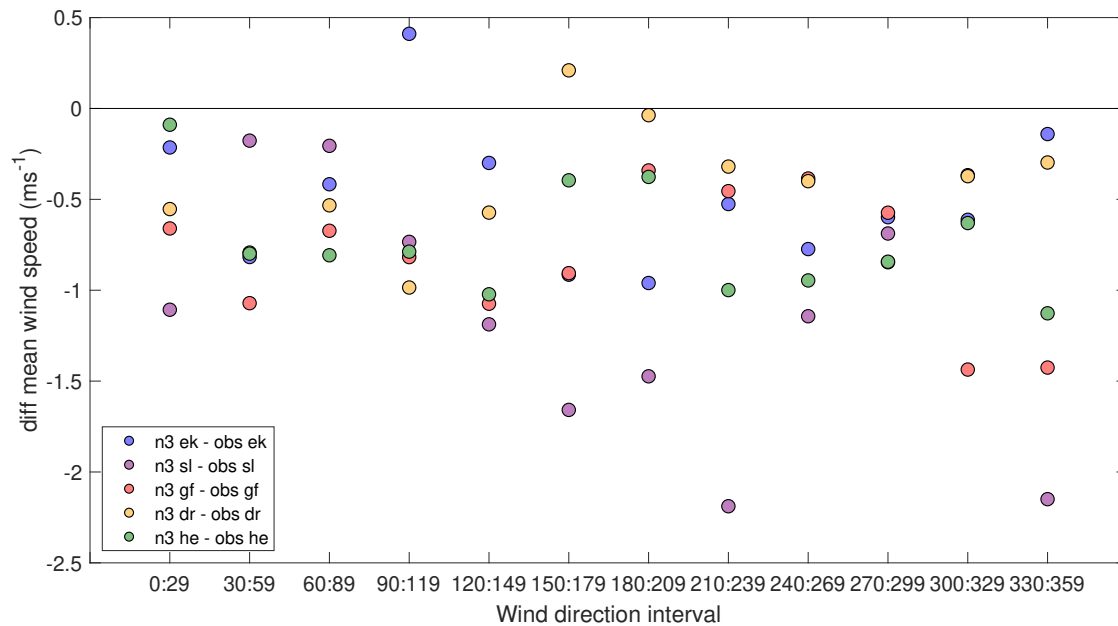
Figure B2. Differences in modeled (n3) and observed (obs) mean wind speeds (ms^{-1}) for each wind direction interval ($30°$-intervals). Abbreviations: ek: Ekofisk; sl: Sleipner; gf: Gullfaks C; dr: Draugen; he: Heidrun.

440   the modeled speeds. For the Sleipner site, wind directions between 150:240 and 330:30 result in differences between the modeled and observed mean wind speeds of -2.5 ms^{-1} to -1 ms^{-1}, indicating some structural disturbance in the flow field.

References

AG, S.: Siemens 6.0 MW Offshore Wind Turbine, Tech. rep., 2011.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., Rooy, W. d., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson,
445   K. I., Lenderink, G., Niemelä, S., Nielsen, K. P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D. S., Subias, A., Tijm, S., Toll, V.,
Yang, X., and Køltzow, M. : The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system, Monthly Weather
Review, 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2017.

Berge, E., Byrkjedal, , Ydersbond, Y., and Kindler, D.: Modelling of offshore wind resources. Comparison of a meso-scale model and
measurements from FINO 1 and North Sea oil rigs, in: European Wind Energy Conference and Exhibition 2009, EWEC 2009, vol. 4, pp.
450   2327–2334, 2009.

Byrkjedal, Ø. and Åkervik, E.: Vindkart for Norge, Tech. rep., Kjeller Vindteknikk, 2009.

Byrkjedal, Ø., Harstveit, K., Kravik, R., Løvholm, A. L., and Berge, E.: Analyser av offshore modellsimuleringer av vind, Tech. rep., Kjeller
Vindteknikk, 2010.

Cory, K. and Schwabe, P.: Wind Levelized Cost of Energy: A comparison of Technical and Financing Input Variables, Tech. rep., National
455   Renewable Energy Laboratory, 2009.

Dörenkämper, M., Olsen, B., Witha, B., Hahmann, A., Davis, N., Barcons, J., Ezber, Y., García-Bustamante, E., González-Rouco,
J. F., Navarro, J., Sastre-Marugán, M., Sīle, T., Trei, W., Žagar, M., Badger, J., Gottschall, J., Sanz Rodrigo, J., and Mann, J.: The
Making of the New European Wind Atlas – Part 2: Production and Evaluation, Geoscientific Model Development, 13, 5079–5102,
https://doi.org/10.5194/gmd-13-5079-2020, 2020.

460 Emeis, S.: Wind Energy Meteorology, Springer, 2 edn., 2018.

Furevik, B. R. and Haakenstad, H.: Near-surface marine wind profiles from rawinsonde and NORA10 hindcast, Journal of Geophysical
Research Atmospheres, 117, 1–14, https://doi.org/10.1029/2012JD018523, 2012.

Haakenstad, H., Breivik, Ø., Furevik, B. R., Reistad, M., Bohlinger, P., and Aarsnes, O. J.: NORA3: A non-hydrostatic high-resolution
hindcast for the North Sea, the Norwegian Sea and the Barents Sea, Submitted to Journal of Applied Meteorology and Climatology,
465   https://folk.uib.no/iso056/Paper_NORA3_Haakenstad_etal.pdf, 2021.

Hasager, C. B., Hahmann, A. N., Ahsbahs, T., Karagali, I., Sile, T., Badger, M., and Mann, J.: Europe's offshore winds assessed with synthetic
aperture radar, ASCAT and WRF, Wind Energy Science, https://doi.org/10.5194/wes-5-375-2020, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Sim-
mons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren,
470   P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J.,
Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Vil-
laume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049,
https://doi.org/10.1002/qj.3803, 2020.

Kempton, W., Pimenta, F. M., Veron, D. E., and Colle, B. A.: Electric power from offshore wind via synoptic-scale interconnection, Proceed-
475   ings of the National Academy of Sciences of the United States of America, 107, 7240–7245, https://doi.org/10.1073/pnas.0909075107,
2010.

Milan, P., Morales, A., Wächter, M., and Peinke, J.: Wind Energy: A Turbulent, Intermittent Resource, pp. 73–78,
https://doi.org/10.1007/978-3-642-54696-9_11, 2014.

NEWA: The New European Wind Atlas, https://www.neweuropeanwindatlas.eu/mission, 2020.

480 Regjeringen: Opening sites at the Norwegian continental shelf for wind power production, https://www.regjeringen.no/no/aktuelt/
     opner-omrader/id2705986/, 2020.

Reichenberg, L., Johnsson, F., and Odenberger, M.: Dampening variations in wind power generation-The effect of optimizing geographic
     location of generating sites, Wind Energy, 17, 1631–1643, https://doi.org/10.1002/we.1657, 2014.

Reichenberg, L., Wojciechowski, A., Hedenus, F., and Johnsson, F.: Geographic aggregation of wind power—an optimization methodology
485 for avoiding low outputs, Wind Energy, 20, 19–32, https://doi.org/10.1002/we.1987, 2017.

Seity, Y., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective-scale operational model,
     Monthly Weather Review, 139, 976–991, https://doi.org/10.1175/2010MWR3425.1, 2011.

Skeie, P., Steinskog, D. J., and Näs, J.: Kraftproduksjon og vindforhold, Tech. rep., StormGeo AS, 2012.

Smith, E. P.: An Introduction to Statistical Modeling of Extreme Values, https://doi.org/10.1198/tech.2002.s73, 2002.

490 Solbrekke, I. M., Kvamstø, N. G., and Sorteberg, A.: Mitigation of offshore wind power intermittency by interconnection of production sites,
     Wind Energy Science, 5, 1663–1678, https://doi.org/10.5194/wes-5-1663-2020, 2020.

St. Martin, C. M., Lundquist, J. K., and Handschy, M. A.: Variability of interconnected wind plants: Correlation length and its dependence
     on variability time scale, Environmental Research Letters, 10, 44 004, https://doi.org/10.1088/1748-9326/10/4/044004, http://dx.doi.org/
     10.1088/1748-9326/10/4/044004, 2015.

495 Vasilyev, L., Christakos, K., and Hannafious, B.: Treating wind measurements influenced by offshore structures with CFD methods, in:
     Energy Procedia, https://doi.org/10.1016/j.egypro.2015.11.425, 2015.

Wiser, R., Jemmi, K., Seel, J., Baker, E., Hand, M., Lantz, E., and Smith, A.: Forecasting Wind Energy Costs and Cost Drivers: the
     Views of the World' S Leading Experts, Tech. rep., IEA Wind, NREL, Berkeley Lab, file:///C:/Users/iso056/AppData/Local/Temp/
     Ade26H1RQxa0ee2avOyS_lbnl-1005717.pdf, 2016.

500 Zheng, C. W., Li, C. Y., Pan, J., Liu, M. Y., and Xia, L. L.: An overview of global ocean wind energy resource evaluations, Renewable
     and Sustainable Energy Reviews, 53, 1240–1251, https://doi.org/10.1016/j.rser.2015.09.063, http://dx.doi.org/10.1016/j.rser.2015.09.063,
     2016.