Anonymous Referee #1

Dear reviewer,

We are grateful for your careful work and appreciate your detailed comments. Below you will find details how we will revise our manuscript based on your review.

With best regards,

Jennie Molinder (née Persson Söderman)

General comments: This article looked at applying two well-known probabilistic techniques from the atmospheric science community to the problem of production loss of wind turbines due to icing. The methods were applied on the NWP model simulations that were used as input to the icing model, which in turn is used as input to the power production model. The paper shows that a traditional model ensemble does not improve results when tested over two wind farms for a two week period, but that using forecasts from neighboring points does improve the forecast.

As discussed in the previous AC, both methods do improve the forecast skill. The referee also recognized this after our previous correspondence.

The article is well structured and the topic is of significant interest to the Wind Energy community. However, the methods are not described in enough detail for reproduction, and then observational data that is used is not described sufficiently nor of sufficient length for it to provide a good validation of the methods being tested.

Methods and observations will be described in more detail including
  - HarmonEPS setup and initialization – See answer to specific comment 1 & 2 below

- The icing model formulation including the effect of different water components types and ice loss - See answer to specific comment 5 & 6 below
- Table of observations with approximate location, mast or nacelle, observed parameter – Table 1 added to manuscript
- And more details according to your specific comments below.

We agree that the period is rather short, which is why we stressed the case study character of our work. The limitation resulted from the fact that global EPS data is required on the boundaries. These data are not operationally archived by ECMWF, only selected periods are available. We picked the most relevant period with available production observations. During this period, one of the three stations is containing two icing episodes and that, when combining all days and stations with icing, we have about 17 days of icing for validation. Given the scarcity of production observations and high computational costs of EPS runs, we still think that the results are robust for these meteorological conditions showing the possible benefit of using probabilistic forecasting for icing related production loss forecasts.

The extension of the dataset is not impossible, but not a simple task due to the need of global boundary data, high computational costs in running and storing the high-resolution ensemble and scarcity of icing and production observations. If regarded necessary, we are willing to find a solution.
We welcome guidance to this question by editor and reviewers.

Finally, the paper mentions that the cloud parameters LWC and MVD are key inputs to the icing model, but they are not analyzed in terms of the spread of the probabilistic approaches, which seems like an important metric to analyze.

A discussion about the mean and standard deviation for LWC and MVD will be added to the revised manuscript. Also addressing your specific comment below on the different water components, we will add a discussion on the impact of all water components and of the MVD and their spread to the results on the example of one icing event.

Added in the Section 3.2.2: Figure 11 with ENS mean MVD and mean spread for Site B. Since there was nearly no rain during the period, the MVD for the liquid water components follows the amount of LWC seen in figure 9c, and the spread was similar during the two weeks, only the figure showing MVD for the

solid water components was added.

Discussion about the MVD:

Added in section 3.2.1: During the icing event however, the ensemble members have a spread of around 25 % of the ENS mean value for liquid and solid water components, leading to a spread in the amount of built-up ice (Fig. 6c).

Added in Section 3.2.2: In Fig. 11a the ENS mean MVD for the solid water components can be seen for this site and the corresponding mean spread is shown in Fig. 11b. The liquid water components ENS mean MVD is not shown, as it follows the LWC amount in this case (Fig. 9c). Large values of MVD (Fig. 11a) coincide with large icing rates (Fig. 9a) stressing the role of MVD for the icing intensity. The large spread of the MVD around January 4 (Fig. 11b) can be connected to the ice load differences of the different ensemble members (Fig. 9a) and also to the simultaneous spread in the production loss forecast (Fig. 10b). This behaviour agrees with the effect of different MVD discussed in Davis et al. (2014).

Added in Section 3.2.2: This variation of the ensemble members is probably due to a large variation of the liquid and solid water content (between 25 and 50 % of the mean amount) resulting in a variation of the calculated MVD.

Specific comments:

1. Section 2.1.1: How does the HarmonEPS used in this study differ from the operational usage of SMHI?

1: The HarmonEPS currently used at SMHI are using a new model version cy40h1 and lagged ensemble members, but it should be noted that the operational version also includes physics perturbations. We will add a reference to the operational setup by Andrae et al. (2017, ALADIN-HIRLAM Newsletter No. 8, available from http://www.umr-

Added to Section 2.1.1: For the generation of the initial conditions for the ensemble members, the so-called PERTANA option is used, where the difference between the control analyses of HarmonEPS and ECMWF is added to the fields from the ECMWF EPS perturbed members. The HarmonEPS setup used here differs from the operational HarmonEPS currently running at SMHI in several aspects. The operational version uses a new model version (cy40h1.1), boundary conditions with the Scaled Lagged Average Forecast method, and also some physics perturbations (Andre and MetCoOp-Team, 2017).

2. P4 L23-24: Do I read correctly that the models are all initialized using the ECMWF control member analysis, and only differ due to boundary conditions?

2: We understand that our description on the initial conditions has to be clarified. In fact, we use the so-called PERTANA option, where the initial fields are combining the HarmonEPS control analysis with the fields from the ECMWF EPS perturbed members. We will revise our description.

Deleted in Section 2.1.1: and the perturbations for each ensemble member come from the boundary conditions of different members of the ECMWF EPS that are added to the analysis of the control member."

Added instead in Section 2.1.1: For the generation of the initial conditions for the ensemble members, the so-called PERTANA option is used, where the difference between the control analyses of HarmonEPS and ECMWF is added to the fields from the ECMWF EPS perturbed members.

3. P5 L12: It would be interesting to see the impact of this averaging on the results. You state they are not here, but perhaps you could at least mention the improvement in error for the different variables in the text.

3: The vertical interpolation was introduced for an earlier study and not validated again for this study. The earlier vertical interpolation used adiabatic lifting, except for inversion cases where a linear interpolation was

used. This caused jumps in wind speed etc. during regime changes. Thus, this more general vertical interpolation was introduced. Although undoubtedly interesting, we think that a more detailed discussion is beyond the scope of this article and plan to shorten this passage in order to keep the focus on the probabilistic forecasting.

For your interest, we add a plot with the wind on model level plus nacelle height, model level plus topographic correction plus nacelle height and our vertical correction. The RMSE is smallest for our vertical interpolation scheme.
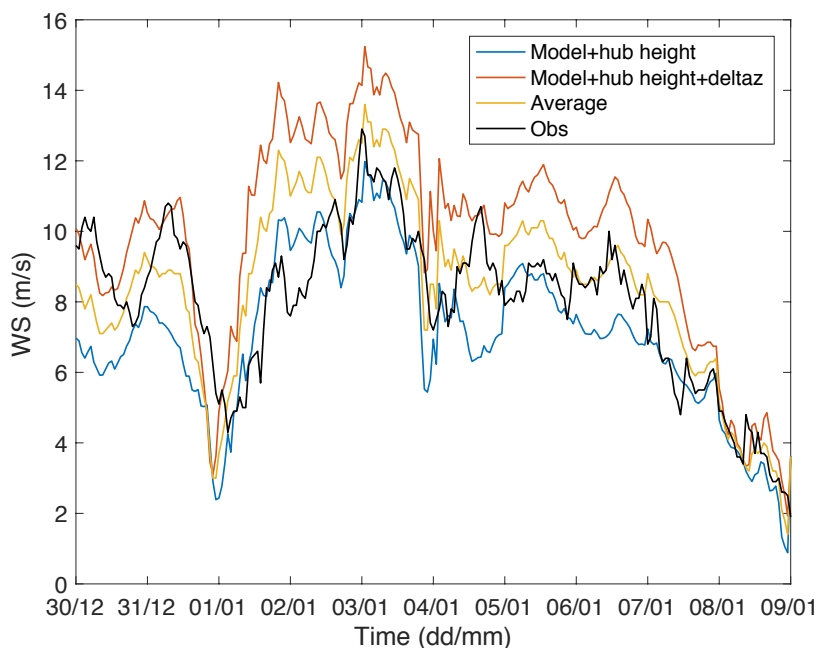

Figure 1. Wind speed at site A for the test period. Observations from nacelle are shown in black, modelled wind speed from model surface plus hub height in blue, modelled wind speed plus difference to real topography plus hub height in red, and the average of both modelled wind speeds in yellow.

Deleted in Section 2.1.2:

Adiabatic lifting between model terrain and real terrain has yielded unsatisfying results.

Deleted in Section 2.1.2:

As important parameters for wind power, such as wind, temperature and moisture, vary strongly with the distance from the surface, a realistic choice for the vertical interpolation is vital. By choosing the averaging method from

Eq. 1 for height interpolation these error sources were reduced (not shown here).

Added in Section 2.1.2: This averaging method was used since

4. P5 L12: What are all parameters, and which parameters was the lifting validated against?

4: All model parameters used in the icing model are corrected. We will add a clarification. No validation for this method is included here. We think that it could be done in a separate study.

Added in Section 2.1.2: all atmospheric parameters that serve as input to the icing model.

5. P6 L31-32: You note that different forms of water are fed separately through the model, how exactly is this carried out? Are you just running the model 4 times and summing the results? It is unclear to me how that would impact the accretion efficiency, which is a heat balance that depends on the mass flux of water impacting the structure.

5: Yes, we are running the model several times and summing the results. For simplicity, we have treated all water components in the same way for alpha3. This assumption can be investigated. However, we think that it is not crucial for the main focus of the paper, namely the application of probablistic forecasting for wind power in cold climate. We will add a clarification on the calculation of alpha3 for the different water components.

Added in Section 2.1.3: Extended Eq. 2 to include the summing of the results and an ice loss term.

Added in Section 2.1.3: For simplicity, the accretion efficiency α_3 is calculated in the same manner for the liquid and solid water components.

6. P7 L2: Is the MVD only calculated for cloud water?

6: The MVD is calculated for each of the water components since this is a necessary input to the icing model. In a revised version of the paper we will describe this more clearly.

Changed in Section 2.1.3: *Meteorological inputs needed for the ice calculations are temperature, wind speed, liquid water content, relative humidity and median volume droplet size. The latter is not directly available from the present NWP models, so a value is estimated using the liquid water content and the concentration of droplets.*

To in Section 2.1.3: Meteorological inputs needed for the ice calculations are temperature, wind speed, liquid/solid water content, relative humidity and median volume droplet size. The latter is not directly available from the present NWP models, so a value is estimated for all water components based on the liquid/solid water content and the concentration of droplets. The concentration of cloud droplets is set to a constant of 100 cm−1 except in the case of precipitation, when the number of droplets is instead based on output from the NWP model.

7. *P7 L4: What is the empirical ice shedding model?*

7: The ice shedding is basically a constant multiplied with the melting equation; this will be described better in a revised version of the paper.

Added in Section 2.1.3: The ice shedding is simulated by multiplying a constant of 8 with the melting term, increasing the process of removing the ice by a factor 8.

8. *P7 L15: In not familiar with the term effect curve, is this just the power curve?*

8: Yes, we are sorry for using the direct Swedish translation. This will be corrected.

Corrected

9. *P11 L26: You mention that there could be an issue with warm turbines, but you mentioned that you have a mixture of mast and nacelle data. Did you investigate if the bias was different between the two sources?*

9: The bias was not yet investigated in detail. We will examine the bias for

mast and nacelle data separately and add a remark to the manuscript.

Added in Section 3.1: This behavior might be caused by a spin-up problem in the model, since it is changing with forecast length. Additionally, a negative temperature bias that is not changing with forecast length can be attributed to the warm turbine affecting the measurements. A difference in the temperature bias between Mast and WT measurements amounts on average to -0.4 $°C$ for mast and -1.1 $°C$ for WT.

10. *P12 L10: could you list the % improvement of the wind speed and RH?*

10: Yes, the improvements for wind speed and RH is 7 and 12% respectively using ENSngb. This will be added in the article.

Deleted in Section 3.1: Similar behavior was found for the forecasted wind speed and relative humidity (not shown).

Added in Section 3.1: 7 % for the wind speed and 12 % for the relative humidity forecast error averaged for all sites and all forecasts

11. *P13 L9: The shift from 10 sites to 3 was confusing when reading this paper, perhaps remind the reader that you only had production data at 4 sites, and had to discard one. I am not sure why you only looked at two of the three sites here though.*

11: Thank you for the comment; we will describe this more clearly in a revised version of the paper.

Added in Section 2.2: Table 1. Observation sites with approximate latitude, description if the measurements were made from a mast or at the wind turbine nacelle (WT) and available production data (x).

Changed in Section 3.2: *Results from two of the three sites, here A and B, with consistent production observations will be presented in more detail here, and the average skill and spread for the different approaches for the*

three sites are discussed.

To in Section 3.2: Results from two of the three sites, here A and B, with consistent production observations will be presented in more detail in the following to point out some interesting features when using the probabilistic forecasts. Site C is not shown, but had icing during about three days of the two weeks. In Section 3.2.3, the average skill and spread for the different approaches for the three sites are discussed. Site A and C was during the two weeks affected by a frontal passage, while site A more experienced general cloudiness and was less affected by this front.

12. Section 3.2: You discuss quite often the LWC values and how they impact the ice growth. It is unclear to me what this parameter is, since your model includes two types of liquid water hydrometeors and two solid hydrometeors. Could you describe what it is, and if it doesn't already make sure it includes all of the relevant hydrometeors.

12: The LWC includes in this section refers to the liquid water content and not the solid parts. In our icing model, liquid water is a necessary condition for ice formation.

As you point out, it would be valuable to include all hydrometeors in the discussion. In a revised version of the manuscript we will, also according to your general comment above, add some discussion about the water components and related spread.

Changed: Figure 6c and 9c is changed from showing only the LWC to show both the total liquid water components and solid water components. In section 3.2.1 and 3.2.2 the different water components are discussed in more detail.

13. Section 3.2: In the experiment period and available data section, you mention that you have production data from each turbine at each site. It is unclear how you aggregated the production data to get a singular value. Can you state approximately how many turbines were at each of the two sites you used?

13. Yes, we will describe the aggregation of the production data and give an approximate number of wind turbines together with more detailed description of the sites in a revised manuscript.

Added in Section 3.2: From three of the sites, production data from each turbine is available for the period. The approximate location of the sites can be seen in Table 1 together with a specification if mast or nacelle (WT) observations were available and if the site had production data. The three sites with production data are at some distance from each other (Table 1). Site A is located on a hill with relatively high terrain west and north of the site, and somewhat lower terrain toward the south and east. In the location of site B there is instead similar terrain height to the south and west, while the terrain to the north and east is lower. The hill of site C extends mainly in the south-north direction, with lower terrain to the west and east. Site B and C have around 10 wind turbines, while site A consists of 20 turbines. All sites are surrounded with forest and with some lakes at lower levels. For each site one single value of production data is calculated by averaging production data from wind turbines without an error code. No de-icing system was used on the wind turbines included in the study.

Deleted in Section 3.2: For one of the four sites with production data, we found that wind speed measurements were unreliable, as the instrument was heavily affected by icing. This influence could be confirmed in webcam pictures. Thus, this stations wind speed data and its production data was omitted.

14. While it is understandable that you cannot list the wind farms themselves, can you at least describe how far apart they are, and if they are exposed to similar weather patterns.

14. We will add a characterisation and approximate location for each wind farm.

Added in Section 2.2: Table 1 with approximate location with a description

of the sites in the text

Added in Section 3.2: Sites A and C were affected by a frontal passage during the two weeks, while site A experienced general cloudiness and was less influenced by this front.

15. *Fig 6. It is hard to see the observations clearly, perhaps you could use a color like red that would stand out more.*

15: Thank you, we will make the figures clearer.

Changed: Red color for observations, grey color for CM instead of green and more relevant y-axis for the ice load figures.

16. *Figure 9: Why was there no ice during the beginning of the period, even though there was still a fairly large amount of LWC?*

16: Thank you for this observation! A quick analysis of the relevant meteorological parameters suggest this is because of very low wind speeds around the 29/12. We will examine this episode more carefully and add a comment about this in the revised manuscript.

Added in Section 3.2.2: In the beginning of the two weeks, there is one day with a large amount of liquid and solid water content, however, no ice started to build up during this day, mainly because of too low wind speeds.

17. *Table 1: What is the mean RMSE the mean of?*

17: It is the mean of the three sites. We will clarify this.

Removed in figure Table description: Mean

Added in table description: averaged over the three sites with production data.

Anonymous Referee #2

Dear reviewer,

We are grateful for your work and appreciate your comments. Below you will find details how we will revise our manuscript based on your review.

With best regards,

Jennie Molinder (née Persson Söderman)

General comments 1. First of all I would like to state that it is an important subject with a clear application and obvious benefits. 2. How much new stuff is offered here? Neighbourhood method was suggested by Mittermaier [2014] but perhaps not applied. Using ensembles for short-term prediction of icing: To my knowledge this is new.

3. It is a weakness that only a very short time period studied: 2 sites in 2 weeks. One site has two icing events; the other one. So all in all the authors are presenting and discussing three (3) icing events and comparing those with an ensembles modelling. This is a somewhat weak evidence.

3: We agree that the period is rather short, which is why we stressed the case study character of our work. The limitation resulted from the fact that global EPS data is required on the boundaries. These data are not operationally archived by ECMWF, only selected periods are available. We picked the most relevant period with available production observations. During this period, one of the three stations is containing two icing episodes and that, when combining all days and stations with icing, we have about 17 days of icing for validation. Given the scarcity of production observations and high computational costs of EPS runs, we still think that the results are robust for these meteorological conditions showing the possible benefit of using probabilistic forecasting for icing related production loss forecasts.

The extension of the dataset is not impossible, but not a simple task due to the need of global boundary data, high computational costs in running and storing the high-resolution ensemble and scarcity of icing and production observations. If regarded necessary, we are willing to find a solution. We welcome guidance to this question by editor and reviewers.

4. There is a short discussion at the end of paper. However, in order to judge the practical feasibility of using ensembles for ice prediction I miss more details on the calculations e.g. time elapsed for model runs and a discussion of whether it is feasible using the present approach or is likely to become feasible using input from National weather services. By feasible I am thinking on feasible for owners of wind farms to e.g. obtain that service from specialized consultants.

4: Thank you for the comment, we will include a discussion on feasibility and delivery times in the revised version. Since the introduction of the European INSPIRE directive, more and more national weather services are providing their forecasts as open data. Given the high computational costs of EPS, it seems beneficial, when national weather service can provide high-resolution EPS data for wind farm owners and specialized consults.

Changed in Section 4: *Finally, a NWP-ensemble forecast is computationally expensive to run and requires extensive infrastructure. Thus, often only national meteorological services can operationally produce ensemble forecasts. The open-data policy following the European INSPIRE directive will make these kinds of forecasts openly available. Then many users to include this forecast data into their own icing and production models which are far less computationally expensive. This will allow for a wide application of probabilistic forecasting for wind power in cold climates.*

To in Section 4: Finally, a NWP-ensemble forecast is computationally expensive to run and requires extensive infrastructure. Thus, often only national meteorological services can operationally produce ensemble forecasts. Many weather services currently run an operational ensemble forecast and the open-data policy following the European INSPIRE directive Many national weather ser- vices run currently an operational ensemble forecast that is often disseminated as open data following the European INSPIRE directive. The delivery time for such data are around 3 hours after analysis time, thus making it possible to use the 06-UTC model run for next-

day wind power forecasts. Many users will then be able to include this forecast data into their own icing and production models, which are far less computationally expensive. This will allow for a wide application of probabilistic weather forecasting for wind power in cold climates.

Detailed comments 5. On page 8 l5 10 sites are mentioned, What happened to the 8 remaining? I miss a reason for eliminating these 8 sites

5: We have 10 sites with meteorological observations, but only three sites with production data. We apologize for the confusion in the description and will describe this in more detail in a revised version.

Added: Table 1 showing more clearly that three of the sites have production data.

Added in Section 2.2: From three of the sites, production data from each turbine is available for the period.

Changed in Section 3.2: *Results from two of the three sites, here A and B, with consistent production observations will be presented in more detail here, and the average skill and spread for the different approaches for the three sites are discussed.*

To in Section 3.2: Results from two of the three sites, here A and B, with consistent production observations will be presented in more detail in the following to point out some interesting features when using the probabilistic forecasts. Site C is not shown, but had icing during about three days of the two weeks. In Section 3.2.3, the average skill and spread for the different approaches for the three sites are discussed. Sites A and C were affected by a frontal passage during the two weeks, while site A experienced general cloudiness and was less influenced by this front.

6. The neighbourhood principle is based on either following the landscape or using same height. I suppose you know where the sites are and might have added a short (anonymized) description of the landscape without revealing the location of the site?

6: Yes, we will add a characterisation and approximate location for each wind farm.

Added to section 2.2: The approximate location of the sites can be seen in Table 1 together with a specification if mast or nacelle (WT) observations were available and if the site had production data. The three sites with production data are at some distance from each other (Table 1). Site A is located on a hill with relatively high terrain west and north of the site, and somewhat lower terrain toward the south and east. In the location of site B there is instead similar terrain height to the south and west, while the terrain to the north and east is lower. The hill of site C extends mainly in the south-north direction, with lower terrain to the west and east. Site B and C have around 10 wind turbines, while site A consists of 20 turbines. All sites are surrounded with forest and with some lakes at lower levels. For each site one single value of production data is calculated by averaging production data from wind turbines without an error code. No de-icing system was used on the wind turbines included in the study.

7. *Effect curve is termed power curve by many in the wind community.*

7: Ok thank you, we will change this.

Corrected

8. *It is not clearly stated whether de-icing equipment is included for the turbines studied.*

8: Thank you, it should be mentioned as you say that no de-icing system is used. We will add this to the manuscript.

Added to section 2.2: No de-icing system was used on the wind turbines included in the study.

Technical corrections 9. Figures are generally too small especially figures 4 to 10. 10. The choice of colour of the individual curves makes it hard to distinguish the curves especially figs 6, 9 and 11

9 & 10: Thank you, we will revise these aspects in the next version of the paper.

Changed: Now red color for observations, grey color for CM instead of green in figure 6,7,9 and 10. Red observation in figure 12 (old 11). More relevant y-axis for the ice load figures. Larger figures (we do not know how large the figures will be in a published version).

Probabilistic forecasting of wind power production losses in cold climates: A case study

Jennie Molinder[1], Heiner Körnich[2], Esbjörn Olsson[2], Hans Bergström[1], and Anna Sjöblom[1]

[1]Uppsala University, Uppsala, SWE

[2]Swedish Meteorological and Hydrological Institute, Norrköping, SWE

*Correspondence to:* Jennie Molinder (jennie.molinder@geo.uu.se)

Abstract.

The problem of icing on wind turbines in cold climates is addressed using probabilistic forecasting to improve next-day forecasts of icing and related production losses. A case study of probabilistic forecasts was generated for a two-week period. Uncertainties in initial and boundary conditions are represented with an ensemble forecasting system, while uncertainties in the spatial representation are included with a neighbourhood method. Using probabilistic forecasting instead of one single forecast was shown to improve the forecast skill of the ice-related production loss forecasts and hence the icing forecasts. The spread of the multiple forecasts can be used as an estimate of the forecast uncertainty and of the likelihood for icing and severe production losses. Best results, both in terms of forecast skill and forecasted uncertainty, were achieved using both the ensemble forecast and the neighbourhood method combined. This demonstrates that the application of probabilistic forecasting for wind power in cold climate can be valuable when planning next-day energy production, in the usage of de-icing systems, and for site safety.

1 Introduction

Wind power production in cold climates experiences significant problems with production losses because of icing. Icing on the turbine blades reduces the energy production due to change in the aerodynamic balance, generation of vibration and increased load (Kraj and Bibeau, 2010). Furthermore, site safety is an issue since falling ice poses a threat to the public and to maintenance. Despite these complications, a substantial part of the wind power production is located in cold climate regions. This geographical choice results from both the possible higher production in lower temperatures where the air is more dense than in warmer regions, and from the low population density, which reduces public safety risks and disturbance. According to the World Market Update 2012 (Wallenius et al., 2013), more than 24 % of the global wind energy capacity was located in cold climate regions at the end of 2012 and most of these turbines experience between light to heavy icing. In order to plan for
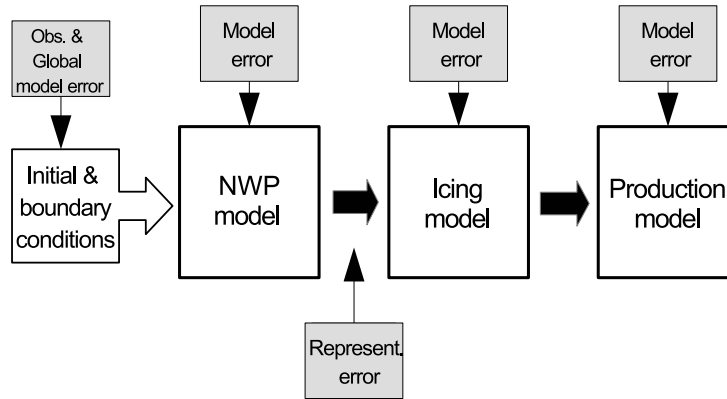
Figure 1. The modelling chain for forecasting icing related production losses. Uncertain parts are pointed out.

next-day energy production and site safety, short-range forecasts of icing and related production losses are vital tools for the energy market.

Forecasting icing and related production losses is challenging due to uncertainties in both the meteorological conditions and the modelling of the involved processes, e.g (Bergström et al., 2013; Davis et al., 2014). A common approach for the modelling chain is shown in Figure 1. A Numerical Weather Prediction (NWP) model is used to forecast meteorological parameters that serve as input to an icing model. Finally, a statistical production model calculates the icing-related production losses from the forecasted wind and icing. It should be noted that all steps in the modelling chain of Figure 1 contain uncertainties, either in the model formulation or the required input data. The NWP model and the icing model suffer from lack of knowledge about the physical processes, from the numerical discretization as well as from simplifications to make the models computationally affordable for operational usage (Yano et al., 2015). Initial conditions for the NWP model are also uncertain due to errors in the meteorological observations and assumptions in data assimilation methods (Megner et al., 2015). Forecasting wind power production requires high horizontal resolution in the order of kilometres to capture wind fields at 100 meters above ground in the Scandinavian mountains and also to model small-scale atmospheric phenomena leading to icing (Bergström et al., 2013). Due to the computational needs of such models, the domain size is limited and lateral boundary conditions are provided by a host model, adding further uncertainties in the modelling chain. Finally, a statistical production model is based on a limited set of previous forecast validations and on assumptions about the functional relationship between wind, ice and production, also resulting in uncertainties. Because of these error sources in the modeling chain, the forecasted production losses are uncertain.

This issue has been addressed by using different NWP models that result in different estimations of ice load and icing intensity, and hence production losses (Ronsten et al., 2012; Bergström et al., 2013). Here, we address this problem in another way.

A common approach for uncertainty quantification in NWP is to use ensemble forecasting (Leith, 1974). It is known that in a non-linear dynamical system such as the weather, the largest uncertainty results from initial errors growing rapidly with
5 forecast time. These errors can be represented by re-running the model multiple times, starting from slightly different initial conditions (Leutbecher and Palmer, 2008). This collection of forecasts is generated by an Ensemble Prediction System (EPS). Global EPS have been run since the early 90s at e.g. the European Centre of Medium-range Weather Forecasts (ECMWF). In the beginning the focus of ensemble forecasting lay mostly on medium-range global forecasting. In the last 10-15 years, meso-scale EPS has been developed at different weather centers (e.g. the Metoffice (Bowler et al., 2008) and NOAA (Du et al.,
10 2003)) addressing the uncertainties on the short range, i.e. during the first 48 hours of the forecast.

An ensemble forecast can be utilized in several ways. The ensemble mean generally has a lower error than a single forecast, because the less predictable parts have been filtered out when averaging the ensemble members (WMO, 2012). This method was used by Al-Yahyai et al. (2011) to estimate the average wind speed over Oman in different seasons, and the ensemble mean reduced the forecast error compared to a single forecast. The difference between, or the spread of, the ensemble members can
15 represent the uncertainty of the forecast. An ensemble forecast can also be used probabilistically to estimate the likelihood of a specific event, for example the timing of a sudden change in wind speed (Schäfer, 2014). Furthermore, EPS for short-range forecasting has been investigated for wind energy purpose, e.g. (Pinson and Kariniotakis, 2010; Traiteur et al., 2013). Traiteur et al. (2013) studied the use of short-range EPS for 1-hour wind speed forecasting and showed that a statistically calibrated EPS outperforms other forecast methods. Here, we will employ an EPS for forecasting icing-related wind power
20 losses.

An additional uncertainty arises due to the fact that kilometre-scale phenomena, such as convective clouds, have faster forecast error growth than phenomena on larger scales, such as the position of a low-pressure system. Thus, a forecasted small-scale cloud can be misplaced by some tens of kilometres generating an error in spatial representation. This uncertainty has implications for how the forecast can be interpreted at a specific wind turbine location. Mittermaier (2014) suggested the
25 use of the neighbourhood method in order to address this misplacement of small-scale features in forecasts. In this method, a selected number of the surrounding grid points to an observation site are treated as equally likely forecasts. These forecasts can then be used in the same way as an ensemble, accounting for the uncertainties in the representativeness at each wind turbine location. An approach to generate probabilistic forecasts of wind power called adapted resampling was already used earlier by Pinson and Kariniotakis (2010), where the value of probabilistic forecasting for trading and wind power management was
30 demonstrated and it was suggested that methods of wind power forecasting should not rely directly on point forecasts as input.

In this case study, probabilistic next-day forecasts for wind power in cold climates has been run for a two-week period in winter 2011/12. The modeling chain (Figure 1) is extended with meso-scale ensemble forecasts and the neighbourhood method. These extensions address the uncertainties in the initial and boundary conditions as well as the representation error of the NWP part. In order to examine the impact of these terms separately, different combinations of ensemble forecasting and the
35 neighbourhood method are examined as the uncertainty quantification of the forecast for icing and related production losses.

Thus, it will be investigated whether these probabilistic methods add value to specific challenges of wind power forecasting in cold climates.

The models in the modelling chain are described in Section 2.1. The specific experiment period and available observational data are described in Section 2.2. The different approaches for the uncertainty quantification are presented in Section 2.3 and the verification methods in Section 2.4. The results in Section 3 are divided into two parts: Meteorological parameters in Section 3.1 and forecasts of icing and production losses in Section 3.2. Concluding remarks are given in Section 4.

2 Method

2.1 Description of the Models

2.1.1 NWP model

As the NWP model, the Ensemble Prediction System HarmonEPS is used. HarmonEPS is a non-hydrostatic, convective permitting model intended for predictions of probabilities of high-impact weather events. It is based on the ALADIN-HIRLAM shared system and contains two packages of physical parameterizations, AROME and ALARO, of which the AROME package (cy38h1.2) was used here in the HARMONIE-AROME configuration (Bengtsson et al., 2017). HARMONIE-AROME has been used for operational weather forecasts at the Swedish Meteorological and Hydrological Institute (SMHI) since 2014 (Müller et al., 2017), also as an ensemble in HarmonEPS since 2016. In the present study, the horizontal resolution of the model is 2.5 km and it has 65 vertical levels. The model domain can be seen in Figure 2. The lateral boundary conditions come from the global EPS at ECMWF with an horizontal resolution of 30 km and are updated at 00 UTC and 12 UTC. A spinup period of three weeks were used to generate the start of the forecast period. The HarmonEPS ensemble consists of 10 perturbed members and one control member. The number of ensemble members were chosen based on a short-range EPS study, by Du et al. (1997), where it was shown that 8-10 ensemble members are sufficient for at least 90 % of the possible benefit of using an EPS. Since an EPS is computationally demanding to run, 10 members were therefore considered to be sufficient for the present study. The control member is using 3D-variational data assimilation of conventional observations as well as satellite observations from the instruments AMSU-A and AMSU-B, with 6 hour cycling. ~~and the perturbations for each ensemble member come from the boundary conditions of different members of the ECMWF EPS that are added to the analysis of the control member.~~ For the generation of the initial conditions for the ensemble members, the so-called PERTANA option is used, where the difference between the control analyses of HarmonEPS and ECMWF is added to the fields from the ECMWF EPS perturbed members. The HarmonEPS setup used here differs from the operational HarmonEPS currently running at SMHI in several aspects. The operational version uses a new model version (cy40h1.1), boundary conditions with the Scaled Lagged Average Forecast method, and also some physics perturbations (Andre and MetCoOp-Team, 2017). As the control member has no perturbations on the initial and boundary conditions, it should statistically outperform the other ensemble members.
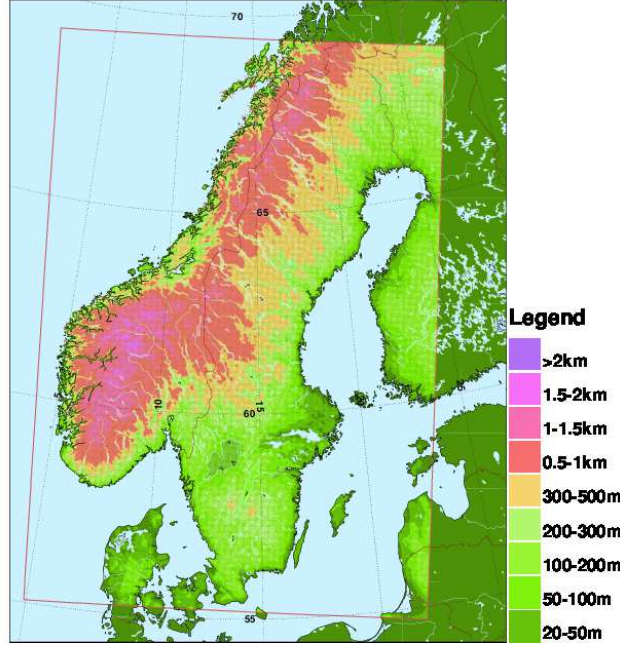
Figure 2. NWP model domain. Colours represent topography described in legend.

2.1.2 Postprocessing of the NWP data

Despite of increasingly higher resolution, the NWP models still lack some topographic details. The height of mountaintops in the model terrain remains in most cases below the actual height. The NWP output parameters are therefore adjusted to account for the difference between model terrain and real topography. ~~Adiabatic lifting between model terrain and real terrain has~~

5 ~~yielded unsatisfying results.~~ ~~Thus,~~ The following vertical interpolation is used for the NWP output; here for the example of temperature (T):

$$T_i = \frac{T(h_m + \Delta h + h_{nacelle}) + T(h_m + h_{nacelle})}{2} \tag{1}$$

where T_i is the vertically interpolated temperature, h_m is the model terrain height, Δh is the difference between the real terrain height and the model terrain height, and $h_{nacelle}$ is the height of turbine nacelle. Effectively, the forecast made at the actual

10 terrain height plus nacelle height, and the forecast made at the model terrain height plus nacelle height are averaged in Eq. 1. This averaging method was employed since using only the forecast data at the actual height over sea level for the wind turbine

5

nacelle can result in using atmospheric parameters well above the terrain compared to the turbine height. On the other hand, using only the forecast data at the height of the turbine nacelle above the model terrain can result in atmospheric data at lower height than the actual height of the wind turbine. ~~As important parameters for wind power, such as wind, temperature and moisture, vary st~~ ~~By choosing the averaging method from Eq. 1 for height interpolation these error sources were reduced (not shown here).~~ This interpolation is done for all atmospheric parameters that serve as input to the icing model. It should be noted that in the case where model terrain height is higher than the real terrain height, only the model terrain height is used. However, no such grid point was found in the current study.

As a bias in the forecasted meteorological parameters could be damaging to the results, bias correction was considered at each time step before using the parameters as input to the icing and production model. However, no such correction is made, since no reliable information on bias for each station at each time step could be derived.

The neighbourhood method following Mittermaier (2014) is used in order to capture the local uncertainty of the NWP data. Averaging forecasts made at several grid points around an observation site results in a better forecast than one single forecast from kilometre-scale NWP. Furthermore, the spread of the forecast from the neighbouring grid points provides an estimate of the forecast uncertainty. Here, the 25 nearest grid points to an observation site are chosen as equally likely forecasts.

Since these grid points are some kilometres apart from the turbine site, the height difference of the local topography can be several hundred meters. Two versions of the neighbourhood method was tested. In the first version, which is also the version presented in the result section (Section 3.2.3), the same height above sea level was used for all grid points, resulting occasionally in a height above ground much larger than the wind turbine height. The other version tested was a terrain-following method, where the same height above ground was used for all grid points. The different versions are further discussed in the result section.

2.1.3 Icing Model

The meteorological parameters forecasted by the NWP model are used to calculate ice loads utilizing a cylindrical ice accretion model, based on an equation often referred to as the Makkonen equation:

$$\frac{dM}{dt} = \sum_{i=1}^{5} \alpha_1 \alpha_2 \alpha_3 W v A - L \tag{2}$$

where M is the mass of ice, t is the time, the sum represent the sum of the ice accumulation due to the different hydrometeors rain, cloud water, snow, graupel and cloud ice, α_1, α_2 and α_3 are efficiency coefficients, W is liquid water content, v is wind speed and A is the cross-sectional area of the cylinder on which the ice accumulation is calculated (Makkonen, 2000) and L is the ice loss term which is an addition to the original Makkonen model. L includes functions for melting, sublimation, wind erosion and ice shedding. The efficiency coefficients take into account aspects of the object where the ice is accumulated, such as the possibility for water adhering to the surface. A detailed description of the coefficients is found in Makkonen (2000). Meteorological inputs needed for the ice calculations are temperature, wind speed, liquid/solid water content, relative humidity and median volume droplet size. The latter is not directly available from the present NWP models, so a value is estimated for

all water components based on the liquid/solid water content and the concentration of droplets. The concentration of cloud droplets is set to a constant of $100\ cm^{-1}$ except in the case of precipitation, when the number of droplets is instead based on output from the NWP model.

In addition to the original Makkonen equation, which only accounts for ice accretion due to cloud water, ice accretion due to cloud ice, graupel, snow and rainwater are included in the icing model. It is assumed that snow and graupel are only contributing to the ice accretion if rain or cloud water is also present since dry snow easily rebounces after the collision with the turbine. The sticking efficiency α_2 is different for snow and graupel compared to cloud water. Based on Nygaard et al. (2013) where both $\alpha_2 = 1/v^{0.5}$ and $\alpha_2 = 1/v$ were discussed, $\alpha_2 = 1/v^{0.75}$ is used here. For simplicity, the accretion efficiency α_3 is calculated in the same manner for the liquid and solid water components. The different forms of water in the cloud are fed separately into the equations using their forecasted concentrations from the NWP model. The equations for calculating droplet number concentrations for cloud ice, rain, snow and graupel have been taken from the AROME microphysics scheme (Seity et al., 2011). The median volume droplet diameter is calculated according to a scheme for cloud water by Thompson et al. (2008).

Formulas for melting, shedding, sublimation and wind erosion are also additions to the model compared to the original Makkonen equation. Melting of ice is calculated using an energy balance equation, which includes an empirical ice shedding. The ice shedding is simulated by multiplying a constant of 8 with the melting term, increasing the process of removing the ice by a factor 8. The sublimation is calculated using eq. 19 in Mazin et al. (2001) which utilizes wind speed and relative humidity in the calculations. The wind erosion is calculated by multiplying a hourly rate coefficent of $10\ gm^{-2}(ms^{-1})^{-1}$ with the wind speed when the wind speed is greater than $5\ ms^{-1}$, otherwise the erosion is zero. Here, only the wind speed at the nacelle is used. If the actual winds at the rotating wind turbine blade is used, the wind erosion coefficient needs to be reduced approximately by a factor of ten (Davis et al., 2016).

A more detailed documentation of the icing model can be found in Bergström et al. (2013). It should be noted that there are some differences in the model version since used in Bergström et al. (2013) as described above, i.e, the additional wind erosion calculations and the height interpolation.

2.1.4 Production Model

The production model consists of two parts, one part for the potential production, and one for the production loss.

Ice-free seasonally varying power curves were calculated for each wind turbine at every wind park, using a minimum of two years of production and wind speed observations. Only production observations with temperatures above $5°C$ were used to ensure that the blades are free of ice. The power curves are then used with forecasted wind speed to calculate the potential production.

The production loss forecast requires the modelled parameters of ice intensity, ice load and wind speed as input. It uses two matrices separating the losses due to ice load and icing intensity (Bergström et al., 2013). For a specific wind speed, ice load and icing intensity, the model yields a production loss in percent. Only one of the matrices is used for each forecast depending on which gives the highest production loss. The matrices were constructed manually using hindcasts of ice load and wind speed from a two-month period in 2010 combined with observed production values from one specific wind park which showed

7

good agreement between observed and modelled icing. Due to contractual reasons the wind park must not be specified. The matrices were generated by fitting 0 and 100 % losses against observations and then by interpolating the values inbetween. The empirical functions for production loss, determined for one specific wind farm, were used in the production forecasts for all wind farms. Generally, the icing intensity influences the production losses more than the ice load (Bergström et al., 2013).

5 Finally, the potential production is combined with the forecasted loss to provide the actual forecasted production output.

The observed production loss was calculated from the ratio between the observed production and the potential production, given the observed wind speed and the ice-free power curves. A value of 30 % means 30 % less production than the potential production.

2.2 Experiment period and available data

10 The case study is based on a two-week period. HarmonEPS was run 26 December 2011 to 7 January 2012, initializing a 42-hour long forecast at 00, 06, 12 and 18 UTC, with hourly output. The forecasted values from 18 to 42 hours starting from 06 UTC are used as the "next-day" forecast in the evaluation of the forecasted production losses and production.

Observations are available from 10 wind parks in Sweden that must not be specified due to contractual reasons. At some sites the observations were made from meteorological masts and at others at the turbine nacelle. All 10 sites measure temperature,

15 wind speed, wind direction, relative humidity, pressure and ice load. The sites are located from northern to southern Sweden between 250 and 1,000 m above sea level and the measurement height above ground is between 60 and 150 m. From three of the sites, production data from each turbine is available for the period. The approximate location of the sites can be seen in Table 1 together with a specification if mast or nacelle (WT) observations were available and if the site had production data. The three sites with production data are at some distance from each other (Table 1). Site A is located on a hill with relatively

20 high terrain west and north of the site, and somewhat lower terrain toward the south and east. In the location of site B there is instead similar terrain height to the south and west, while the terrain to the north and east is lower. The hill of site C extends mainly in the south-north direction, with lower terrain to the west and east. Site B and C have around 10 wind turbines, while site A consists of 20 turbines. All sites are surrounded with forest and with some lakes at lower levels. For each site one single value of production data is calculated by averaging production data from wind turbines without an error code. No de-icing

25 system was used on the wind turbines included in the study.

The meteorological parameters are measured every 10 minutes with the multi-instrument Quatro-Ind H (Lambrecht Meteorological Instruments, Germany), except at one site where the multi-instrument WXT510 (Vaisala, Finland) is used. Since the NWP forecast output is hourly, only the 10-minute observations at every full hour are used in the verification.

The ice load is measured with an IceMonitor Combitech AB (2016). The IceMonitor measures ice on a rotating cylinder

30 according to ISO 123494 specifications ISO 12494 (2000). Some problems have been identified with this instrument. One is that it may stop rotating, and the ice is then only accumulated on one side of the cylinder. Another issue is that the ice can cause the rod to lift and thus to measure an incorrect ice load (Thorsson et al., 2015). Therefore, it is difficult to use the ice measurements quantitatively, but they are still used here to get an approximate observation of the icing.

Table 1. Observation sites with approximate latitude, description if the measurements were made from a mast or at the wind turbine nacelle (WT) and available production data (x).

Site	A	B	C	D	E	F	G	H	I	J
Latitude	$65^o N$	$67^o N$	$58^o N$	$68^o N$	$68^o N$	$68^o N$	$62^o N$	$60^o N$	$60^o N$	$58^o N$
WT/Mast	WT	WT	WT	Mast	Mast	Mast	Mast	WT	WT	Mast
Prod.data	x	x	x							

The following quality controls were conducted for the observations of temperature (T), wind speed (WS) and relative humidity (RH). Values for $T < -40\,°C$, $WS < 0\,m\,s^{-1}$ and $RH < 0\,\%$ are removed from the data set as being unrealistic. Also, if the standard deviation of the 10-minute averaged wind speed was zero, the observation was removed. Finally, in order to remove unrealistic jumps in the observations: if the difference between the current observed value and the next deviates more than three times the standard deviation of this difference for the entire period, the observation was removed. ~~For one of the four sites with production data, we found that wind speed measurements were unreliable, as the instrument was heavily affected by icing. This influence could be confirmed in webcam pictures. Thus, this stations wind speed data and its production data was omitted.~~

For the production data, data was only used when no error code from the site was given. Thus, the reduction in observed production should be caused by icing.

2.3   Uncertainty quantification approaches

Two methods for uncertainty quantification are employed in this study; ensemble forecasting and the neighbourhood method.

Ensemble forecasting is used here to account for uncertainties in the initial and boundary conditions within the NWP model HarmonEPS as described above. The possibility to include model errors in the ensemble prediction system was omitted in order to stay as close as possible to the operational setup of the NWP model.

The neighbourhood method accounts for uncertainties in the representativeness of the forecast for a specific location of the wind turbines, e.g. the site of a wind turbine.

The control member (CM) from the ensemble forecast is used as baseline, as it reflects the use of a single deterministic forecast. In order to quantify the role of the different uncertainty sources on the forecast uncertainty, the benefit of using different combinations of the two above methods compared to baseline is investigated.

Four different combinations of the two methods were studied. These four uncertainty quantification approaches are in addition to CM presented in Figure 3 and are described below.

– CMngb (Control Member Neighbourhood): The model chain starts with a single NWP forecast, namely the control member from the ensemble forecast. Next, the neighbourhood method is added providing multiple forecast input to the icing and production model. This approach results in 25 forecasts from the 25 neighbouring grid points.
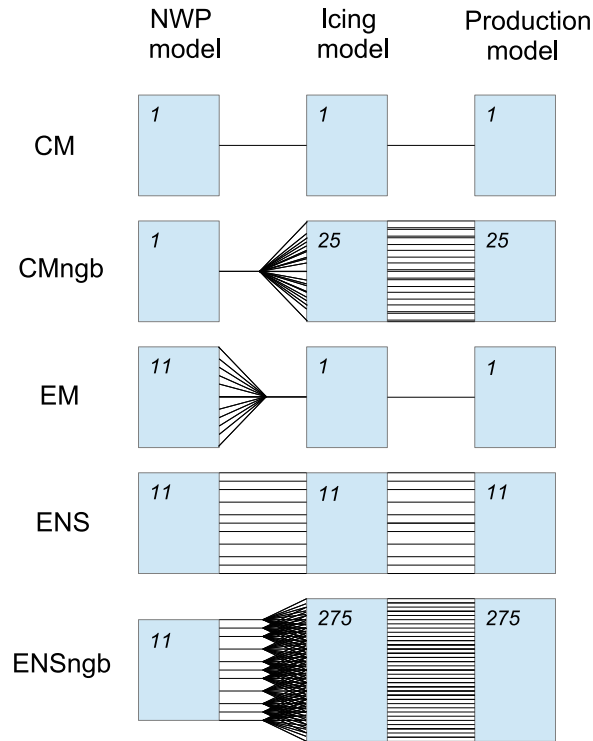
Figure 3. Uncertainty quantification approaches. The commonly used approach CM and the four uncertainty quantification approaches. CM: Control member, a single forecast used in each forecast step. CMngb: Control member together with the neighbourhood method. EM: Ensemble mean, averaging the ensemble members after the first modelling step. ENS: Ensemble, using all ensemble members throughout all modelling steps. ENSngb: Ensemble together with the neighbourhood method. The number in each box represents the number of forecast members in each forecast step.

- EM (Ensemble Mean): The model chain starts with the ensemble forecast from HarmonEPS. The 11 ensemble members (the 10 perturbed members and the control member) are then averaged before the icing forecast, providing the single statistically best meteorological input for the icing and production model. The uncertainty of the icing and production forecasts cannot be determined.

5 - ENS (Ensemble): The first modelling step is based on the ensemble forecast. All ensemble members are then used each as input to the icing and production loss models, resulting in 11 forecasts of icing and production loss. The 11 forecasts give an estimation of the uncertainty in the icing and production forecasts.

– ENSngb (Ensemble Neighbourhood): The ensemble and the neighbourhood method combined. The neighbourhood method is added to each ensemble member after the NWP model step, resulting in 25 forecasts for each ensemble member and a total of 275 forecasts used as input to the icing and production model.

2.4 Verification methods

5 The forecast skill is assessed by the root mean squared error (*RMSE*), the mean forecast error called *bias*, and the unbiased forecast error, ie. the standard deviation (*std*) of the forecast error. They are connected by:

$$RMSE^2 = bias^2 + std^2 \tag{3}$$

The calculation of the forecast error as the deviation between forecasted and true value uses an observation as replacement for the truth. Thus, the observational error needs to be taken into account in the calculation of the error terms. Due to the lack
10 of a consistent estimate for the observational error, the observational error is neglected. This leads to an overestimation of the RMSE, bias and std.

The magnitude of the bias of the meteorological parameters in this study varied between the observation sites and with forecast lengths. Thus, the average bias for each forecast length was estimated and removed from the $RMSE$ to get std.

By averaging the ensemble members for the same valid time, they can be treated in the same way as a single forecast, and
15 the same skill scores can be used allowing to compare the skill of the different probabilistic approaches with the skill of the single deterministic forecast denoted as CM.

The spread of the ensemble forecast contains important additional information compared to a deterministic forecast, such as a situation-dependent estimate of the forecast uncertainty. The spread $SPRE$ is defined as:

$$SPRE = \sqrt{\frac{1}{M-1} \sum_{i=1}^{M} (x_i - \bar{x})^2} \tag{4}$$

20 where x_i is the i-th forecast member, $\bar{x}$ is the ensemble mean and M is the number of forecast members.

In a perfect probabilistic forecast, any of the simultaneous forecast members would be statistically indistinguishable from the truth. If this is not the case, the forecast uncertainty is over- or underestimated, i.e. providing too wide or too narrow range of forecast outcomes, respectively. In order to verify the ensemble spread ($SPRE$), it is compared to the forecast skill of the ensemble mean in terms of unbiased forecast error (*std*) following Johansson (2017).

25 Both the std and the $SPRE$ are statistically expected to increase with increasing forecast length. In a perfectly calibrated forecast, the spread should be as large as the skill. The so-called spread/skill-relationship of the forecasts, $SPRE/std$, should therefore equal one. As the unbiased forecast error std is overestimated due to the neglected observational error as mentioned above, the spread/skill-relationship will consequentially be underestimated. This error could be corrected with an appropriate estimate of the observational error (Schwartz et al., 2014).

11

The skill of the forecasts made with the different approaches from section 2.3 and the spread/skill relationship are presented below. Since a relatively short period of forecasts is studied with only two icing events, it is not possible to test the significance of the statistical measures. The results should therefore only be considered to show the potential benefit of the use of probabilistic forecasting for wind power in cold climates.

5 ## 3   Results

3.1   Meteorological parameters

For the two-week period examined in this study, the skill of the basic meteorological model performance is presented in Figure 4 in terms of bias and unbiased forecast error for relative humidity, wind speed and temperature for the 42-hour forecasts. The statistics is based on all forecasts at all 10 observation sites for each forecast length. The unbiased forecast error averaged

10 over the forecast window amounts to about 7 %, 2.3 m s^{-1} and 0.9 $°C$ for relative humidity, wind speed and temperature, respectively. The forecast error is expected to increase with forecast length, which is the case for temperature (Figure 4c). The unbiased forecast error of the wind speed and relative humidity (Figure 4a/b), however, displays a flat shape suggesting that the error is saturated during the forecast window. This behaviour points to a difficulty in the analysis, i.e. the initialization of the wind field in the model. One aspect in this regard is the lack of wind observations in the planetary boundary layer to initialize

15 the meso-scale wind field.

The temperature bias increases from -0.3 $°C$ to close to -1 $°C$ after 42 hours (Figure 4c). This behavior might be caused by a spin-up problem in the model, since it is changing with forecast length. Additionally, a negative temperature bias that is not changing with forecast length can be attributed to the warm turbine affecting the measurements. A difference in the temperature bias between Mast and WT measurements amounts on average to -0.4 °C for mast and -1.1 °C for WT. The bias

20 of the wind speed (Figure 4b) decreases slightly with forecast length from 1 m s^{-1} to 0.5 m s^{-1} after 42 hours. The bias in the relative humidity (Figure 4a) displays only small changes with forecast length and remains generally small with absolute values less than 1 %.

In general, the meteorological model displays good performance for the basic meteorological parameters. The lack of forecast error growth in wind speed and relative humidity suggests that the model forecast could be improved by assimilating more

25 observations, especially for wind and humidity. A better initial state of the forecast might even improve the bias behavior.

The meteorological performance of the different uncertainty quantification approaches in terms of spread and skill is shown in Figure 5. The skill as defined as the unbiased forecast error std and spread for the temperature forecasts is displayed for each approach. The baseline approach of a deterministic forecast (CM) shows the largest forecast error. For the examined period, the benefit of using neighbourhood and/or ensemble methods can already be seen in the first hours of the forecasts,

30 and increases with forecast length as expected. The CMngb has the smallest improvement of forecast error, but still suggesting that the neighbourhood method is valuable, if an ensemble forecast is not available. Even higher improvement is achieved by the approaches ENS and EM that provide by construction the same values here. The largest reduction of the forecast error is achieved using both the ensemble forecast method and the neighbourhood method (ENSngb), with an average reduction of
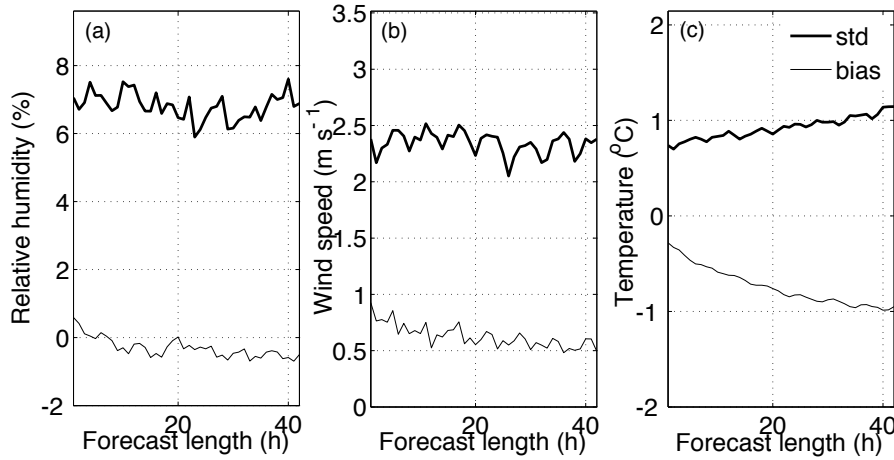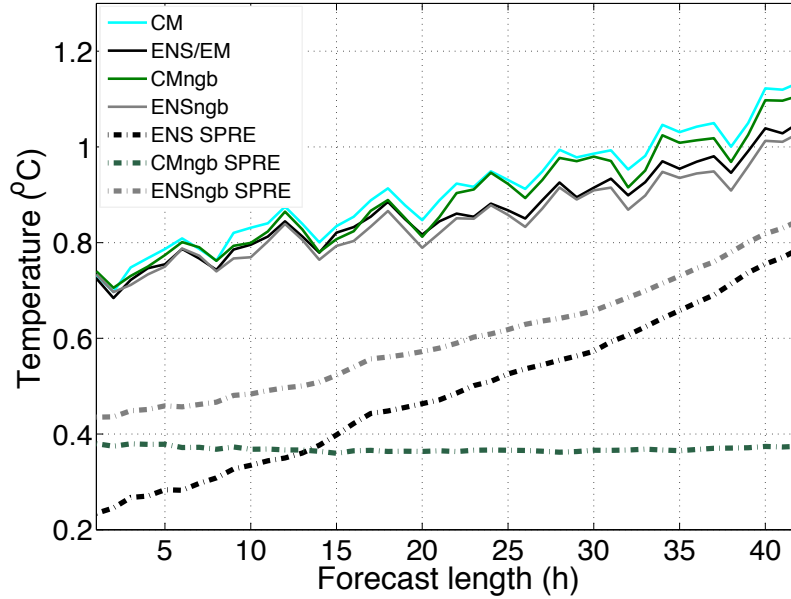
Figure 4. CM: Unbiased forecast error std and bias for a) relative humidity in percent, b) wind speed in m s^{-1} and c) temperature in ^{o}C for increasing forecast length. The statistical measures are calculated for the two-week period and averaged over all available stations. The meteorological observations at 60 to 150 m height above ground from 10 sites (9 for wind speed) are used.

9 % for the temperature, 7 % for the wind speed and 12 % for the relative humidity forecast error averaged for all sites and all forecasts averaged for all sites and all forecasts. ~~Similar behavior was found for the forecasted wind speed and relative humidity (not shown).~~

Figure 5 also displays the forecast spread of the approaches ENS, ENSngb and CMngb. The spread is always clearly lower than the unbiased forecast error, which means that the forecast uncertainty is underestimated in all approaches. This behaviour might result from neglected uncertainty sources in the modelling chain (Fig. 1) The employed ensemble represents uncertainties in the initial and boundary conditions, but it does not take into account uncertainties in the model physics or numerical formulations. There are methods to account for these uncertainties, such as stochastic physics, which increase the spread (Bouttier et al., 2015), but they are not in the scope of this study. It is also important to consider the observational error when validating the spread-skill relationship of the forecast, as discussed in Section 2.4. Since the forecast error estimate is not corrected for the observational error, the underdispersiveness of the approaches is not as large as it appears in Fig. 5.

On its own, the ensemble forecasting method (ENS) has better skill, spread and spread/skill relationship than the neighbourhood method (CMngb) for this period, around 0.6 and 0.3 respectively for all meteorological parameters. Specifically, the spread/skill relationship improves from around 0.3 for CMngb to around 0.6 for ENS for temperature, wind speed, and relative humidity. The approach ENSngb results in the best spread/skill relationship, around 0.7. This implies that the neighbourhood method is adding additional information about the forecast uncertainty, which is valuable for both the forecast skill and the forecast spread. The spread resulting from the neighbourhood method is constant with forecast length (CMngb in Figure 5), since it represents the internal variations of the weather over the scale of the neighbourhood domain, but it does not take into account uncertainties due to initial boundary conditions, or model formulations that increase with increasing forecast length.

13

Figure 5. Unbiased forecast error std for all uncertainty quantification approaches for the temperature in ^{o}C and the spread, $SPRE$, for the three approaches consisting of multiple forecasts as a function of increasing forecast length. The std is calculated for the average forecast in case of multiple forecasts i.e. approach CMngb, ENS and ENSngb.

3.2 Forecasted Icing and Production losses

The forecasted icing is, as mentioned before, more difficult to validate since the observations of icing are unreliable. Alternatively, the forecasted production loss for the three sites with production data is also considered as a measure of the forecast skill of the icing. Results from two of the three sites, here A and B, with consistent production observations will be presented in more detail in the following to point out some interesting features when using the probabilistic forecasts. Site C is not shown, but had icing during about three days of the two weeks. In Section 3.2.3, the average skill and spread for the different approaches for the three sites are discussed. Sites A and C were affected by a frontal passage during the two weeks, while site A experienced general cloudiness and was less influenced by this front.

3.2.1 Site A

Two icing events were forecasted at site A during the two weeks. During the first icing event, around the 30/12, only a small amount of ice is forecasted, while during the second icing event, starting around 1 January, both the modeled and observed ice load amount to several kg m^{-2}. In Figure 6a the forecasted icing using the ENS approach and the CM approach is presented together with the ensemble mean of the ENS approach and observed ice load. The spread of the ensemble members from

14

the ENS, which displays the forecast uncertainty of the icing, are presented in Figure 6b. During the icing events, the spread signals uncertainty in the ice amount. Largest forecast uncertainty is found at the end of the second icing event with around 1.5 kg m^{-2}. The magnitude of the overall forecast uncertainty, or spread, amounts to about 50 % of the ice load here.

In this period, first the build up and then the loss of ice happen almost simultaneously for all the ensemble members. This is especially visible in the spread for the second icing event with values close to zero at the start and end of the event. Forecasted and observed production loss for the site allow studying the timing during the event in more detail (Fig. 7a). The forecasted production loss starts about 12 hours later than the observed production loss, pointing to a problem with the agreement between the ensemble members. A closer examination reveals low forecasted liquidand solid water content in the beginning of the icing period by all the ensemble members(Figure 6c), resulting in no accumulation of ice by the icing model. Instead the ice starts to accumulate in the model when solid water componentssnow and graupel are forecasted, in addition to the liquid water content. Interestingly all ensemble members behave very similarly with this timing (not shown). This could be related to an error in the modelled cloud characteristics or in the icing model causing a too slow ice buildup.

Additionally, the ensemble fails to describe the end of the icing period (6-9 January) where the modelled production loss drops to zero for the remaining period, while the observed production loss is high around 8 Januray (Fig. 7a). The end of the two weeks also lacks forecasted liquid and solid water content, and thus no further buildup by the icing model is possible. Again, the ensemble was overconfident with all members displaying a similar behaviour. By including the neighbourhood method to the ensemble, there is some improvement in the end of the two weeks adding some forecasted liquid water content at the end of the simulation (Fig. 8) in some members. However, the amount of forecasted water content was still too low to generate any notable buildup by the icing model.

A closer look at the weather situation shows that site A is affected by a frontal passage during the second week. The ensemble members all have the front passage at the same time, resulting in similar cloud cover and similar similar values for cloud cover as well as liquid and solid water content. This problem results from the insufficient representation of the uncertainties in the boundaries. The boundary data from different members of the ECMWF EPS prescribes very similar surface pressure patterns for this event resulting in an overconfidence in the arrival time of the cloud front. Furthermore, small-scales are initialized for all ensemble members in the same way using the control member analysis. During the icing event however, the ensemble members have a spread of around 25 % of the ENS mean value for liquid and solid water components, leading to a spread in the amount of built-up ice (Fig. 6c). In Davis et al. (2014) it was shown using a similar model that the resulting ice load was highly sensitive to a variation of the ingoing median volume droplet size of the water droplets, suggesting that this is an uncertain part of the icing model. Here, the droplet size is calculated from the liquid/solid water content before being used in the icing model. This is discussed further in the section below (Section 3.2.2).

3.2.2 Site B

At site B only one icing event was forecasted and observed, starting already on the first days of the period. This site was not as strongly affected by frontal passages as site A, and had more general cloudiness during the period. Figure 9a-c shows the
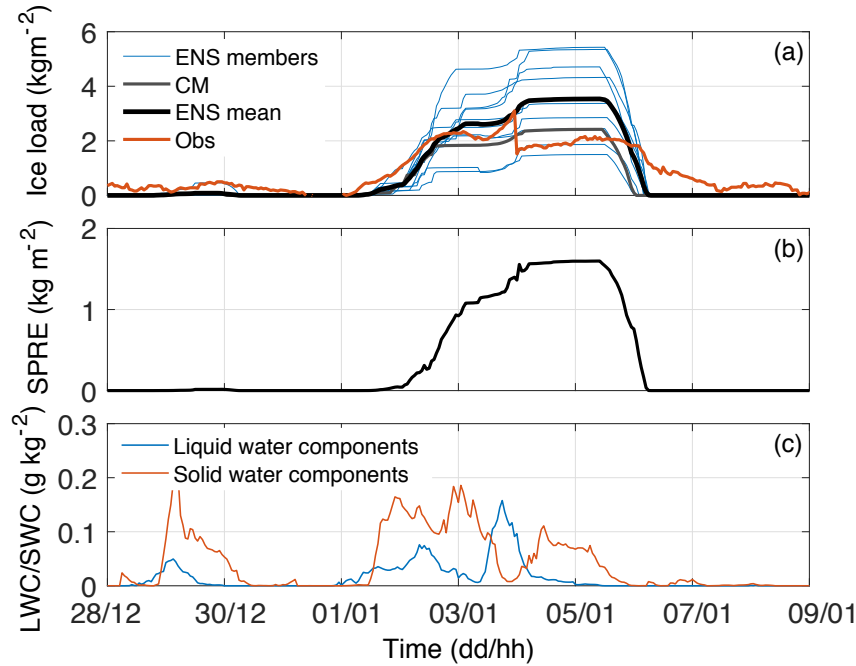
Figure 6. Site A. a) Forecasted ice load in kg m^{-2}. Blue lines are ensemble members, grey line is the CM and black line is the ENS mean. The ice load observations are in red. b) The related spread of the ensemble members in ENS. c) Forecasted ENS ensemble mean liquid water content in g kg^{-1} for the period.

forecasted and observed ice load, spread of forecasted ice load and liquid water content. In Figure 10a-b the forecasted and observed production loss as well as the related forecast spread is presented.

In the beginning of the two weeks, there is one day with a large amount of liquid and solid water content, however, no ice started to build up during this day, mainly because of too low wind speeds. At the start of the actual icing event, the ice growth is first small, but with increasing pace after January 1, 2012, reaching the largest load of about 2.5 kg m^{-2} on January 4 (Fig. 9a). The observed production loss shows a similar behaviour (Fig. 10). In Fig. 11a the ENS mean MVD for the solid water components can be seen for this site and the corresponding mean spread is shown in Fig. 11b. The liquid water components ENS mean MVD is not shown, as it follows the LWC amount in this case (Fig. 9c). Large values of MVD (Fig. 11a) coincide with large icing rates (Fig. 9a) stressing the role of MVD for the icing intensity. The large spread of the MVD around January 4 (Fig. 11b) can be connected to the ice load differences of the different ensemble members (Fig. 9a) and also to the simultaneous spread in the production loss forecast (Fig. 10b). This behaviour agrees with the effect of different MVD discussed in Davis et al. (2014).

Generally, the forecasted ice and production loss displays a reasonable agreement with obsevations. However, there are two interesting deviations. Firstly, in the first half of the time series the forecast is generally underestimating the observed production loss (Fig. 10), while the forecasted ice load agrees well with the observations (Fig. 9a). The reason of the disagreement

16
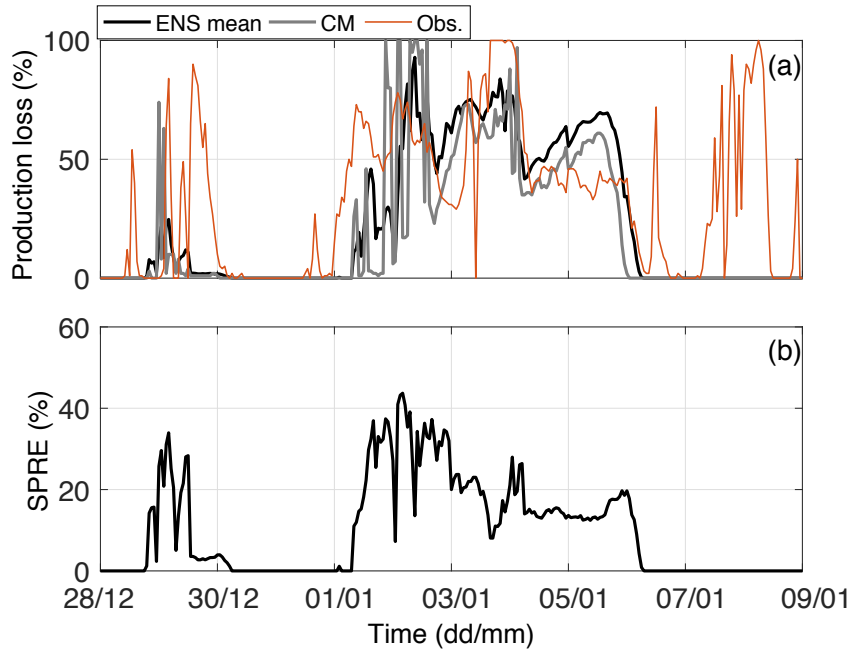
Figure 7. Site A. a) Forecasted production loss in percent. Black line is the ensemble mean using ENS, grey line is the CM and red line is the observed production loss. b) The related spread of the ensemble members in ENS.

seems to be related to deficiencies in our production model. Secondly, the observed drop in both ice load and production loss at the end of the period is delayed in all forecasts, opposite to the behaviour at site A.

The high forecast spread, or forecasted uncertainty of the icing forecast (Fig. 9b) shows the value of having a probabilistic forecast. This variation of the ensemble members is probably due to a large variation of the liquid and solid water content (between 25 and 50 % of the mean amount) resulting in a variation of the calculated MVD. The benefit of an ensemble in this case can also be seen in Figure 10a where the ensemble mean (black line) of the production loss forecast starts to decrease, getting closer to observations, while the CM is remaining nearly at 100 %. The forecast spread of the production loss is also increasing during the last day of the period, suggesting an enlargened uncertainty in the forecast. Generally, forecasting the end of an icing period has been shown to be challenging due to the difficulties in modelling ice loss Davis. On the other hand the start of the stronger icing period from the 3rd of January is well timed in the forecast data, and furthermore, the spread of the production loss forecast increases from approximately 5 % to 30 % simultaneously. This increased spread is a useful indicator for the uncertainty of the start of the ice period and provides additional information to the actual forecast.
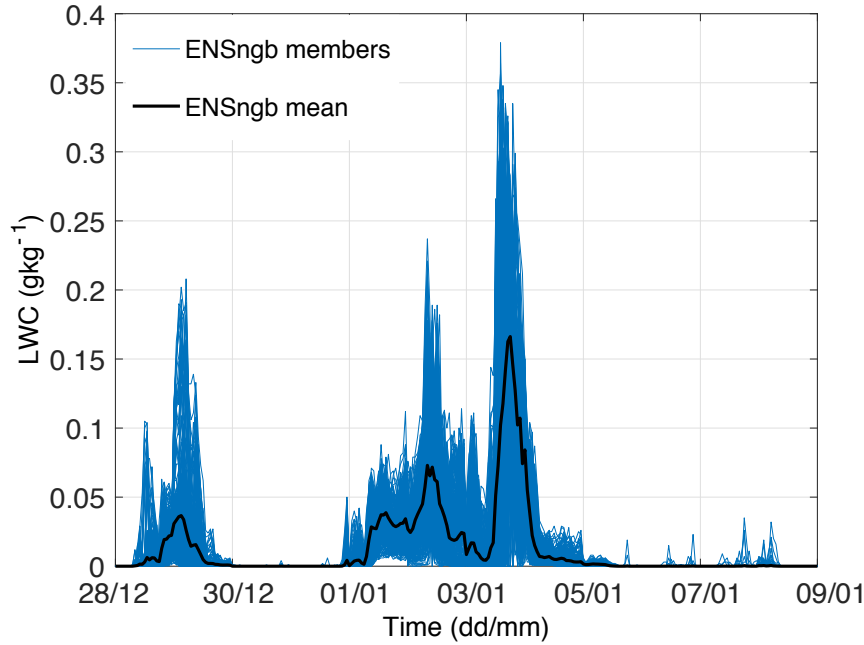
Figure 8. Site A. Forecasted liquid water content in g kg^{-1} using the ENSngb approach.

3.2.3 Forecast performance of the different approaches

In Figure 12a the forecasted next-day production using the four different uncertainty quantification approaches are compared with the single forecast of CM and observed production for site A. For the approaches that generate multiple forecasts, the average of the forecast members is presented. The production is calculated for a 2 MW-turbine.

5   The different approaches are following the observed production most of the time. As the figure suggests visually and the RMSE calculation below confirms, the ENSngb mean is the most skillful approach for this site. A closer inspection of Fig. 12a reveals some typically behaviour for the approaches. All approaches overestimate the production during the start of the second icing period around the 1 January, which agrees with the underestimated production loss in Figure 10a. The different model forecasts tend to overestimate the production occasionally. This probably happens due to an overestimation of the potential

10   production, since the wind speed has a positive bias (figure 4b). A bias correction of the wind speed would be useful to reduce this error.
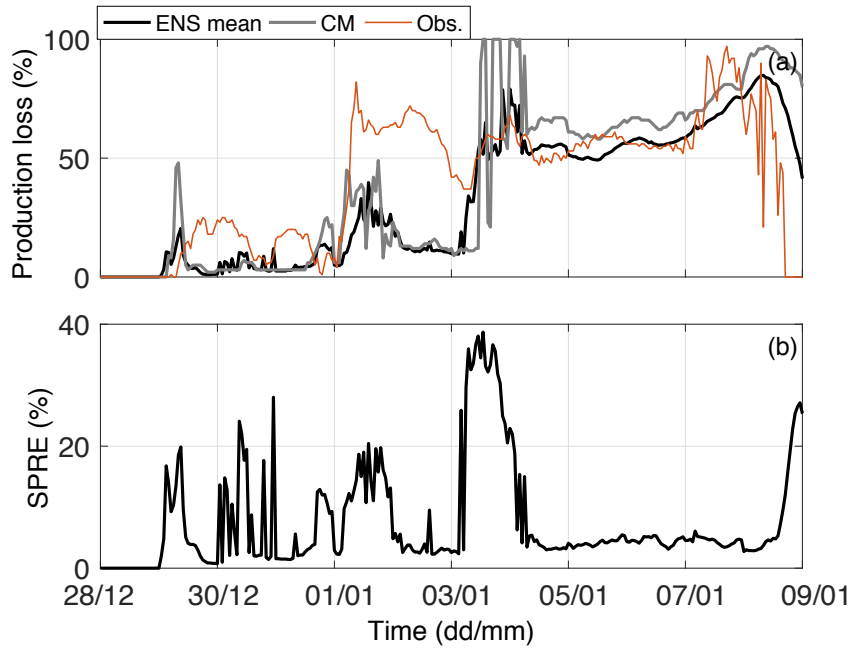
     In Figure 12a the single forecast of CM has stronger variations than the other approaches. These variations are partly a sign of uncertainty in the forecast and belong to unpredictable phenomena. They are filtered out when averaging over the members of the probabilistic approaches. However, it is important to realize that this filtering creates a so-called unoccupied

15   average, ie. the smoothed state is unrealistic, since part of the variance is now included in the spread of the forecast members describing the forecast uncertainty. This spread for the uncertainty quantification approaches can be seen in Figure 12b. The
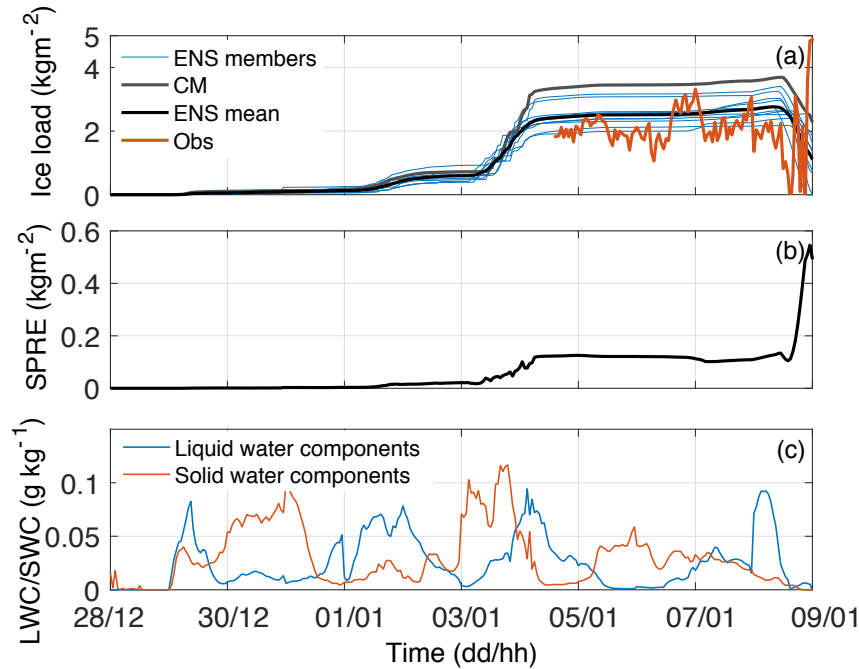
Figure 9. Same as Fig. 6, but for Site B.

ensemble (ENS) provides larger spread than the neighbourhood method applied to a single forecast (CMngb). Similar to the meteorological parameters, the largest spread, and thus forecasted uncertainty, is generally seen for the combination of ensemble and neighbourhood method (ENSngb). The average spread/skill ratio of the ENSngb production forecast amounts to around 0.7, compared to 0.5 for ENS. The increased spread for the ENSngb should therefore provide the best estimate for the

5   actual forecast uncertainty, even if the model is still overconfident.

As a summary of the forecast performance, Table 2 yields the mean error as $RMSE$ of the 06 UTC +18-42 h forecasts for the production loss and for the production, averaged over the three observation sites where production data was available. The forecast quality of the production forecast using the different approaches generally follows the order in the first step of the modelling chain, i.e. for the meteorological parameters (Figure 3). The benefit of the neighbourhood method for the production

10 loss forecast can be seen comparing the $RMSE$ forCM and the CMngb. However, the production loss $RMSE$ is the same for ENS and ENSngb, suggesting that the neighbourhood method neither contribute to or reduce the forecast skill when added to the ENS approach. Using the ENS or ENSngb approach compared to CM results in a reduction of the production loss forecast error from 26 to 21 %. It should be noted that the usage of the ensemble mean as the input to the icing and production loss model (approach EM) deteriorates the forecast quality to 29 % compared to the ensemble-based approach ENS with 21 %,

15 where the output from each member is calculated through the entire chain. This results from the non-linearity of the icing and production loss model. The increased forecast quality for the forecasted meteorological parameters by the ensemble mean is lost by the usage of the unoccupied average as input into the icing and production model.

19

Figure 10. Same as Fig. 7, but for Site B.

For the production forecast, the best forecast is provided by the ENSngb approach, while the worst comes from the single forecast of CM. Here, adding the neighbourhood method to the ENS approach reduces the $RMSE$ (Table 2). Using the ENSngb approach compared to CM results in a reduction of the production forecast error from 0.49 to 0.41 MW or by 16 % relative to the CM forecast.

5   The two different versions of neighbourhood methods, ie. terrain-following or constant height described in Section 2.1.1 displays some differences for the production forecast. Assuming that wind power is installed at higher elevation than the surroundings, the terrain-following version provides neighbours from lower absolute heights with higher moisture content, and thus more atmospheric icing, and more surface-affected wind fields compared to the constant-height version. The quality of the resulting production loss forecast is very similar for the two versions, resulting in the same $RMSE$. For the final production

10   forecast however, an improvement is seen from the constant-height version compared to the terrain-following one which can be traced back to better wind forecasts from the constant-height neighbourhood method.

     It should be pointed out once more that the statistical significance of the results could not be assessed since this is a case study with a limited sample size, but the results are consistent in the different analyses and supports theoretical expectations. The improved forecast skill of the production loss and of the production, using the four probabilistic forecasting approaches

15   instead of the single forecast of CM, also suggests that the icing forecast is improved, even though it is not possible to validate with the available ice load observations.
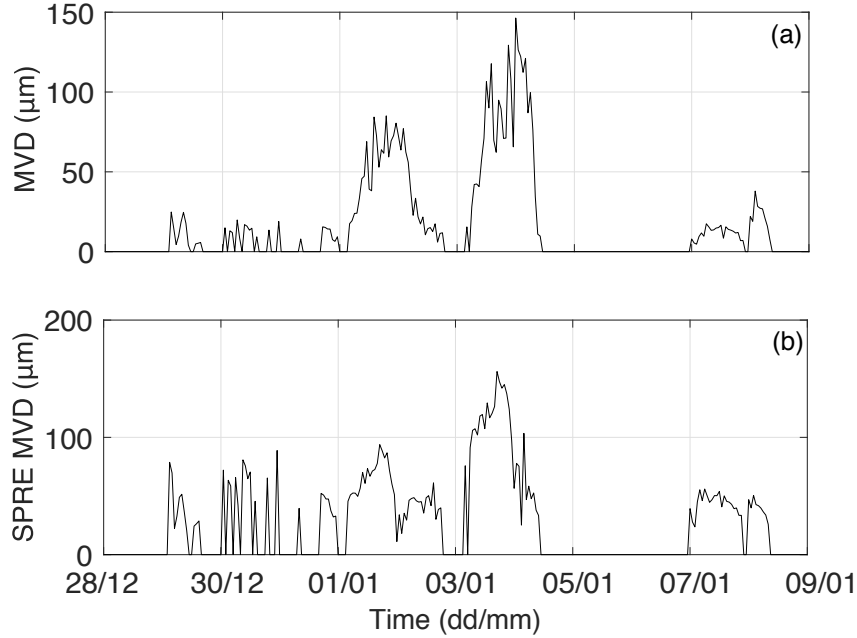
Figure 11. Site B. a) ENS mean MVD for the solid water components. b) Mean spread of MVD for the solid water components.

Table 2. $RMSE$ of the different approaches for production and for production loss forecasts averaged over the three sites with production data. The production is calculated for a 2 MW-turbine.

Approach	Prod (MW)	Prod loss (%)
CM	0.49	26
CMngb	0.44	23
EM	0.47	29
ENS	0.44	21
ENSngb	0.41	21

4 Concluding Remarks

The problem of predicting next-day production losses due to icing of wind turbines has been addressed with the usage of probabilistic forecasting. Two methods, ensemble forecasting and the neighbourhood method, have been used in four different uncertainty quantification approaches to produce probabilistic forecasts. Improved skill and estimations of the forecast uncertainty were both investigated in this two-week case-study. The main results are:

5
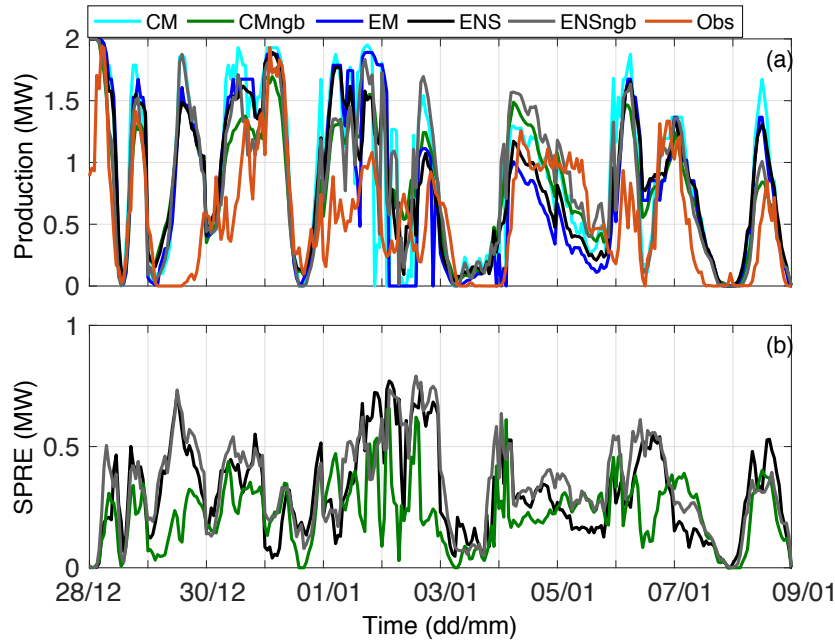
Figure 12. a) Production forecast in MW at site A for the different approaches (Fig. 3). For CMngb, ENSngb and ENS the lines present the average of the multiple forecasts. Observed production is given with the dash-dotted line. The production is calculated for a 2 MW-turbine. b) The spread of the multiple members in CMngb, ENSngb and ENS.

- Using probabilistic forecasting improves the forecast skill for the meteorological parameters, the icing and the icing-related production loss compared to the commonly used approach with one single deterministic forecast.

- The spread of the multiple forecasts can be used as an estimation of the forecast uncertainty, also for icing and related production losses. However with the current model setup, the uncertainty is underestimated both for the meteorological parameters and for the production.

- The approach where both the uncertainties in initial/boundary conditions and the representativeness of the wind turbine are represented, ENSngb, has the highest skill for the next-day production forecast. This suggests that both errors should be taken into account when generating a probabilistic forecast.

Improving the skill by the use of an ensemble forecast is a useful contribution to wind power forecasting in cold climate. Additionally, a reliably forecasted uncertainty can be of great value for end-users as probabilistic forecasts of icing events and related production losses. Even though the spread of the forecast is too low and, hence, the forecast uncertainty underestimated, it could be utilized if the spread was calibrated, see e.g. Veenhuis (2012) or Sloughter et al. (2012). Knowing the likelihood for

22

icing, the end-user can employ site-specific cost-loss ratios in the decision making for the trading process, the use of de-icing systems and for the safety of people working at the wind farm.

To further develop the use of probabilistic forecasting in this area, it is important to note that we are not taking into account all of the uncertainties in the modelling chain, e.g. errors resulting from approximations made in the different models. It is known that the icing model contains numerous uncertain parameters, such as the sticking efficiency in case of snow and wind erosion (Davis et al., 2014; Nygaard et al., 2013). The inclusion of these uncertainties into the entire modelling chain is currently ongoing research.

The weather is the most fundamental part of the modelling chain and due to its chaotic behaviour the focus here is the meteorological model. As the forecast skill for wind speed and relative humidity at nacelle height appeared to be saturated during the first 42 forecast hours, more effort is required in order to improve the initial state of the forecast, e.g. through data assimilation methods or inclusion of local observations of humidity and wind, by radar and other instruments. From the meteorological spread/skill relationship, it can be concluded that also more spread is needed. The lack of spread in the ensemble was especially visible during a frontal passage where all members had very similar boundary conditions and thus the same timing of the frontal passage. Better spread from the boundary conditions can be achieved by a smart selection of the global host model ensemble members (Molteni et al., 2001). Furthermore, full-scale data assimilation for all ensemble members, and not only for the control member, would allow for a better spread in the initial state of the ensemble members.

Finally, a NWP-ensemble forecast is computationally expensive to run and requires extensive infrastructure. Thus, often only national meteorological services can operationally produce ensemble forecasts. Many national weather services run currently an operational ensemble forecast that is often disseminated as open data following the European INSPIRE directive. The delivery time for such data are around 3 hours after analysis time, thus making it possible to use the 06-UTC model run for next-day wind power forecasts. Many users will then be able to include this forecast data into their own icing and production models, which are far less computationally expensive. This will allow for a wide application of probabilistic weather forecasting for wind power in cold climates.

References

Al-Yahyai, S., Charabi, Y., Al-Badi, A., and Gastli, A.: Nested ensemble NWP approach for wind energy assessment, Renewable Energy, 37, 150–160, https://doi.org/10.1016/j.renene.2011.06.014, 2011.

Andre, U. and MetCoOp-Team: ALADIN-HIRLAM Newsletter, http://www.umr-cnrm.fr/aladin/meshtml/NL8-final.pdf, 2017.

5 Bengtsson, L., Andere, U., Aspelien, T., Batrak, Y., Calvo, J., De Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hartal, M., Ivarsson, K.-I., Lenderink, G., Miemelä, S., Pagh Nielsen, K., Onvlee, J., Rontu, L., Samuelsson, P., Santos Munoz, D., Subias, A., Tijm, S., Toll, V., Yang, X., and Ødegaard 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.

Bergström, H., Olsson, E., Söderberg, S., Thorsson, P., and Undén, P.: Wind power in cold climates, Ice mapping methods, Tech. rep., Elforsk,
10 Stockholm, 2013.

Bouttier, F., Raynaud, L., Nuisser, O., and Ménétrier, B.: Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX, Quarterly Journal of the Royal Meteorological Society, https://doi.org/10.1002/qj.2622, 2015.

Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B., and Beare, S. E.: The MOGREPS short-range ensemble prediction system, 722, 703–722, https://doi.org/10.1002/qj, 2008.

15 Combitech AB: The Ice load surveillance sensor Icemonitor, Tech. rep., http://www.rwis.net/res/pdffiles/IceMonitor{_}Product{_}Sheet.pdf, 2016.

Davis, N., Hahmann, A. N., Clausen, N. E., and Žagar, M.: Forecast of icing events at a wind farm in Sweden, Journal of Applied Meteorology and Climatology, 53, 262–281, https://doi.org/10.1175/JAMC-D-13-09.1, 2014.

Davis, N. N.: Icing Impacts on Wind Energy Production by, Ph.D. thesis, DTU Wind Energy.

20 Davis, N. N., Pinson, P., Hahmann, A. N., Clausen, N.-e., and Žagar, M.: Identifying and characterizing the impact of turbine icing on wind farm power generation, Wind Energy, 16, 1503–1518, https://doi.org/10.1002/we, 2016.

Du, J., Mullen, S. L., and Sanders, F.: Short-Range Ensemble Forecasting of Quantitative Precipitation, Monthly Weather Review, 125, 2427–2459, https://doi.org/10.1175/1520-0493(1997)125<2427:SREFOQ>2.0.CO;2, 1997.

Du, J., MiMego, G., Tracton, S., and Zhou, B.: NCEP Short-range ensemble forecasting (SREF) system: Mulit-IC, multi-model and multi-
25 physics approach, Tech. rep., Enviromentl Modeling Centre/NCEP/NOAA/, Washinton DC, USA, 2003.

ISO 12494: Atmospheric icing of structures (ISO/TC 98/SC3), Tech. rep., 2000.

Johansson, Å.: Another look at Spread and Skill, Revised version to be submitted in Monthly Weather Review, 2017.

Kraj, A. G. and Bibeau, E. L.: Phases of icing on wind turbine blades characterized by ice accumulation, Renewable Energy, 35, 966–972, https://doi.org/10.1016/j.renene.2009.09.013, 2010.

30 Leith, C. E.: Theoretical Skill of Monte Carlo Forecasts, Monthly Weather Review, 102, 409–418, 1974.

Leutbecher, M. and Palmer, T. N.: Ensemble forecasting, Journal of Computational Physics, 227, 3515–3539, https://doi.org/10.1016/j.jcp.2007.02.014, 2008.

Makkonen, L.: Models for the growth of rime, glaze icicles and wet snow on structures, Philosophical Transactions of The Royal Society Lond., 358, 2913–2039, https://doi.org/10.1098/rsta.2000.0690, 2000.

35 Mazin, I. P., Korolev, A. V., Heymsfield, A., Isaac, G. A., and Cober, S. G.: Thermodynamics of icing cylinder for measurements of liquid water content in supercooled clouds, Journal of Atmospheric and Oceanic Technology, https://doi.org/10.1175/1520-0426(2001)018<0543:TOICFM>2.0.CO;2, 2001.

Megner, L., Tan, D. G. H., Körnich, H., Isaksen, L., Hor, A., Stoffelen, A., and Marseille, G.: Linearity aspects of the ensemble of data assimilations technique, Quarterly Journal of the Royal Meteorological Society, 141, 426–432, https://doi.org/10.1002/qj.2362, 2015.

Mittermaier, M. P.: A Strategy for Verifying Near-Convection-Resolving Model Forecasts at Observing Sites, Weather and Forecasting, 29, 185–204, https://doi.org/10.1175/WAF-D-12-00075.1, 2014.

5 Molteni, F., Buizza, R., Marsigli, C., Montani, A., Nerozzi, F., and Paccagnella, T.: A strategy for high-resolution ensemble prediction. I: Definition of representative members and global-model experiments, Quarterly Journal of the Royal Meteorological Society, 127, 2069–2094, 2001.

Müller, M., Homleid, M., Ivarsson, K.-I., Køltzow, M., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., Dahlgren, P., Kristiansen, J., and Randriamampianina, Roger Ridal, M.: AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction Model, Weather and Forecasting, 32, 609–627, https://doi.org/10.1175/WAF-D-16-0099.1, 2017.

10 Nygaard, B. E. K., Àgústsson, H., and Somfalvi-Tóth, K.: Modeling Wet Snow Accretion on Power Lines : Improvements to Previous Methods Using 50 Years of Observations, Journal of Applied Meteorology and Climatology, 52, 2189–2203, https://doi.org/10.1175/JAMC-D-12-0332.1, 2013.

Pinson, P. and Kariniotakis, G. N.: Conditional Prediction Intervals of Wind Power Generation, IEEE Tranactions on Power Systems, pp. 1845–1856, https://doi.org/10.1109/TPWRS.2010.2045774, 2010.

15 Ronsten, G., Wallenius, T., Hulkkonen, M., Baring-Gould, I., Cattin, R., Durstewitz, M., Krenn, A., Laakso, T., Lacroix, A., Tallhaug, L., Byrkjedal, Ø., and Peltola, E.: IEA Wind Task 19, State-of-the-Art of Wind Energy in Cold Climates, Tech. rep., iea wind, 2012.

Schäfer, M.: Entwicklung eines Windrampendetektors zur Unterstützung der Netzintegration von Windenergie, Mäster theses, Johannes Gutenberg Universität Mainz, Germany, 2014.

20 Schwartz, C. S., Romine, G. S., Smith, K. R., and Weisman, M. L.: Characterizing and Optimizing Precipitation Forecasts from a Convective-Premitting Ensemble Initialized by a Mesoscale Ensemble Kalman Filter, Weather Forecasting, 29, 1295–1318, https://doi.org/http://dx.doi.org/10.1175/WAF-D-13-00145.1, 2014.

Seity, Y., Brousseau, P., 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.

25 Sloughter, J. M., Gneiting, T., and Raftery, A. E.: Probabilistic Wind Vector Forecasting Using Ensembles and Bayesian Model Averaging, Monthly Weather Review, 141, 2107–2119, https://doi.org/10.1175/MWR-D-12-00002.1, 2012.

Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme . Part II : Implementation of a New Snow Parameterization, Monthly Weather Review, 136, 5095–5115, https://doi.org/10.1175/2008MWR2387.1, 2008.

30 Thorsson, P., Söderberg, S., and Bergström, H.: Modelling atmospheric icing: A comparison between icing calculated with measured meteorological data and NWP data, Cold Regions Science and Technology, 119, 124–131, https://doi.org/10.1016/j.coldregions.2015.07.003, 2015.

Traiteur, J., Callicutt, D. J., Smith, M., and Baiday Roy, S.: A Short-Term Ensemble Wind Speed Forecasting System for Wind Power Applications, pp. 1763–1774, https://doi.org/10.1175/JAMC-D-11-0122.1, 2013.

35 Veenhuis, B. A.: Spread Calibration of Ensemble MOS Forecasts, Monthly Weather Review, 141, 2467–2482, https://doi.org/10.1175/MWR-D-12-00191.1, 2012.

Wallenius, T., Peltola, E., Lehtomäki, V., Antikainen, P., Bluemink, G.-J., Wadam-Gagnon, M., Dillingh, J., Clausen, N.-E., and Klintström, R.: World Market Update 2012, International Wind Energy Development, Tech. rep., Navigant Research, 2013.

WMO: Guidelines on Ensemble Prediction, Chair, Publication Board, Geneva, 2012.

Yano, J.-I., Bengtsson, L., Geleyn, J.-F., and Brozkova, R.: Towards a Unified and Self-Consistent Parameterization Framework, World Scientific, 2, https://doi.org/10.1142/9781783266913_0030, 2015.