Wind Energy

WIND ENERGY
Wind Energ. 2016; **19**:1503–1518
Published online 30 September 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/we.1933

RESEARCH ARTICLE

Identifying and characterizing the impact of turbine icing on wind farm power generation

Neil N. Davis[1], Pierre Pinson[2], Andrea N. Hahmann[1], Niels-Erik Clausen[1] and Mark Žagar[3]

[1] Department of Wind Energy, Technical University of Denmark, Roskilde, Denmark
[2] Department of Electrical Engineering, Technical University of Denmark, Lyngby, Denmark
[3] Vestas Wind Systems A/S, Aarhus, Denmark

ABSTRACT

Wind park power production in cold climate regions is significantly impacted by ice growth on turbine blades. This can lead to significant errors in power forecasts and in the estimation of expected power production during turbine siting. A modeling system is presented that uses a statistical modeling approach to estimate the power loss due to icing, using inputs from both a physical icing and a numerical weather prediction model. The physical icing model is that of Davis *et al.*[1] with updates to the simulation of ice ablation. A new approach for identifying periods of turbine blade icing from power observations was developed and used to calculate the observed power loss caused by icing. The observed icing power loss for 2 years at six wind parks was used to validate the modeling system performance. Production estimates using the final production loss model reduce the root mean squared error when compared with the empirical wind park power curve (without icing influence) at five of the six wind parks while reducing the mean bias at all six wind parks. In addition to performing well when fit to each wind park, the production loss model was shown to improve the estimate of power when fit using all six wind parks, suggesting it may also be useful for wind parks where production data are not available. Copyright © 2015 John Wiley & Sons, Ltd.

KEYWORDS
icing; power production

Correspondence
Neil N. Davis, Department of Wind Energy, Technical University of Denmark, Roskilde, Denmark.
E-mail: neda@dtu.dk

Received 16 September 2014; Revised 17 July 2015; Accepted 31 August 2015

NOMENCLATURE

Abbreviation	Description
Power curves	
PC_m	manufacturer's power curve
PC_i	icing threshold power curve
PC_e	turbine-specific empirical power curve
PC_{park}	park-specific empirical power curve
Ice mass	
ICE_g	hourly accumulated ice mass
ICE_r	hourly ablated ice mass
ICE_t	total accumulated ice mass
Estimated power production	
ALL	estimated power using model fit from all wind parks
PARK	estimated power using model fit from each corresponding wind park

1. INTRODUCTION

Production loss due to icing is one of the main challenges for expanding wind energy in cold climate regions, which for wind energy includes regions where temperatures commonly fall below the normal operating temperatures of wind turbines and regions where conditions are favorable for ice growth on wind turbine blades. The BTM World Market Update 2012 included a chapter on cold climate turbines,[2] where it was estimated that 24% of the currently installed turbine capacity is located in cold climate regions. Heavy icing regions, where icing is expected to have a major impact on annual energy production, contain 4% of installed turbines worldwide. The wind-based electricity production in moderate to heavy icing regions is expected to grow from 11.5 GW at the end of 2012 to 19.5 GW by 2017.[2] Therefore, accurate estimates of icing production loss are needed both for the proper siting of these new wind farms and to provide more accurate production forecasts for power system operation and electricity markets. Improved production forecasts in turn provide better returns to developers (because of lower regulation costs in electricity markets), while better supporting power system operators in their various tasks.[3]

Before modeling icing-related power loss, one first has to determine whether the power loss is due to icing or other factors. Homola *et al.*[4] provided a comprehensive review of ice sensors for wind turbines. These included direct sensors placed on a nearby mast or the turbine nacelle and indirect sensors such as webcam images. They concluded that ice measured using direct sensors did not match the periods of decreased power production. The use of webcam imagery was the best approach; however, the lack of available light in winter was a limiting factor. For the sites used in this study, webcam imagery was not available, thus an indirect and data-driven approach for identifying icing was developed, based on a deviation from the turbines' empirical power curve. Relying on power measurements for identifying icing is reasonable when studying the impact of icing on power production. However, if ice detection is required for other purposes, such as the safety of maintenance personnel, other approaches may be more appropriate.

To capture the impact of icing on power loss, a physically based ice model is needed to estimate the amount of ice on the turbine blades. The ice model needs to simulate both ice growth and ice ablation processes. Atmospheric ice growth occurs when atmospheric water in any phase sticks to and freezes on an exposed structure.[5] Ice ablation refers to the removal of ice by any method. Common ice ablation methods are melting, sublimation, shedding, and erosion. Melting is a phase change process where ice converts to liquid water and occurs when the ambient temperature is above 0 °C. Sublimation, the conversion of ice to water vapor, depends on atmospheric humidity. Ice shedding occurs when large pieces of ice fall off or are thrown from a structure, while erosion occurs because of the abrasive force of the wind removing small pieces of ice. This is the same process that erodes soil and rocks over time.[6,7]

The iceBlade model, described in Davis *et al.*,[1] was used as the icing model for the current study. It includes a modified version of the Makkonen model[8] for growth of atmospheric ice on rotating cylinders. The modifications make it more applicable for wind turbine icing and include an adjustment of the incoming wind speed based on the rotational speed of the turbine, a heat transfer coefficient for airfoils, and a static cylinder diameter. The iceBlade model includes methods for sublimation and total shedding of ice and was improved in this study through the addition of a term for wind erosion.

Statistical approaches are generally preferred for estimating icing-related power loss because of the computational expense of physical approaches. The resulting statistical methods can be broadly divided into two different approaches. The first approach is to create a three-dimensional power curve, where power output is modeled directly as a continuous function of both wind speed and accumulated ice mass (e.g., Byrkjedal[9]). Alternatively, Karlsson *et al.*[10] presented a three-dimensional power curve that used ice duration rather than accumulated ice mass. The second approach employs statistical techniques to combine multiple fields from icing and meteorological models into a production loss estimate. The power loss estimate is then applied to a standard power production estimate. Baltsheffsky[11] designed a neural network to obtain production estimates that accounted for the impact of icing, using as input results from an icing model similar to iceBlade, as well as more common numerical weather prediction (NWP) model outputs. Davis *et al.*[12] fit a linear regression model using several output variables from both iceBlade and an NWP model, demonstrating that the mean bias (MB) and root mean squared error (RMSE) of production estimates could be reduced by including terms important to the icing process. The present study builds on a similar approach through the use of a generalized additive model (GAM).[13,14] A GAM is a generalized linear model that can use linear predictors that depend on unknown smooth functions, allowing them to be fit without relying on possibly rigid parametric restrictions. Since the relationship between icing parameters (e.g., ice mass, duration, and density) and turbine production is unknown, the replacement of the linear model used in Davis *et al.*[12] with a GAM is a logical next step.

This study builds a modeling system that improves the production estimate at wind parks with significant icing using NWP, icing, and power loss models. This system could potentially be employed either as a forecast tool or to aid in siting future wind parks. However, neither application is explored directly, as this study is a proof of concept. This study will also explore the possibility to obtain a general model for estimating icing-related production losses by fitting two models to the wind park power loss. The first includes only data from the wind park being modeled, while the general fit uses data from all six wind parks. This latter strategy would eventually allow for the application of the general model to new and existing wind farms where the necessary data for local estimation (e.g., past power production) may not be available.

Wind Energ. 2016; **19**:1503–1518 © 2015 John Wiley & Sons, Ltd.
DOI: 10.1002/we

The remainder of the paper is laid out as follows. Section 2 describes the observational datasets available for this study, as well as the data pre-processing. Subsequently, Section 3 presents both the meteorological and revisited iceBlade models, as well as a detailed description of the proposed statistical approach to the modeling of icing-related power loss. Section 4 gathers the results from our empirical investigation, while Section 5 contains the discussion and a set of concluding remarks.

2. INDIRECT ICING IDENTIFICATION IN AVAILABLE DATA

2.1. Datasets

Observations were collected from a subset of turbines at six wind parks in Scandinavia. For confidentiality reasons, the parks are anonymized, and the power production was normalized by the rated power. All parks were equipped with Vestas V90 turbines and had reduced production during the winter that was at least in part caused by icing. This dataset is a subset of a larger dataset being used for an inter-comparison study as part of the Icewind project[15] and was selected to ensure uniformity across the turbine type. Even though only V90 turbines were used, the method presented should be applicable across different turbines, with small changes to the turbine-specific variables in the iceBlade model. The parks are identified with capital letters relating to the full Icewind dataset. The data cover two winter seasons from June 2010 until July 2012 and include 10 min records of nacelle-measured temperature and wind speed, power output of the turbine, and several quality assurance (QA) flags from the SCADA system. The QA flags were used to remove data when the turbine was not in normal operation. This was typically due to winds below the turbine's cut-in speed or when the turbine was placed in a derated (curtailed) state. The criteria for data removal were significantly strict that they would include periods where the turbine was out of normal operation for only 10 of the 600 s resolution of the data to also capture start/stop periods. Additional data were removed for wind speeds below 5.5 m s^{-1} because there was a large deviation in power production at lower wind speeds during non-icing periods. There should be minimal impact on the production loss estimates because of their removal, as the power output is small at low wind speed.

2.2. Inferring icing events

The observational dataset did not include a direct measure of icing; therefore, observed icing periods were determined empirically. A common approach for identifying icing is to find power and wind speed observations at cold temperatures with a significant deviation from the manufacturer's power curve (PC_m). These points are commonly identified using a flat 20% deviation. However, this approach will not identify all of the icing events that occur at higher wind speeds. The percentile deviation suggests that there is a larger variability in power at higher wind speeds, but the data show the opposite. At wind speeds above the rated wind speed, the turbine's control system uses the surplus energy to reduce the variance in power output leading to lower variability. In a previous study,[1] a simple curve was fit to transition from a 20% deviation below rated power to a 2% deviation above rated power. However, this curve only matched the data from a few turbines in this study. Therefore, a more statistically robust approach was developed for identifying icing using the standard deviation of the observed power.

An ice threshold curve (PC_i) was defined as a curve of wind speed and power that separates iced points from non-iced points when the temperature was below 0 °C. PC_i was based on the empirical power curve (PC_e) and standard deviation curve of each turbine. PC_e was used rather than PC_m to account for the local conditions at each turbine, since these effects were found to be significant for several turbines.

PC_e was fit for each turbine using a GAM with the formula

$$p = s(v) \tag{1}$$

where p is the observed power and $s(v)$ is a smooth function of the nacelle-measured wind speed (v). A thin plate regression spline was used as the smooth function. To ensure that the data used in the fitting process did not include iced points, only points with a temperature above 5 °C were used. The R package *mgcv*[16,17] was used for fitting the GAM, and the k-dimension, which limits the maximum degrees of freedom in the model, was set to 20. The k-dimension was determined by fitting several GAM models and visually comparing the fit with the cloud of observed power for temperatures above 5 °C. The model was found to be sensitive to values of k less than 13, as the effective degrees of freedom for all turbines was between 10 and 13 when k was set to 20.

The variance of power at different wind speeds was found by calculating the standard deviation of power for each 0.1 m s^{-1} bin of wind speed. The standard deviations were then smoothed using a LOESS smoother[18] with a span of 0.4. The smoothing removed outliers where there were fewer points, providing a more representative curve of the variance. PC_i was then calculated by

$$PC_i = PC_e - 2\sigma_s \tag{2}$$

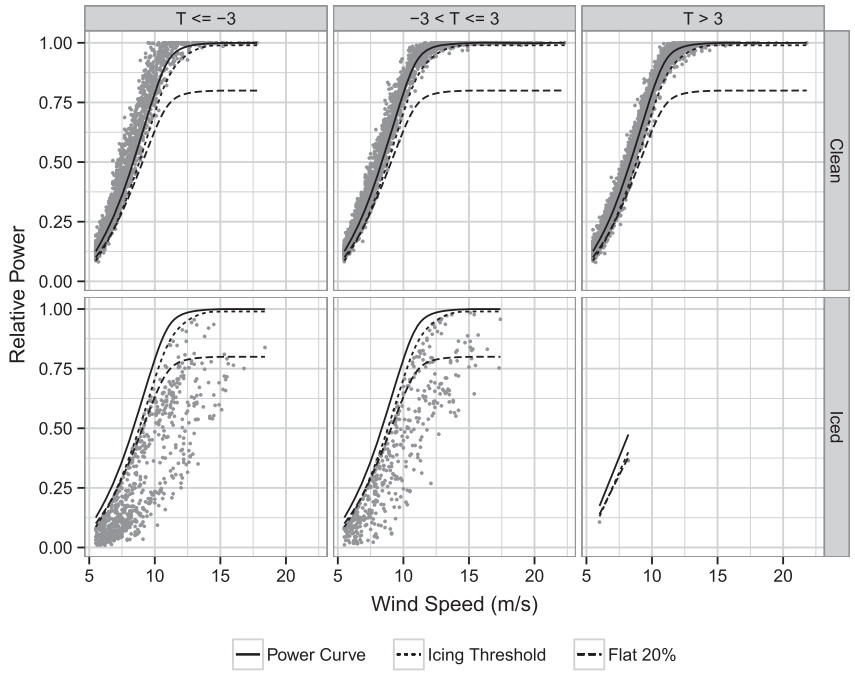

Figure 1. Observed power data and empirical power curve (PC_e) for a single turbine. The columns identify different temperature bins, while the rows separate points identified as clean or iced. The solid line is PC_e for points with temperature above 5 °C; flat 20% (long-dashed line) is a 20% deviation from PC_m, and Icing Threshold (PC_i; short-dashed line) is the threshold derived from the standard deviation-based approach.

where σ_s is the smoothed standard deviation at a given wind speed. At some points above rated power $\sigma_s = 0$, for these points $PC_i = 0.99 \times PC_e$. Figure 1 shows the observed power data, PC_e, and PC_i, for one of the turbines in the dataset. The top-right plot approximates the reference data for non-iced turbine production. This plot shows that the standard deviation-based method captures the shape of PC_e better than the 20% threshold.

After fitting PC_i, icing points were identified in a three-step process. First, preliminary iced points were identified as points when the temperature was below 3 °C, the turbine was operating normally, and the power was below PC_i. The second step adjusted the ending time of icing to include points removed during the QA process and points that were not considered icing by the temperature threshold. This limited the end of an icing event to the point when the power returned above PC_i. Finally, a criterion was added that required an icing event to have a minimum duration of 30 min (three consecutive data points). As seen in the top row of Figure 1, there are several points in all temperature bins that are below PC_i, because of this requirement. From the iced point plots (bottom row), it could be seen that the new approach identifies many icing points above 10 m s^{-1} that would not be captured by the flat 20% deviation method. Since no direct measurement of icing was available, a more robust evaluation of these two approaches was not possible.

After the QA and icing detection were carried out using 10 min values, the data were aggregated into park-averaged values at the top of the hour (minutes equal to 0) to match the output of the meteorological model (Section 3.1). Since icing events are typically more than 1 h in duration, this should have limited impact on the analysis. Park averaging was performed differently for the continuous variables, the QA flag, and the icing flag. For the continuous variables (temperature, power, and wind speed), the arithmetic mean of all non-QA turbines in the park was used. The park's QA state was determined based on a threshold of 50% of the turbines being in a QA state, while park icing was identified if two or more of the turbines were determined to be iced. An empirical power curve for the entire wind park (PC_{park}) was fit using equation (1), for all points with a mean park temperature above 5 °C.

3. DESCRIPTION OF THE METEOROLOGICAL AND STATISTICAL MODELS

All models were run for the 2 years of data from June 2010 until July 2012, with hourly temporal resolution. The meteorological model and icing models were evaluated for the same time period. However, for the statistical power loss models, the first year of data from June 2010 until July 2011 was used for model fitting, and the second year of data was used for model evaluation.

Wind Energ. 2016; **19**:1503–1518 © 2015 John Wiley & Sons, Ltd.
DOI: 10.1002/we

3.1. Meteorological model

The NWP output is from a global modeling study carried out at Vestas Wind Energy Systems. The study was undertaken to aid in understanding the long-term wind climatology at specific locations. The simulation started on January 1, 2000 and was run for 54 h with a 6 h spin-up. The meteorological variables targeted in the setup were wind speed, wind direction, temperature, and density. However, the entire model output was archived allowing for this dataset to be utilized for ice-related studies.

The Weather Research and Forecasting (WRF) model[19] was run at a 3 km grid resolution with 62 vertical levels, 17 of which were located within the lowest 1 km of the surface. The input and boundary conditions were from the Global Forecast System analysis,[20] and the boundary conditions were updated every 6 h. The physical parameterizations used were the WRF single moment five-class microphysics scheme,[21] rapid radiative transfer model for longwave radiation,[22] Dudhia shortwave radiation scheme,[23] Eta similarity surface layer physics, Noah land surface model,[24] MYJ planetary boundary layer (PBL) scheme,[25] and Kain–Fritsch cumulus scheme.[26]

The following outputs from WRF were provided from a grid cell near the center of each of the wind parks for the following surface variables: terrain height, surface shortwave and longwave radiation, and precipitation for the years 2010 through 2012. The impact of using a single-grid cell was investigated in Davis *et al.*[1] and found to not have a large impact on the model results. Height-dependent model output was provided for wind speed, temperature, pressure, water vapor mixing ratio, and four mixing ratios representing different cloud parameters (hydrometers) produced in the model microphysics (cloud water, rain, ice, and snow). These fields were interpolated from the model levels to heights of 40, 80, 120, 160, and 200 m above the model topography.

3.2. Icing model

The icing model used in this study is the iceBlade model described in Davis *et al.*[1] with a few modifications. The original version of the iceBlade model represents the turbine blade as a 1 m long cylinder with a 0.144 m radius, based on a NACA64_A17 airfoil with a chord length of 3.0 m. The airfoil geometry was taken from the NREL 5 MW offshore wind turbine.[27] The inflow wind speed for the iceBlade model is based on the turbines' rpm at a distance of 85% of the blade length. The meteorological conditions remain constant for the entire hour between the meteorological model output times. Ice is modeled to form straight out from the leading edge, retaining a constant diameter. Ice ablation occurs through sublimation and shedding algorithms. The shedding algorithm removes all ice from the blade when the temperature reaches 0 °C. The low threshold temperature is in part a result of the underestimation of temperature, commonly called a cold bias, in the WRF model shown in Section 4. While the 0 °C threshold worked reasonably well at all six sites in this study, it may need to be adjusted in the future for different climates, turbines, or NWP models. The iceBlade model was driven using only the 80 m WRF results after an analysis showed minimal differences when other vertical levels were used (not shown).

Modifications to the original version of the iceBlade model were required for the parks in the study. It was found that at several of the park's periods of icing were not identified by the iceBlade model. A previous study[1] found that the combination of WSM5 microphysics and the MYJ PBL schemes, used in the global WRF simulation, produced large amounts of cloud ice compared with other setups. Based on these results, the cloud liquid water content was calculated as

$$\mathrm{LWC} = (q_c + q_i) \times \rho_{air} \tag{3}$$

where q_c is the cloud water mixing ration, q_i is the cloud ice mixing ratio, and ρ_{air} is the air density. Including q_i improved the model results for the icing periods and should better match the actual cloud composition.

The second adjustment to the original iceBlade model was to the ice ablation rate. The dominant ice removal method in iceBlade at temperatures below 0 °C is the sublimation of ice. At these temperatures, sublimation is a relatively slow process as the saturation vapor pressure is low. The higher the saturation vapor pressure, the faster ice can sublimate. The observed power production data recovered quickly, even during periods where the temperature was well below 0 °C, suggesting that the ice was ablating quickly. We believe this is due to the wind erosion of ice on the blade. A literature review did not uncover any studies of wind erosion specific to ice on structures. Wind erosion is caused by both the force of the wind removing loose ice particles and the impact of wind blown particles breaking off pieces of the ice.[6] Therefore, the wind kinetic energy was chosen as the main input to the wind erosion term. A range of values was tested, and the model performance was evaluated using the equitable threat score (ETS), false alarm rate, and miss rate.[28] The optimal fit was found to be

$$E = 5 \times 10^{-6} v^3 \tag{4}$$

where E is the wind erosion kg s^{-1} and v is the wind speed. The ETS was relatively insensitive to the value of the fit.

The iceBlade model was run under two settings: blade cylinder and standard cylinder. In the standard cylinder setting, the enhanced velocity from the blade rotation was turned off and the cylinder size was decreased from 0.144 to 0.03 m. From the iceBlade model, the accumulated ice mass (ICE_g), ablated ice mass (ICE_r), total ice mass (ICE_t), and an ice flag were produced at each time step. ICE_t is the running total of ICE_g and ICE_r. The ice flag used a threshold of 0.1 kg m^{-1} to denote when ice was on the blade.

In addition to the iceBlade model, a simplified ice model was created using temperature and relative humidity. This model only included an ice flag that identified ice as any time the WRF model temperature was below 0 °C and the relative humidity was greater than 95%.[29] This value will be used as a reference value for comparison with the physical model results.

3.3. Statistical power loss models

Instead of modeling power including icing effects, the power loss caused by icing was modeled. This approach allows for state-of-the-art methods for power production to be applied separately from the icing-induced power loss. The power loss, p_l, was defined as

$$p_l = p_{park} - p \tag{5}$$

where p is the observed power and p_{park} is the power based on PC_{park} as defined in Section 2.

A hierarchical modeling approach[30] was implemented to allow the fitting of different models for iced and non-iced periods. The hierarchical model was a combination of a decision tree[31] and a GAM. The decision tree was pruned at the first level based on the iceBlade ice flag. The ice flag was true when $ICE_t > 0.1$ kg m^{-1}. Separate GAM models were fit for the true and false ice cases. GAMs were chosen since their semi-parametric methodology is useful given the uncertainty of the relationship between the various iceBlade outputs and the power loss.

The GAMs were built using a combination of known physical relationships and statistical analysis. Initially, all of the icing and relevant meteorological terms were correlated with the power loss and each other using Kendall's rank correlation.[32] This helped to identify the key terms for fitting the model. Several terms with high correlation to power loss were dropped, e.g., the specific humidity, as they also had high correlation with more relevant predictors such as temperature. After this initial screening, a cross-validation approach[33] was used to select the terms of the model, analyze the difference between the various smooths available for GAMs, and select the k-dimension for each term.

For the cross-validation, the first year of data was randomly split into 12 sections, each with approximately 500 data points. The sections were defined to ensure a similar ratio of observed ice to non-iced points as in the entire first year dataset. Eight of the sections were combined for use as the training dataset, with the remaining four sections being used as the test dataset. By selecting all possible combinations for the test dataset, a total of 495 test datasets were able to be used for validation, which can be found using the four combination of the 12 sections. Figure 2 shows density functions of the MB and RMSE that were plotted to study the model performance.

The MB is defined as

$$\mathrm{MB} = \overline{(m - o)} \tag{6}$$

where m is the model result and o is the observed value, and the overbar signifies the arithmetic mean. Even though power loss was the modeled term, the cross-validation and later model evaluation was performed against the observed power rather than the observed power loss, allowing for comparisons with the non-iced estimate (PC_{park}). The estimated power loss was converted back to power as

$$p_e = p_{park} - p_{el} \tag{7}$$

where p_e is the estimated power and p_{el} is the estimated loss.

Using cross-validation, to evaluate the models, the final model was found as,

$$p_{el} = \begin{cases} s(v) + s(i_g) + s(i_t) & \text{if } ICE_t > 0.1 \\ s(T) + s(p) & \text{otherwise} \end{cases} \tag{8}$$

where $s(v)$, $s(i_t)$, $s(i_g)$, $s(T)$, and $s(p)$ are smooth functions of wind speed, ICE_t, ICE_g, temperature, and atmospheric pressure. The smooth functions were fit using thin plate regression splines[34] with k-dimensions of 15 for $s(v)$, 40 for both $s(i_t)$ and $s(i_g)$, and 7 for both $s(T)$ and $s(p)$. The k-dimension in a GAM sets the upper limit on the degrees of freedom

Wind Energ. 2016; **19**:1503–1518 © 2015 John Wiley & Sons, Ltd.
DOI: 10.1002/we

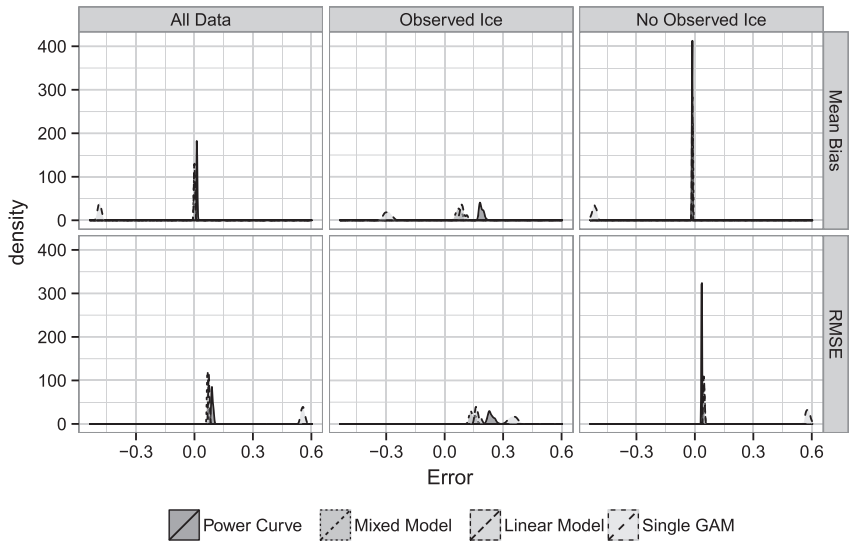

Figure 2. Density plots of power MB (top row) and RMSE (bottom row) for 495 cross-validation simulations for Park A and the first year of data. The columns show different splits of the dataset. The colors and line-types signify four different estimates of the power. Power curve is PC_{park}; mixed model adjusts PC_{park} using the hierarchical model with GAM fits; linear model is the hierarchical model with linear fits, and single GAM is fit without using the categorical model.

given by the smoothing functions. Since the empirical power curve already provided a good estimate of the non-iced data points, the non-iced GAM was not studied in detail.

The mixed model approach significantly outperformed a single GAM, by improving the iced points without significantly reducing the performance of the non-iced points (Figure 2). The single GAM shown in the figure used

$$p_{el} = s(v) + s(i_g) + s(i_t) \tag{9}$$

which was the same as the portion of the final model used during icing conditions, but the results were similar for other GAM models as well. The mixed model approach benefits greatly from the correct identification of icing periods, and therefore, from the changes made to the iceBlade model ablation algorithm previously described in Section 3.2.

After the model parameters were determined, an evaluation was created for the second year of data by training it on the entire first year of data. To investigate the applicability of a general model, two different sets of data were used to fit the final model. The first model used data from each individual wind park to determine the fit (PARK). The second fit used the data from all of the wind parks (ALL). This comparison allowed for the determination of how independent the response of the different parks was to the statistical model. If the relationship is found to be universal, the model could also be used for siting of turbines in nearby or similar areas without existing turbines.

4. RESULTS OF THE EMPIRICAL INVESTIGATION

The results from the component models, WRF and iceBlade, as well as the results from the statistical power loss models were evaluated. The WRF model was evaluated for temperature revealing a cold bias. The cloud mixing ratios could not be verified, as there were no cloud-related measurements at the wind parks. The iceBlade model was evaluated by studying both time-series and statistical measures. The evaluation highlights the improvements of the iceBlade model over the standard cylinder approach, as well as the importance of the wind erosion term on the model results. An example of the cross-validation analysis shows the improvement of the fit using the new approach. Finally, the results from the power loss modeling are presented to show the changes in RMSE and MB.

4.1. WRF results

Figure 3 compares the park-averaged nacelle temperature to WRF-simulated temperature at each of the wind parks. The WRF model does a reasonable job capturing the overall signal of the temperature, with a high correlation (Table I) and a relatively tight grouping of points around the regression line. However, there is a cold bias of approximately 2.5 °C at each park that is larger at colder temperatures.

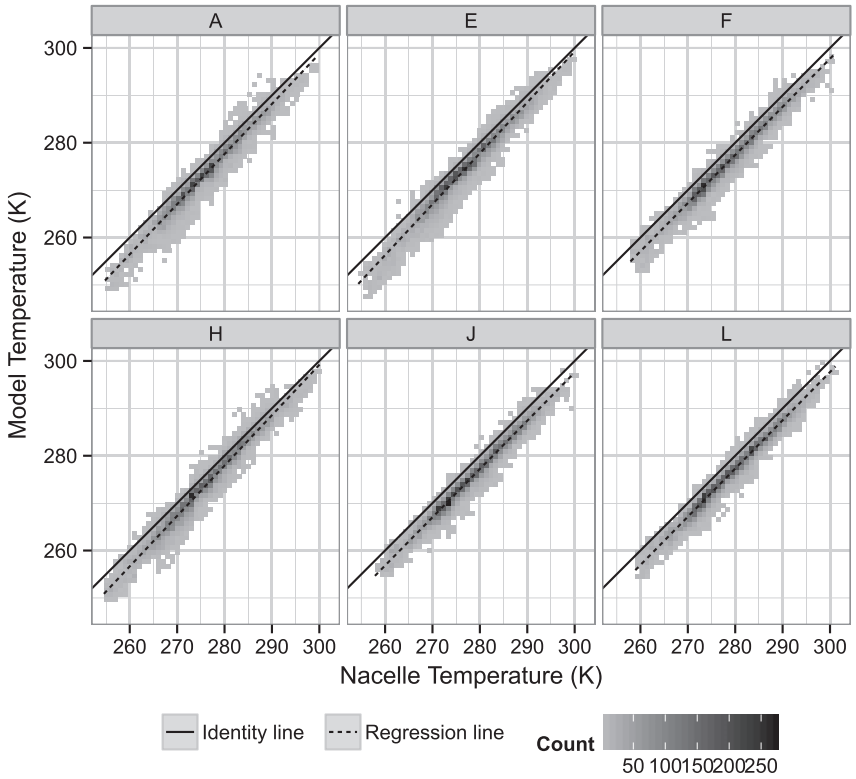

Figure 3. Comparison of nacelle and WRF-modeled temperature, interpolated to 80 m above ground level, at each wind park (A–L). Regression (dashed line) and identity (solid line) lines are shown to aid in the evaluation. The data are binned into 1 by 1 °C cells and shaded by the number of points in the bin (count).

Table I. MB, RMSE, Pearson's correlation coefficient (Cor_R), linear slope (slope), and RMSE with the MB removed from the model data (RMSE_adj) of WRF-simulated temperature compared with the nacelle-averaged temperature for each wind park labeled A–L.

Park	MB (°C)	RMSE (°C)	Cor_R	Slope	RMSE_adj (°C)
A	−2.72	3.22	0.98	1.06	1.72
E	−2.58	3.17	0.98	1.08	1.85
F	−2.64	3.00	0.99	1.02	1.43
H	−2.37	2.96	0.98	1.06	1.76
J	−2.82	3.13	0.99	1.02	1.36
L	−2.72	3.02	0.99	1.02	1.33

Table I includes comparison statistics of the modeled and observed temperatures. RMSE_adj is defined as

$$RMSE\_adj = \sqrt{\overline{[(m - MB) - o]^2}} \tag{10}$$

The large difference between RMSE and RMSE_adj shows that much of the model error is a result of the bias with an RMSE_adj around 1.5 °C at each park. However, as conditional bias correction was not performed, the large slope of the temperature bias would lead to an over-correction of warmer points and under-correction at colder points.

4.2. iceBlade results

Since the ice observations were defined as a simple yes/no flag, a contingency table[28] was used to evaluate the iceBlade ice flags that were created using a threshold mass of 0.1 kg m^{-1}. A contingency table (Table II) compares modeled

Wind Energ. 2016; **19**:1503–1518 © 2015 John Wiley & Sons, Ltd.
DOI: 10.1002/we

events with observed events at each hour and evaluates how many times a result is a *true positive* (a), *false alarm* (b), *miss* (c), or *true negative* (d). From the contingency table, several scores can be calculated. The false alarm rate (F), defined as

$$F = \frac{b}{b+d} \tag{11}$$

and the miss rate (M), defined as

$$M = \frac{c}{a+c} \tag{12}$$

are scores that identify errors in the model and have a range of 0 to 1, where lower values are better. The ETS[28] takes into account both false alarms and missed events and is designed to provide reasonable scores for rare events, such as turbine icing. The ETS is defined as

$$ETS = \frac{a-a_r}{a+b+c-a_r} \tag{13}$$

where a_r is $(a+b)(a+c)/n$ and n is the total number of events. The ETS has a range of $-1/3$ to 1, with 1 being a perfect model and a value below zero suggesting that the chance estimate is more skilled than the actual model.

Figure 4 shows a time-series of the iceBlade-simulated ice mass, for a 1 m long section of the turbine blade at wind park A, with the periods colored by contingency table result.

The blade setting results in more periods of significant ice growth than the standard cylinder setting. This, in turn, leads to much larger values of ICE_t. For the standard cylinder setting, there are no hours where ICE_g is above the 0.1 kg m^{-1}

Table II. Contingency table.

Event modeled	Event observed	
	Yes	No
Yes	a (True positive)	b (False alarm)
No	c (Miss)	d (True negative)

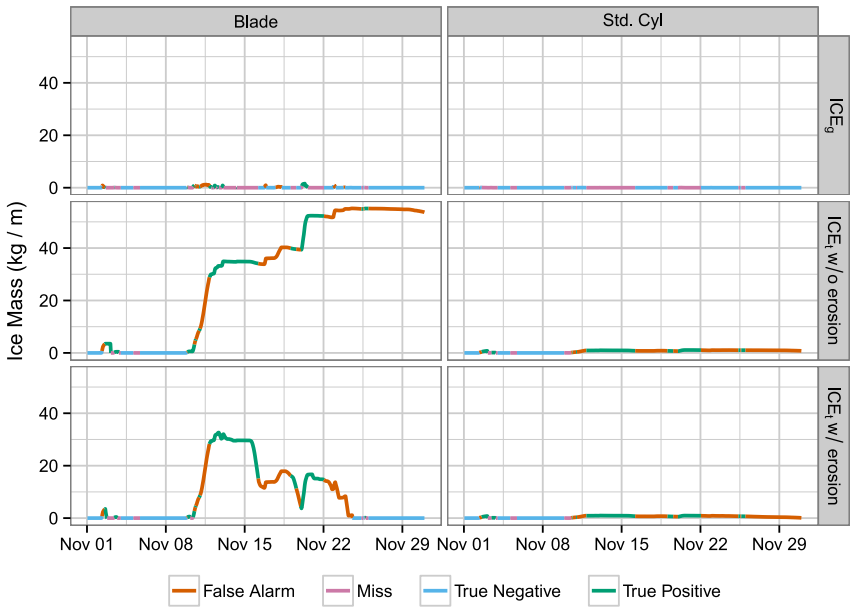

Figure 4. Time-series of modeled ice mass (kg m^{-1}) at wind park A for November 2010. The columns show the results using the iceBlade method (left) and the standard cylinder Makkonen model (right). The rows show ICE_g (top row), ICE_t without wind erosion (middle row), and ICE_t with wind erosion (bottom row). The colors denote the four possible outcomes of the model results compared with observations.

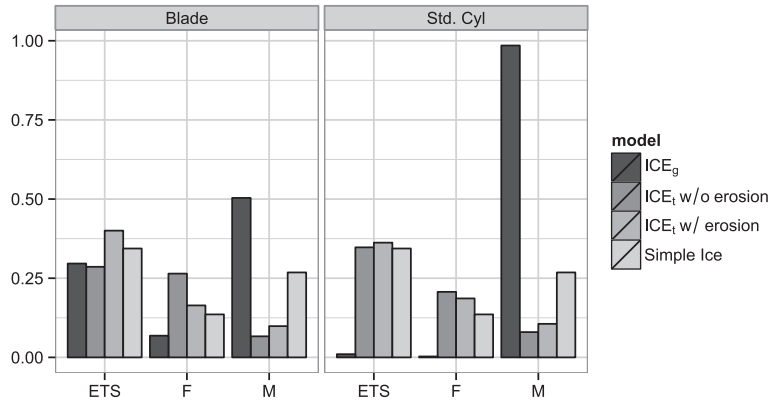

Figure 5. Values of the ETS, false alarm rate (F), and miss rate (M) for the icing period prediction at wind park A. The gray scale identifies the four icing model results for the blade (left) and the standard cylinder (right) geometries.

threshold, during this period. However, ICE_t shows an icing period that lasts from 9 November until the end of the month. The wind erosion term has a large impact on blade icing, but much less on the standard cylinder, because of the higher wind speeds with the blade icing setting. The increased ablation from the wind erosion term leads to fewer false alarms. However, even in results using the blade icing method, a large number of false alarms occur during this period.

Figure 5 shows M, F, and ETS for the entire 2 year period at park A. This figure includes results from the simple ice model described in Section 3.2. The simple ice model performed very well in the ETS and F scores but had a larger M than any of the ICE_t results. For both the blade and cylinder methods, ICE_g has a very high M and correspondingly low ETS, despite a low F. The inclusion of the wind erosion term reduces F but also increases M. The F and M values are similar in magnitude for ICE_t with wind erosion, which was a goal of the fitting algorithm for the wind erosion term. The large decrease in F compared with the smaller increase in M leads to a large improvement in ETS when the erosion term is added to the model. As in the time-series (Figure 4), the improvements from adding the wind erosion term are found to be much smaller for the standard cylinder. Based on this analysis, the blade method with wind erosion was the model chosen for the statistical modeling.

4.3. Statistical modeling results

As described in Section 3.3, Figure 2 shows an example of the cross-validation density plot used to evaluate the different models of power loss. The density plot shows the probability density function of the model error from the 495 model fits created during the cross-validation. A density plot can be thought of as a histogram with a very small bin width so the peaks of the density surface identify the values that occur most frequently. The comparison shown includes the final mixed model fit (mixed model) equation (8), a model with the same terms but where the GAM is replaced by an ordinary least squares linear model (linear model), PC_{park} (power curve), and the single GAM model fit using equation (9) (single GAM). The linear and mixed models show similar results for the mean model bias, with large overlap of the density curves. However, the mixed model has significantly lower RMSE for the observed ice times and lower RMSE overall. As expected, PC_{park} had very high bias and error, because of the lack of icing correction. Both model fits show significant improvement compared with PC_{park} for the iced data. However, for the non-iced points, the addition of the statistical model increases both MB and RMSE. Since the improvement for the icing points is much larger, the total dataset shows a substantial decrease in error as well. As a comparison, the single GAM performs very poorly in the cross-validation. We were unable to find any relevant parameters for icing that improved the estimate compared with PC_{park}. This is due to the multidimensionality of the problem. By limiting the icing parameters to only those hours when they are active, a better fit can be found.

Figure 6 shows density plots of the hourly absolute error for the final model fit, with PARK showing the model fit to data only at that park, and ALL was fit using data from all parks. While in Figure 2, the densities were based on the results from many model fits; here, they are based on each modeled hour. For the first year, the PARK model always performed best. However, at parks E, F, and L, the ALL model results are similar to the PARK model. This was expected as the fit should be more accurate for individual parks when only their data was used. However, for the second year, the ALL model performs best at five of the six parks. In the first year, the RMSE is improved at all parks for both PARK and ALL, compared with PC_{park} (Table III). In the second year, PARK performs the same or worse than PC_{park} at five of the six parks for RMSE. Site E presents an interesting case for comparison between the bulk statistics, such as RMSE, and the information contained in the density plot. For the second year, both PARK and ALL errors were more likely to fall below 0.1 than the power curve

Wind Energ. 2016; **19**:1503–1518 © 2015 John Wiley & Sons, Ltd.
DOI: 10.1002/we

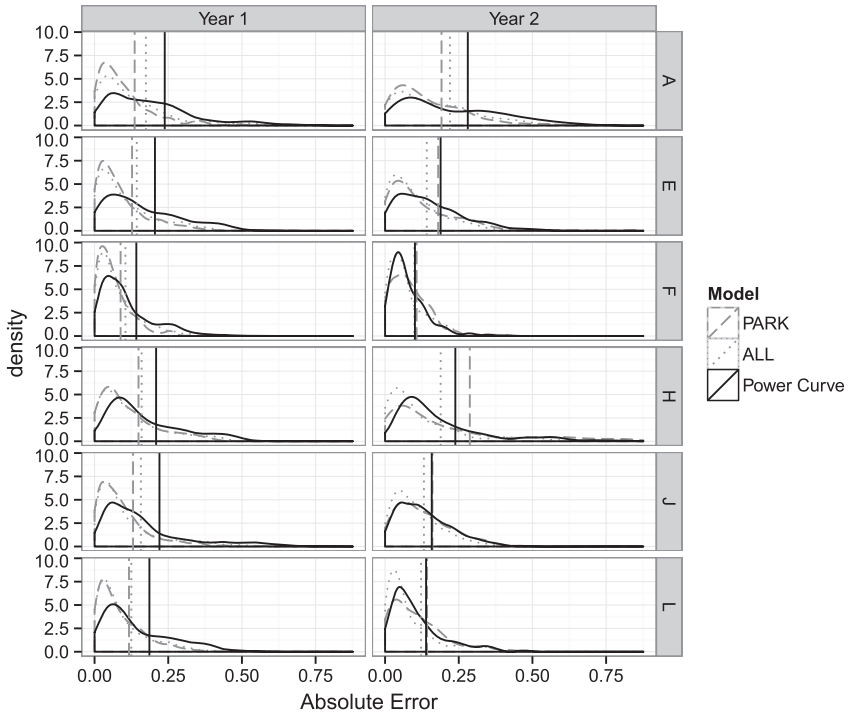

Figure 6. Density plot of absolute error from the power prediction model for points with observed icing at each wind park for PC_{park} (power curve; solid black line), the model fit for park (PARK; long dashed gray line), and the model fit with all parks (ALL; dotted gray line). The vertical lines show the RMSE value. Year 1 is the year used for fitting the models, and Year 2 is the verification year.

Table III. RMSE and MB for the unadjusted power curve, the model fit to the park (PARK), and the model fit to all parks (ALL) for hours with observed icing (Hours). Minimum values for each metric and park are in bold.

		Year 1			Year 2		
Park	Model	RMSE	MB	Hours	RMSE	MB	Hours
A	PARK	**0.14**	**0.07**	669.00	**0.19**	**0.14**	741.00
A	ALL	0.17	0.12	669.00	0.22	0.17	741.00
A	Power curve	0.24	0.19	669.00	0.28	0.23	741.00
E	PARK	**0.13**	**0.06**	691.00	0.18	**0.01**	751.00
E	ALL	0.14	0.08	691.00	**0.14**	0.09	751.00
E	Power curve	0.20	0.16	691.00	0.19	0.15	751.00
F	PARK	**0.09**	**0.04**	774.00	0.11	**0.05**	330.00
F	ALL	0.10	0.05	774.00	**0.10**	0.05	330.00
F	Power curve	0.14	0.11	774.00	0.10	0.08	330.00
H	PARK	**0.15**	0.10	534.00	0.29	0.19	469.00
H	ALL	0.16	**0.10**	534.00	**0.19**	**0.11**	469.00
H	Power curve	0.21	0.17	534.00	0.24	0.18	469.00
J	PARK	**0.13**	**0.06**	882.00	0.16	0.12	555.00
J	ALL	0.16	0.08	882.00	**0.13**	**0.10**	555.00
J	Power curve	0.22	0.17	882.00	0.16	0.13	555.00
L	PARK	**0.12**	0.07	743.00	0.14	0.10	459.00
L	ALL	0.12	**0.06**	743.00	**0.12**	**0.07**	459.00
L	Power curve	0.19	0.14	743.00	0.14	0.11	459.00

method, but PARK produced a few points that had very large errors. When the errors were averaged, PARK and PC_{park} had the same RMSE, while the RMSE was reduced for ALL.

Table III also shows the number of hours with observed icing at all sites. Parks F and J had significant changes in the number of icing hours from 1 year to the next. It is important to point out that the number of icing hours did not always

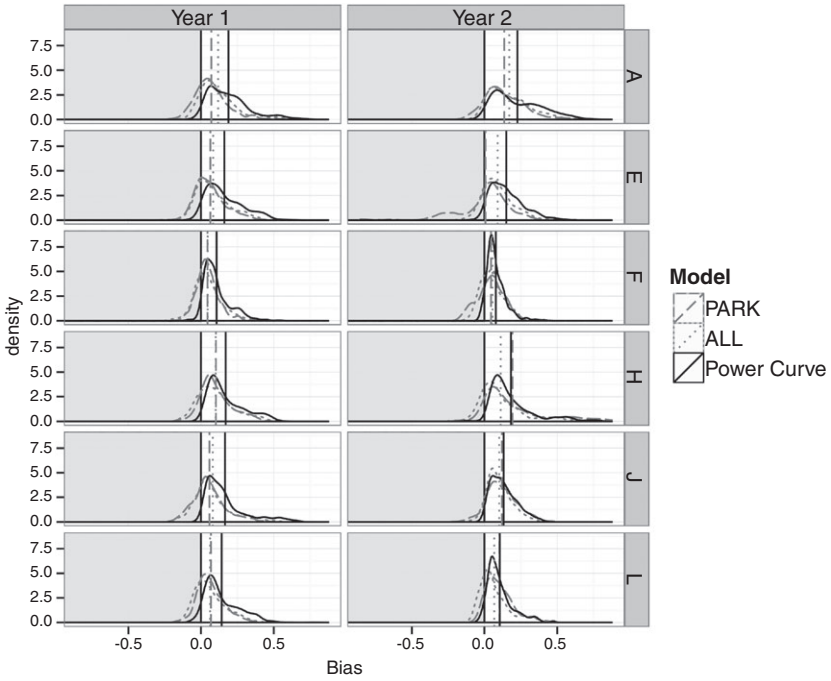

Figure 7. Same as Figure 6 but for bias and MB statistics. Light gray shading denotes bias values less than 0.

correspond to more icing-related production loss. For example, at site E, Year 2 had less icing, yet PC_{park} showed a smaller MB in Year 2.

Figure 7 shows the corresponding model bias. Both PARK and ALL shift the distributions so they are more centered on zero compared with PC_{park}, and the MB generally has the same trend. As with the error plot, the performance of PARK and ALL varied greatly depending on the site and year being modeled, and PARK generally performed better for the first year, while ALL generally performed better for the second year. The conflicting results between the bulk and comprehensive error analyses for Year 2 at Park E in the error plot were also found in the bias results. Here, PARK had a large number of bias values less than −0.2, signifying a fairly large bias on either side of zero. However, the MB was only 0.01, as the large positive and negative biases canceled out. ALL has fewer points far away from zero, at site E, but has a larger MB than PARK. Provided with both the density plot and bulk statistics, it is clear that the ALL model is the best model at this site. However, when examining only the bulk statistics, PARK and ALL perform similarly.

5. DISCUSSION AND CONCLUSIONS

Modeling icing occurrence on a wind turbine is a challenging task. This study has shown that the iceBlade model, driven by inputs simulated using the WRF model, can reasonably capture periods of icing at several wind parks. Additionally, the error in estimated power can be improved by using a statistical model to relate icing to the power loss.

5.1. Simulated cold bias

The WRF-simulated temperatures were found to have a cold bias of approximately 2.5 °C compared with the nacelle temperature (Table I). This is consistent with results found in other studies using the WRF model.[35,36] The temperature bias has a large impact on the simulation of icing, as both the ice accretion and ice ablation terms are highly sensitive to temperature, particularly around 0 °C. The cold bias was approximately 0.5 °C greater during periods of ice accretion than the values in Table I. The impact of the colder temperatures is one factor in the rapid ice accretion that occurs on 10 November when the observations did not show any icing signal (Figure 4). During this event, the observed temperature was between −5 and −7.5°C, while the WRF-simulated temperature was approximately 2.5 °C colder throughout the period. While both temperatures are cold enough for ice to form, the colder temperatures in WRF would lead to a larger percentage of incoming water to freeze on impact.

Wind Energ. 2016; **19**:1503–1518 © 2015 John Wiley & Sons, Ltd.
DOI: 10.1002/we

There are three possible sources of the cold bias. The first is a warm bias in the nacelle temperatures, compared with the ambient temperature. The nacelle temperatures are known in the industry to have a wide range of accuracies, and the large number of observed icing points that occur between 0 and 3 °C (Figure 1) suggests a warm bias in the measurements. Inaccurate NWP model terrain height is the second potential source of the modeled temperature bias. However, it was found that the WRF model terrain heights were always lower than the observed heights, leading to warmer modeled temperatures Finally, the WRF parameterizations used in this study may be contributing to the cold bias. An earlier study investigating different WRF physics parameterizations[1] found that the combination of the WSM5 microphysics and MYJ PBL scheme had a significantly larger cold bias than other schemes. Therefore, the accuracy of the model temperature estimate could be improved by using different physics options.

5.2. Ice modeling

The results of iceBlade were evaluated against a binary dataset for icing that was created using an ice threshold curve as described in Section 2. In Davis *et al.*,[1] the blade setting was found to significantly outperform the standard cylinder setting for detecting periods of icing. However, in this study, these two methods produced similar ETS scores, with the blade setting only slightly outperforming the cylinder setting after the wind erosion term was added (Figure 5). However, ICE_t was significantly different between the two methods (Figure 4). The larger and more dynamic ice mass values of the blade setting consistently returned better estimates of the production loss. The standard cylinder tends to produce more periods of ice growth, but when ice does form on the blade, the accumulation is larger. More ice takes longer to be removed leading to longer total icing periods. Therefore, while the overall skill scores of the two models are similar, the periods of icing are quite different.

The inclusion of the wind erosion term was key to the performance of the blade setting. Without the wind erosion term, the blade setting included a significant number of false alarm events and overall produced too much ice (Figure 5). This highlights the importance of accurate modeling of the ice ablation process. It should be noted that the algorithm for wind erosion included in this study is uncertain. Wind erosion can be caused by two factors: the force of the wind blowing loose pieces of ice off the blade and the abrasive force of particles in the atmosphere contacting the ice and breaking off pieces of the ice. Neither of these forces are well understood for turbine icing; therefore, an empirical fit was used. However, because there were no direct ice mass measurements on wind turbines, the empirical fit was made using only the iced/not-iced flag created from the power observations. This provided little input when trying to fit a function to the complex time-series of the ice mass shown in Figure 4.

McDonough *et al.*[37] presented an icing model inter-comparison that showed a strong agreement for the onset of icing between the various models but a large spread in the results for the ending of an icing event. The uncertainty in the ending of ice events agrees with the results presented here and suggests that more understanding is needed into the ablation of ice on wind turbines. Given the importance of icing to the wind industry, a dedicated measurement campaign into ice ablation would be the ideal setup for understanding how this process occurs on a turbine blade. In addition, because the power threshold approach cannot be used when turbines are not in normal operating mode, the measurement of icing directly on the turbine blades is also important for the control of deicing systems and health and safety evaluations.

5.3. Power loss modeling

Given the large uncertainties in the physical icing model, the modeling of ice-related power loss is a challenge. The use of a hierarchical modeling approach was found to be a large improvement over a one-size-fits-all model (Figure 2). The problem presented by icing is an application that is well suited for the hierarchical approach, as the parameters that affect power during icing, mostly the ice mass and shape, do not have any impact on the power loss during non-icing periods. Of course, this requires the physical models to do a reasonable job of capturing periods with icing

There is no consensus on the parameters that should be used to best model power loss in the literature. Some studies have relied on ICE_t,[9] another study suggested that only periods of ICE_g were important,[38] still others have suggested using the duration of icing.[10] Our approach was similar to the approach of Baltsheffsky[11] in that a large number of terms were included, and statistical analysis aided the final selection. In the end, only ICE_t, ICE_g, and the wind speed remained significant. Many other parameters were highly correlated with the ice mass terms and removed to avoid duplicating information. Given the importance of ICE_t and ICE_g, it can be assumed that as the physical models improve, the power loss model should also improve. Therefore, the best way to improve the prediction of production loss due to icing is to invest in improvements to the NWP models, physical icing models, and observational data collection. It is expected that using an NWP model setup for the specific wind park locations and designed for estimating the parameters important for icing should also improve the prediction. Additionally, appropriate post-processing could be applied to the NWP model results to provide a point forecast.

It was somewhat surprising that the model including data from all wind parks during the fitting process (ALL) performed better than the model fit with only that parks data (PARK) for the second year (Figures 6 and 7). During the model fitting

process, testing different k-dimensions and variables, only the PARK method was used. An analysis of the smoothing terms found that they were fairly linear for ALL. However, for PARK, there were often high-amplitude curves in the ice mass smooths. This suggests that during the cross-validation the k-dimension was set too large for these terms and should be smaller in future work.

During cross-validation only MB and RMSE were used to evaluate the model. The final results for site E, however, demonstrated how these bulk statistics can be a misleading value when examining the model fit (Figure 7). This was particularly true of MB. At this site, PARK had the best MB score for Year 2; however, ALL had many more forecasts close to zero and far fewer large error forecasts. This highlights the importance of looking at all of the data when assessing model performance and not strictly relying on summary statistics. It is possible that the reliance on the bulk statistics lead to the over fitting of the PARK model.

While the model built for power loss showed good results, there are still enhancements that would be worth investigating. The first relates to the choice of power loss as the predicted parameter. While the modeling of power loss produced reasonable results, when converting it back to power, we had to limit the results to a value between zero and one, as the statistical model did not have knowledge of these limits. Additionally, as shown in Section 2, the position on the power curve is related to the magnitude of potential power loss. Therefore, modeling power directly could provide a better result. However, to do this, the mesoscale winds would need to be adjusted so they better match the observed winds. We chose not to do for this study, focusing instead on the icing impact.

Bernstein *et al.*[38] mentioned that, in their experience, production loss was most common during periods of active ice growth. However, when only ICE_g was used, there were many missed periods of lower power production (Figure 5). When fitting the statistical model, ICE_g was found to be significant, which suggests that the impact of ice on power production is different when the ice is growing than during periods of ablation. This was also seen in the observed dataset where there is a large initial drop in power at the onset of icing that diminishes with time. This could be in part due to an increase in roughness during active icing that is quickly eroded when the ice stops growing. Two attempts to model this behavior were undertaken in this study. First, a roughness parameter was estimated based on the ice density modeled by iceBlade, with the assumption that a lower density corresponded to a rougher surface. This term only had a slight correlation with power loss, potentially because it was only updated when ice was actively accumulating. Therefore, the impact of erosion and compaction on the roughness of the surface were not included. A roughness algorithm, which included these effects, would presumably provide a better estimate for the production loss model. The second attempt at capturing the change in production loss with time was by counting the hours since last ice accretion, but this term did not have a significant impact on the model fit. Another extension to the statistical modeling, could be the introduction of a probability of icing rather than a binary decision, this could reduce some of the uncertainty in the icing model results.

6. CONCLUSIONS

In this study we set out to answer two questions: could a model for power loss by icing be developed that improves the estimate of power and could a general model using results from multiple parks provide insight that could potentially be used for wind park siting? The estimate of power was improved by developing a statistical model for power loss using an icing model driven from inputs of a meteorological model. This model reduced both the bias and error of the estimate on a second year of data after being fit for the first year. Additionally, it was shown that fitting a model for a single park is not as useful as fitting for multiple parks; this likely is due to the fitting of multiple parks including a more robust signal compared with the site-specific noise. This suggests that given enough data for many wind parks, a model could be developed that would aid in the siting of wind parks, and provide better returns for developers, while supporting power operators.

ACKNOWLEDGEMENTS

This work was supported financially by the Top-Level Research Initiative (TFI) project, Improved forecast of wind, waves and icing (IceWind), Vestas Wind Systems A/S, and the Nordic Energy Industry.

The statistical modeling and data-analysis was carried out using the R Project for Statistical Computing.[17]

The authors would like to thank the anonymous reviewers of this paper for their valuable comments and suggestions that greatly improved the quality of this paper.

REFERENCES

1. Davis N, Hahmann AN, Clausen NE, Žagar M. Forecast of icing events at a wind farm in Sweden. *Journal of Applied Meteorology and Climatology* 2014-02; **53**: 262–281.

2. Wallenius T, Peltola E, Lehtomäki V, Antikainen P, Bluemink GJ, Wadham-Gagnon M, Dillingh J, Clausen NE, Klintsröm R. Special chapter: cold climate turbines ("CCTs"). In *BTM World Market Update*, BTM-Consult (ed.). Navigant Research: Chicago, Illinois, USA, 2013.
3. Giebel G, Brownsword R, Kariniotakis G, Denhard M, Draxl C, The state-of-the-art in short-term prediction of wind power. *Technical Report*, 2011. http://anemos.cma.fr.
4. Homola MC, Nicklasson PJ, Sundsbø P. Ice sensors for wind turbines. *Cold Regions Science and Technology* 2006; **46**: 125–131.
5. 12494:2001(E). *Atmospheric Icing of Structures*. ISO: Geneva, Switzerland, 2001.
6. Chepil W, Woodruff N. The physics of wind erosion and its control. *Advances in Agronomy* 1963; **15**: 211–302.
7. Laity JE, Bridges NT. Ventifacts on Earth and Mars: analytical, field, and laboratory studies supporting sand abrasion and windward feature development. *Geomorphology* 2009; **105**: 202–217.
8. Makkonen L. Models for the growth of rime, glaze, icicles and wet snow on structures. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 2000; **358**: 2913–2939.
9. Byrkjedal O. Mapping of icing in Sweden—on the influence from icing on wind energy production. *Winterwind*, Skellefteå, Sweden, 2012. [Online]. Availible: http://www.slideshare.net/WinterwindConference/3a-byrkjedal-icingkvt (Accessed 4 April 2012).
10. Karlsson T, Turkia V, Wallenius T, Miettinen J. Production loss estimation for wind power forecasting. *Winterwind*, Sundsvall, Sweden, 2014. [Online]. Availible: http://www.winterwind.se/?edmc=2792 (Accessed 19 February 2014).
11. Baltscheffsky M. Modelling of production losses due to icing for individual turbines in a wind farm—development of techniques for forecasting and site assessment. *Winterwind*, Östersund, Sweden, 2013. [Online]. Availible: http://windren.se/WW2013/04_Baltscheffsky_Magnus_Winterwind_2013.pdf (Accessed 28 June 2013).
12. Davis N, Hahmann AN, Clausen NE, Žagar M, Pinson P. Forecasting production losses by applying the Makkonen icing model to wind turbine blades. *The 15th International Workshop on Atmospheric Icing of Structures*, St. John's, Newfoundland and Labrador, Canada, 2013; 2-54–2-58.
13. Hastie TJ, Tibshirani RJ. *Generalized Additive Models* 1st ed. Chapman & Hall/CRC: Boca Raton, Florida, 1990.
14. Wood SN. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC: Boca Raton, Florida, 2006.
15. Davis N, Hahmann AN, Clausen NE, Pinson P, Žagar M, Byrkjedal O, Karlsson T, Wallenius T, Turkia V, Soderberg S, Baltscheffsky M. Icewind inter-comparison of icing production loss models. *Winterwind*, Sundsvall, Sweden, 2014. [Online]. Availible: http://www.winterwind.se/sundsvall-2014/presentations-2014/?edmc=2809 (Accessed 21 April 2014).
16. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2011; **73**: 3–36. DOI: 10.1111/j.1467-9868.2010.00749.x.
17. R Core Team. *R: a language and environment for statistical ccomputing*. R Foundation for Statistical Computing: Vienna, Austria, 2014. [Online]. Available: http://www.r-project.org (Accessed June 2014).
18. Cleveland WS, Grosse E, Shyu WM. Local regression models. In *Statistical Models in S*, Chambers JM, Hastie TJ (eds). Chapman and Hall: New York, 1992; 309–376.
19. Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang XY, Wang W, Powers JG. A description of the advanced research WRF version 3, 2008.
20. National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce. NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. 2012.
21. Hong SY, Dudhia J, Chen SH. A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Monthly Weather Review* 2004; **132**: 103–120.
22. Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough S. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research* 1997; **102**: 16663. DOI: 10.1029/97JD00237.
23. Dudhia J. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of the Atmospheric Sciences* 1989; **46**: 3077–3107.
24. Chen F, Dudhia J. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part II: preliminary model validation. *Monthly Weather Review* 2001; **129**: 587–604.
25. Janjić ZI. The step-mountain eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly Weather Review* 1994; **122**: 927–945. DOI: 10.1175/1520-0493(1994)122<0927.

26. Kain JS. The Kain–Fritsch convective parameterization: an update. *Journal of Applied Meteorology* 2004; **43**: 170–181. DOI: 10.1175/1520-0450(2004)043<0170.
27. Jonkman J, Butterfield S, Musial W, Scott G, Definition of a 5-MW reference wind turbine for offshore system development. *Technical Report February*, National Renewable Energy Laboratory, Golden, Colorado, 2009.
28. Wilks DS. *Statistical Methods in the Atmospheric Science* second. Academic Press: London, UK, 2006.
29. Fikke S, Ronsten G, Heimo A, Kunz S, Ostrozlik M, Sabata J, Wareing B, Wichura B, Chum J, Laakso T, Makkonen L, COST 727 : Atmospheric icing on structures measurements and data collection on icing : state of the art. *Technical Report 75*, MeteoSwiss, 2006.
30. Cohen J, Cohen P, West SG, Aiken LS. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (3rd ed.)* Erlbaurm: Mahwah, NJ, 2003,728.
31. Breiman L. *Classification and Regression Trees*. CRC Press: Boca Raton, Florida, 1984.
32. Kendall MG. A new measure of rank correlation. *Biometrika* 1938; **30**: 81–93.
33. Picard RR, Cook RD. Cross-validation of regression models. *Journal of the American Statistical Association* 1984; **79**: 575.
34. Wood SN. Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2003; **65**: 95–114. DOI: 10.1111/1467-9868.00374/.
35. Cheng WYY, Steenburgh WJ. Evaluation of surface sensible weather forecasts by the WRF and the eta models over the Western United States. *Weather and Forecasting* 2005; **20**: 812–821. DOI: 10.1175/WAF885.1.
36. Heikkilä U, Sandvik A, Sorteberg A. Dynamical downscaling of ERA-40 in complex terrain using the WRF regional climate model. *Climate Dynamics* 2010; **37**: 1551–1564. DOI: 10.1007/s00382-010-0928-6.
37. McDonough F. Wind turbine icing and power forecast algorithm assessments in Scandinavia. In *Winterwind,* Sundsvall, Sweden, 2014. [Online]. Available: http://www.winterwind.se/sundsvall-2014/presentations-2014/?edmc=2802 (Accessed 4 April 14).
38. Bernstein BC, Hirvonen J, Gregow E, Wittmeyer I. Experiences from real-time LAPS-LOWICE runs over Sweden: 2011–2012 icing season. *Winterwind*, Skellefteå, Sweden, 2012. Available: http://www.slideshare.net/WinterwindConference/3a-bernstein-lapslowice (Accessed: 6 August 2013).