A probabilistic forecasting approach to wind turbine control

Matthew J. Holland* Kazushi Ikeda

Nara Institute of Science and Technology
Ikoma, Nara, Japan

Abstract

In wind-based energy operations, turbine performance, part deterioration rates and regular maintenance costs play a significant role in the economic viability of such projects. Problems to determine in advance the optimal future state of variables such as generator torque or blade pitch angle may be framed as anticipatory control problems, though successful application in real-world operations requires reliable forecasts of the future state of local wind speed. The authors propose a general probabilistic forecasting approach within the geographically robust CRPS minimization estimation framework suited to make use of AMeDAS weather variable observations, with the chief goal of evaluating the utility of the AMeDAS observations in the context of a rotor speed control problem where one seeks to maximize long-term power output at a specific turbine. Tested over a full year's worth of observations from sites across Japan, and using deterministic and non-deterministic evaluation metrics compared against standard references, the capacity of the proposed model as a forecaster at the 10-minute horizon was verified, and in doing so confirmed the potential wide-scale utility of the AMeDAS network in this and related anticipatory control problems.

1 Introduction

The chief source of risk in wind-based energy projects is the inherently intermittent nature of the resource. While well-understood physical systems such as mesoscale atmospheric flows can be observed and can in principle be utilized to make largely deterministic statements about the wind's future state, the complexity and computational cost involved naturally leaves significant uncertainty, thereby motivating stakeholders in wind energy to take interest in the wind's stochastic traits [11]. A large body of both academic and industry-based research illustrates that the economic viability of wind-based power projects, in particular those operated by non-governmental independent power producers (IPPs), largely falls upon three factors: efficient decision-making in liberalized power markets [15], optimization of turbine power output [8], and the minimization of operational costs incurred at wind farms [9]. Modern wind turbines allow operators numerous controllable parameters, and the latter two factors are directly impacted by effective, dynamic control of wind turbines.

Representative work in the literature includes a cascaded non-linear controller from Boukhezzar and Siguerdidjane [1] designed to maximize the kinetic energy of the wind captured by the turbine, while Kusiak and Zhang [9] proposed a framework for minimizing the vibration that shortens the useful life of wind turbines, namely by specifying a model for tower and drive train

*Email: matthew-h@is.naist.jp.

acceleration (measured by two installed accelerometers) in which one optimizes over controllable inputs such as blade pitch and generator torque at a given point in time. Other similar works focused on efficient control in the wind farm setting exist in the literature [12], but the crux of utilizing most such technologies is gaining access to reliable, local, short-term forecasts of future wind speed. "Short-term" here may refer to a temporal resolution of anywhere from a few seconds to upwards of 1–2 hours, though in this study the focus is on the 10-minute time scale.

One paradigm for approaching the wind forecasting problem comes from traditional physical models of atmospheric flows, generally referred to as numerical weather prediction (NWP). The massive volume of continuous data input and computational resources required in NWP methods have historically been cost-prohibitive [10], though as computer hardware improves, so does the viability of NWP, though even today a forecast horizon of below 6–8 hours will almost inevitably see a data-driven statistical approach being the most feasible solution to the problem [14]. Statistical methods traditionally have produced "point forecasts," namely a unary output intended to be interpreted as the best estimate of some future state. The argument has been eloquently made by Pinson [13] among others that while such techniques are simple and user-friendly, in addition to being able to produce numerical point forecasts, there is in fact great value in making "density forecasts" which specify a probability distribution for the future state of the wind, in that by doing so one offers users far richer information about the uncertainty involved with all estimates being made.

In this study the authors focus on the role of short-term probabilistic wind forecasts in turbine control with the goal of maximizing power output. A cost-effective probabilistic approach to a long-term power output maximization problem is proposed which utilizes a networked array of meteorological observation sites distributed across Japan called AMeDAS, and the performance of the proposed method is investigated with AMeDAS observational data from several site clusters from around the country. In section 2 a variable rotor speed turbine power output model is specified, and the basic control task is introduced. In section 3 the methods used for evaluating probabilistic forecast output are introduced, followed by an exposition of the proposed model which adapts to arbitrary forecast locations in section 4. Experimental details and highlights key observations are given in section 5, and 6 gives concluding remarks.

2 Rotor control and turbine output

Wind turbines are systems designed to capture the kinetic energy of wind and convert it into mechanical power. Traditional wind turbine design takes a propeller-like form, and while the detailed models of "blade element theory" are known to describe the system effectively, such models require an array of wind velocity signals, detailed information about blade geometry as well as complex computations, making them inherently cost-prohibitive [17]. A simplified model of instantaneous turbine power output at time t is generally given by

$$P(x_t, \lambda_t, \beta_t) = \frac{1}{2}\rho\pi r^2 x_t^3 C(\lambda_t, \beta_t) \tag{1}$$

where ρ is the air density (m/kg^3), $r > 0$ is the radius of the rotor (m), x_t is the wind velocity (m/s) assumed to be passing laterally through the rotor, and C_t is a performance power coefficient, and while it can be determined by the ratio of mechanical power extracted from the wind by a specific turbine P_t to actual wind power $\rho\pi r^2 x_t^3/2$, it is often modelled as a function of blade pitch β_t (rad) and tip-speed ratio λ_t. Tip-speed ratio is defined as $\lambda_t = \omega_t r / x_t$, where ω_t is rotor speed (rad/s), easily obtained as $2\pi N/60$ where N is rotations

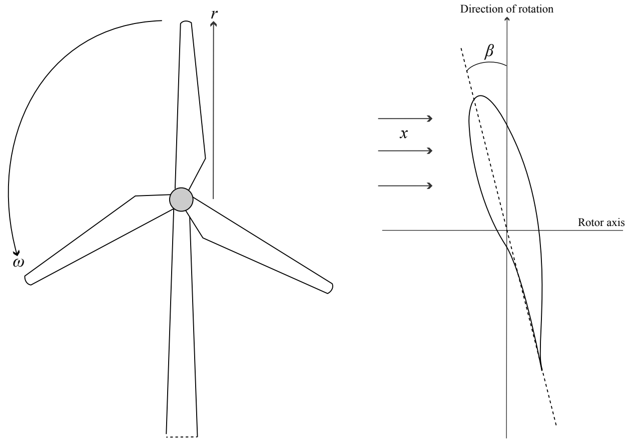

Figure 1: A simplified wind turbine schematic.

per minute. Following Slootweg et al. [17], the power coefficient is modelled as

$$C_t = \theta_0 \left(\theta_1 \lambda'_t - \theta_2 \beta_t - \theta_3 \beta_t^{\theta_4} - \theta_5\right) \exp\left(-\theta_6 \lambda'_t\right), \tag{2}$$

where

$$\lambda'_t = \frac{1}{\lambda_t - \theta_7 \beta_t} - \frac{\theta_8}{\beta_t^3 + 1} \tag{3}$$

and the $\theta_i \geq 0$ are pre-specified constants, for which Slootweg et al. [17] and Wasynczuk et al. [18] provide useful references. In evaluating and comparing the energy yield of specific turbine builds parameter specification is critical, though in our context of a general control problem, standard reference values will be appropriate. In this variable speed wind turbine output model, clearly ω_t and β_t are controllable parameters whose optimal values change depending on instantaneous wind speed, and it is over these parameters which one maximizes C_t and therefore P_t.

While x is easily measured at a high temporal resolution, instantaneous changes in β_t, ω_t are clearly not feasible, and for now letting $\beta = \beta_t$ be constant, and assuming ω_{t+1} must be specified at time t, which is to say

$$\omega^*_{t+1} = \arg\max_{\omega} \frac{1}{2} \rho \pi r^2 \hat{x}^3_{t+1} C_P(\hat{\lambda}_{t+1}(\omega)), \tag{4}$$

where $\hat{x}_{t+1}$ is a step-ahead wind speed forecast, and $\hat{\lambda}_{t+1}(\omega)$ is tip-speed ratio at $t+1$ given this forecast. When $\beta = 0$ the solution takes on a particularly simple form as

$$\omega^*_{t+1} = \frac{\hat{x}_{t+1}}{r\left(\frac{1}{\theta_4} + \frac{\theta_3}{\theta_1} + \theta_2\right)}. \tag{5}$$

Real turbines have manufacturer-specified cut-in and cut-out wind speeds which bound the domain over which power can be generated. The cut-in speed denotes the minimum speed at which in regular operation the turbine will generate power. Cut-out wind speeds are specified to prevent damage to turbines due to excessive vibration, among other risks. True output (as illustrated in Fig. 2 for 2012 specifications of major European turbines) must then be modelled piecewise as

$$P(t) = \begin{cases} P(x_t, \lambda_t, \beta) & \text{if } x_{in} < x_t < x_{out}, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

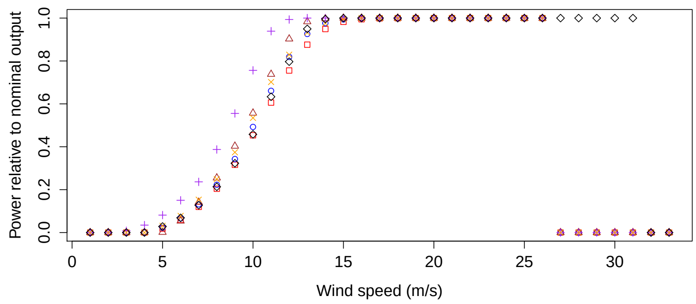

Figure 2: Power curves for six widely-used European-made wind turbines.

leaving us with the key task of making real-time local, short-term wind speed forecasts.

3 Evaluating wind forecasts

For an arbitrary location on the observation network at which one wishes to forecast, here called the "forecast site" and to which an index is assigned $s \in \mathcal{S} = \{0, 1, 2, ...\}$, the task of interest is one in which meteorological observations including wind speed up to present time step t are given, that is the sequence $(\mathbf{x}_{s,T}, \ldots, \mathbf{x}_{s,t})$, though first observation time $T \leq t$ may naturally depend on s. The forecasting task considered is that of specifying parametric distribution F for future wind speed $x_{t+k} \sim F(\boldsymbol{\theta})$ where $k > 0$ is forecast lag. This generally involves specifying some data-driven loss measure $\mathcal{L}(\boldsymbol{\theta})$ to be minimized either sequentially or in a batch-type approach. As this is a probabilistic model, setting $\mathcal{L}(\boldsymbol{\theta})$ to be the negative log likelihood assuming IID observations is a standard general approach, though Gneiting et al. [5] among other meteorological literature report that MLE methods lack robustness in terms of the variable to be forecast and other problem-specific meteorological conditions, and instead suggest the use of *continuous ranked probability score* (CRPS) as a loss measure to be minimized. Given predictive distribution F and observed value x it is defined

$$\text{CRPS}(F, x) = \int_{-\infty}^{\infty} (F(u; \boldsymbol{\theta}) - \mathbf{1}[u \geq x])^2 du, \tag{7}$$

and is both a proper scoring rule [4] and intuitively interpretable as evaluating both the "sharpness" of how F allocates density and how consistent (or "calibrated") it is with the data. $\text{CRPS}(F_{\boldsymbol{\theta}}, x)$ naturally depends entirely on the functional form of F, and a lack of closed-form expressions for CRPS given many distributions presents a challenge in many cases. While in general wind speed is positively skewed and thus non-Gaussian, readily computable expressions exist for the case $F(\boldsymbol{\theta}) = \mathcal{N}(\mu, \sigma^2)$, which is what is considered here. A Normal assumption implies support of $\mathbb{R}$ which is inappropriate for lower-bounded wind speed, and thus one is motivated to "truncate" the variable to $\mathbb{R}_0^+$. Letting wind speed be an observed random variable $X = x$ and letting B denote the event $x \in \mathbb{R}, x \geq 0$, basic probability results give us

$$P[X = x|B] = \mathbf{1}[B]\frac{P[X = x]}{P[B]} \tag{8}$$

and thus assuming the unconditional X is Normal, if define $Z = X|B$, can simply verify that $Z \sim \mathcal{N}^0(\mu, \sigma^2)$, defined by CDF

$$F(z) = \frac{\Phi\left(\frac{z-\mu}{\sigma}\right) - \Phi\left(\frac{-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{-\mu}{\sigma}\right)} \tag{9}$$

where Φ is the standard Normal distribution function. In this case, due to Gneiting et al. [3] there exists a closed-form expression of the CRPS given as

$$\begin{aligned} \mathrm{CRPS}\left(\mathcal{N}^0(\mu,\sigma^2), x\right) &= \mathrm{CRPS}\left(\mathcal{N}, x\right) - 2\sigma\phi\left(\frac{\mu}{\sigma}\right)\Phi\left(-\frac{\mu}{\sigma}\right) \\ &\quad + \frac{\sigma}{\sqrt{\pi}}\Phi\left(-\sqrt{2}\frac{\mu}{\sigma}\right) + \mu\left(\Phi\left(-\frac{\mu}{\sigma}\right)\right)^2, \end{aligned} \tag{10}$$

which is a readily computable error/loss measure, since a straightforward closed-form expression exists for the Normal case $\mathrm{CRPS}\left(\mathcal{N}(\mu,\sigma^2), x\right)$ as well. A simple starting point is to model the predictive at each time step as $\mu_t(\boldsymbol{\theta}_\mu) = \boldsymbol{\theta}_\mu^T \boldsymbol{\varphi}_{t-k}$, $\sigma_t(\boldsymbol{\theta}_\sigma) = \boldsymbol{\theta}_\sigma^T \boldsymbol{\psi}_{t-k}$, where in general $\boldsymbol{\varphi}_{t-k}$ and $\boldsymbol{\psi}_{t-k}$ are real vector-valued functions of the data observed as of time $t-k$. A specific case of this model was successfully implemented by the same authors, on a fixed network of three observation sites in the northwestern United States, one of which was designated the lone forecast site. In their study, *a priori* information and experiment design made it such that selection of *which sites* to use as data input sources was a non-issue. In general however, this is a non-trivial problem, and the authors put forward a systematic approach for automating this task in the next section.

4 Extension to arbitrary-site model

Given an arbitrary location on an observation network at which to forecast, unless the number of nodes on that network is trivially small, it will be inefficient to consider candidate input variables from all possible sites. Specific use is made of data from the AMeDAS observation network in Japan here, though it is desirable that the same approach be in principle applicable to any network of weather observation sites, and as such site-specific information requirements are kept minimal.

As an initial means of filtering the inputs to be considered, apply a threshold to the geographical distance from the forecast site. For site $s \in \mathcal{S}$ given latitude and longitude (a_s, b_s) (rad), using the well-known "haversine" formula and assuming the earth is spherical with radius R, the distance between sites s and s', denoted $d(s,s')$ is accurately estimated by

$$d(s,s') = 2R\,\mathrm{atan}\,2\left(\sqrt{h(s,s')}, \sqrt{1-h(s,s')}\right), \tag{11}$$

where

$$h(s,s') = \sin^2\left(\frac{a_s - a_{s'}}{2}\right) + \cos(a_s)\cos(a_{s'})\sin^2\left(\frac{b_s - b_{s'}}{2}\right). \tag{12}$$

Given some fixed forecasting point $\tilde{s}$ then, the first restriction applied to the pool of candidate inputs can be denoted by

$$\mathcal{S}_{\tilde{s}} = \{s \in \mathcal{S} : d(s,\tilde{s}) \leq D_{\tilde{s}}\}, \tag{13}$$

for some threshold $D_{\tilde{s}} \geq 0$. If $D_{\tilde{s}} = 0$, naturally only inputs from the forecast site will be available, while the limit in the opposite direction will result in the whole network being included in the candidate input pool.

Given the filtered network index $\mathcal{S}_{\tilde{s}}$, next consider the specific inputs that will be evaluated by a model selection algorithm as possible candidates for the elements of $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$. Assuming that for every site $s \in \mathcal{S}_{\tilde{s}}$ one has identically $D+1$ scalar-valued variables being observed, at time t the sequence of observations $(\mathbf{x}_{s,t}, \mathbf{x}_{s,t-1}, \ldots)$ is given, with $\mathbf{x}_{s,t} = (x^0_{s,t}, \ldots, x^D_{s,t})$. Equivalently, for $i = 0, \ldots, D$ a sequence denoted $\mathcal{D}^i_{s,t} = (x^i_{s,t}, x^i_{s,t-1}, \ldots)$ will be available, and a systematic means of determining how far back in time to consider each possible input is desirable, which may be included in the model as-is or subsequently passed through some transformation first. To do this, if assuming $i = 0$ is the wind speed variable index for each site, set a threshold $\tilde{\gamma}$ on cross-correlation to specify the number of lagged inputs to be included as model candidates as

$$m^j_s = \max\{m \in \mathbb{N} : \gamma^j_{s,m}, \gamma^j_{s,m-1} \geq \tilde{\gamma}\} \bigcup \{0\}, \tag{14}$$

where sample cross correlation $\gamma(\cdot,\cdot)$ is denoted

$$\gamma^j_{s,m} = \gamma\left(\mathcal{D}^0_{s,t}, \mathcal{D}^j_{s,t-m}\right), \quad s \in \mathcal{S}_{\tilde{s}},\, j \neq 0,\, m \geq 0 \tag{15}$$

and the observation sequences $\mathcal{D}^i_{s,t-m}$ will have some fixed common window length. In other words, at a given time t making forecasts for $t+k$, from a given site on the filtered sub-network $s \in \mathcal{S}_{\tilde{s}}$, the most lagged observation of the jth variable that will be considered for inclusion in the model would be from time $t - m^j_s$ as defined here. In this study, wind speed and direction are considered as potential inputs, and the thresholds $D_{\tilde{s}}$ and $\tilde{\gamma}$ may be determined systematically at each site by an appropriate cross-validation routine given the historical data, or simply fixed to reasonable values based on empirical observations.

Our goal is to utilize a model which generalizes well to arbitrary sites on the network, and another site-specific concern is diurnal non-stationarity of the wind speed, that is variation in the wind speed distribution with period of roughly 24 hours. As a straightforward means of dealing with an additive diurnal component, consider the n-order harmonic curve denoted

$$f(n, a, b) = a\sin(n\pi H/24) + b\cos(n\pi H/24) \tag{16}$$

and modelling the additive diurnal component of wind speed as a linear combination of these harmonics. While the optimal model order may differ by site, following Hering and Genton [6] a second order model is assumed across all sites, taking the expected general form

$$x = b_0 + f(2, b_1, b_2) + f(4, b_3, b_4). \tag{17}$$

For wind speed data from several sites on the AMeDAS network over two seasonally distinct time periods, fitting the b_i using a standard OLS routine yields clear diurnal components, as well as site-specific differences and distinct seasonality moving from the colder and drier winter months to warmer, humid spring months (Figs. 3 and 4). The transformation applied to the raw wind speed observations then, is to use the residual times series given by

$$x^r_t = x_t - \hat{b}_0 - f(2, \hat{b}_1, \hat{b}_2) - f(4, \hat{b}_3, \hat{b}_4). \tag{18}$$

In using the CRPS as a basis for some sort of loss metric, fixing a window of the most recent $M > 0$ time steps and using the arithmetic mean CRPS over this window

$$\mathcal{L}(\boldsymbol{\theta}, t) = \frac{1}{M}\sum_{i=0}^{M-1} \mathrm{CRPS}(\mathcal{N}^0(\mu_{t-i}(\boldsymbol{\theta}_\mu), \sigma^2_{t-i}(\boldsymbol{\theta}_\sigma)), x_{t-i}) \tag{19}$$

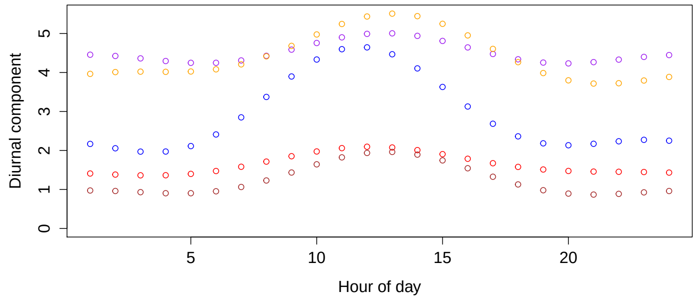

Figure 3: Diurnal component of 5 geographically unique forecast sites during first quarter (winter) of 2011.

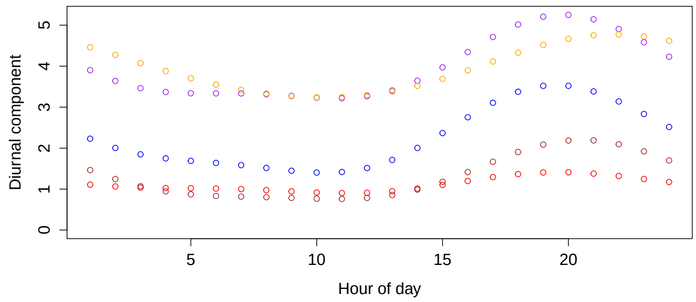

Figure 4: Diurnal component of same sites during second quarter (spring) of 2011.

is the recommended approach in the meteorological literature. Note that while at each time step this model assumes the weights $\boldsymbol{\theta} = (\boldsymbol{\theta}_\mu, \boldsymbol{\theta}_\sigma)$ are constant over these M time steps, the predictive parameters are functions of the data at each time step and thus clearly no assumption of temporal stationarity is made. At each time step then, parameters are estimated as

$$\boldsymbol{\theta}^* = (\boldsymbol{\theta}_\mu^*, \boldsymbol{\theta}_\sigma^*) = \arg\min_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}, t) \tag{20}$$

noting that in general constraints of $\theta_{\sigma,0}, \theta_{\sigma,1} \geq 0$ will be needed in the case of modelling $\sigma_t(\boldsymbol{\theta})$ as a weighted recent volatility term at all sites (typically weighted variance summed across sites) with data included in the final model. The gradient of this loss function may readily be accessed, and thus the efficient BFGS quasi-Newton algorithm implemented in the R language and environment [16] can be effectively utilized with logarithmic barrier functions used to implement the constraints. Model selection details are out of scope here, though simple approaches include the Bayesian Information Criterion (BIC), while somewhat more sophisticated approaches using information criteria based on the Kullback-Leibler divergence between the forecast model and the true underlying distribution may also be effectively implemented. In the experiments discussed in the following section a BIC-based algorithm was used in which candidate features are considered at each site in $\mathcal{S}_{\tilde{s}}$, for k lag and beyond consecutively until no improvement in the BIC value is found, at which point inputs from the next site in $\mathcal{S}_{\tilde{s}}$ is considered, and so forth.

Experiment parameters (fixed turbine)	
Turbine manufacturer	Vestas
Turbine model	112
Turbine radius	56m
Turbine (x_{in}, x_{out})	(3m/s, 27m/s)
Air density	1.25m/kg^3
Power param. $\boldsymbol{\theta}$	$(0.73, 151, 0, 0, 0, 0.003, 13.2, 0, 18.4)$
Distance param. $D_{\tilde{s}}$	50,000m

5 Experiment and results

In Japan there exists a nationwide network of remotely operated meteorological observation sites overseen by the Japan Meteorological Agency (JMA), with over 1,300 land-based sites capable of measuring and recording numerous quantities at varying temporal resolutions, called *Automated Meteorological Data Acquisition System* (AMeDAS). The sites are distributed roughly uniformly across the country, and a large subset of the network consisting of approximately 840 sites (one per 21km^2 on average) is equipped to measure wind speed and direction. The study conducted here uses three sites as forecast points, and data from surrounding sites spread across prefectures including Fukuoka, Toyama, Nagano, and Hokkaido and thus reflects sufficient diversity in terms of topographical characteristics and the stochastic profile of the wind, from extremely windy locations subject to large gusts to locations with calmer, albeit noisier wind.

A power generation simulation for a 10-minute forecast horizon at each site using AMeDAS data was conducted from January 1, 2011 through to December 31, 2011. With an assumption of constant pitch this represents 52,560 disjoint 10-minute intervals for which average power output is taken to be $P(\hat{x}_{t+1}, \hat{\lambda}_{t+1})$ and compare cumulative annual output for various standard reference forecasts against the optimal control specified when using a perfect forecaster. Dealing with very short horizons here, the "persistence" forecast modelled $\hat{x}_{t+k} = x_t$ is generally the strongest standard reference.

In the rotor speed control task, the best-performing standard reference (denoted PER) is compared against the forecaster discussed in the previous section, denoted AT, with a autocorrelation-weighted term in line with the first reference. Noting that while all forecasters take 10-minute data as raw input, the basic output of the references is 10-minute horizon forecasts while the proposed approach is hourly. The shorter time scale tends to be noisy, and thus from the proposed approach its hourly output is used over each disjoint set of six contiguous 10-minute intervals corresponding to that hour. While sensitivity is lost to very recent autoregressive trends, robustness to low-speed noise is gained. Fig. 5 shows a sample of the average wind speeds and power output from the single model turbine considered here, the latter normalized by the maximum output recorded over the course of the year-long test period. This is only a small segment of the time period considered, and output can readily be observed to increase while stronger wind velocities persist. As well, the small values for much of this period are simply a result of far stronger winds (and thus outputs) being observed at time points beyond the 500 hour mark.

Using the estimate of average power output for each of the 52,560 distinct time steps in the test data, Table 1 gives the ratio of total cumulative power generated in the case of the proposed and reference models to the case of using an "oracle," or a perfect forecaster. The results show the proposed approach is clearly competitive and results in what amounts to a substantial

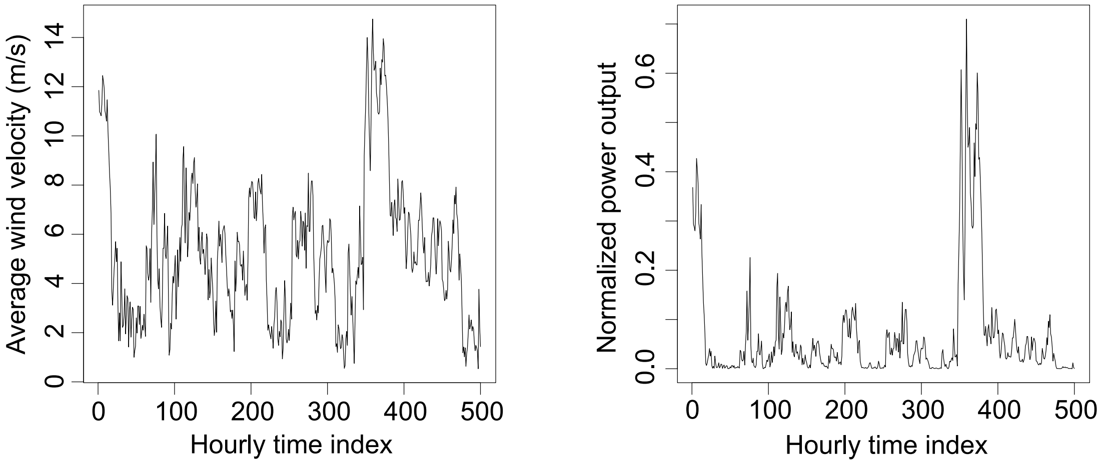

Figure 5: Hourly (averaged) wind velocity and 10-minute average power output (as ratio of max annual output) at Fukuoka forecast site for first 500 hours.

Site No.	1	2	3
PER	0.852	0.920	0.851
AT	**0.860**	**0.921**	**0.859**

Table 1: Cumulative annual power output between 0:09 on January 1st 2011 and 23:59 on December 31 2011, as a fraction of the total final output generated by perfect forecaster.

difference in total power output, on the order of up to 1% of total annual output in the best possible forecast scenario, which has clear economic significance for all but the smallest-scale turbines. A related visualization of these results for site 3 (centred near Hakuba, Nagano) can be seen in Fig. 6, and clearly reflects how particular seasons are more conducive to high-volume wind power generation. That is, month to month, the contribution to total annual output on average changes significantly. A large percentage of the annual output was clearly generated during the second quarter of the period considered, while subsequent periods saw a marked slowdown in output.

Explicitly considering the probabilistic output of the model proposed here at this same site, note that as is shown in Fig. 6, for a random sample of the outputs, the probability integral transform (PIT) histogram generated tends to "lack confidence" [2] in that it tends to assign very large or very small probabilities only rarely, which may be an indication of non-optimal consistency with the data. Considering the highly site-general nature of the approach proposed here, one can reasonably infer that this consistency (or "calibration") would improve should additional site-specific information be explicitly incorporated into the model. Detailed exploration of optimal probabilistic forecasts to make use of AMeDAS network data is out of the scope of this initial study. In any case, these results in the case of a very simple model suggest that in addition to providing value-added point forecasts, the additional information gained regarding the stochastic character and uncertainty surrounding the future state of the wind at a given site of interest (here on the AMeDAS grid) can be considered reasonably

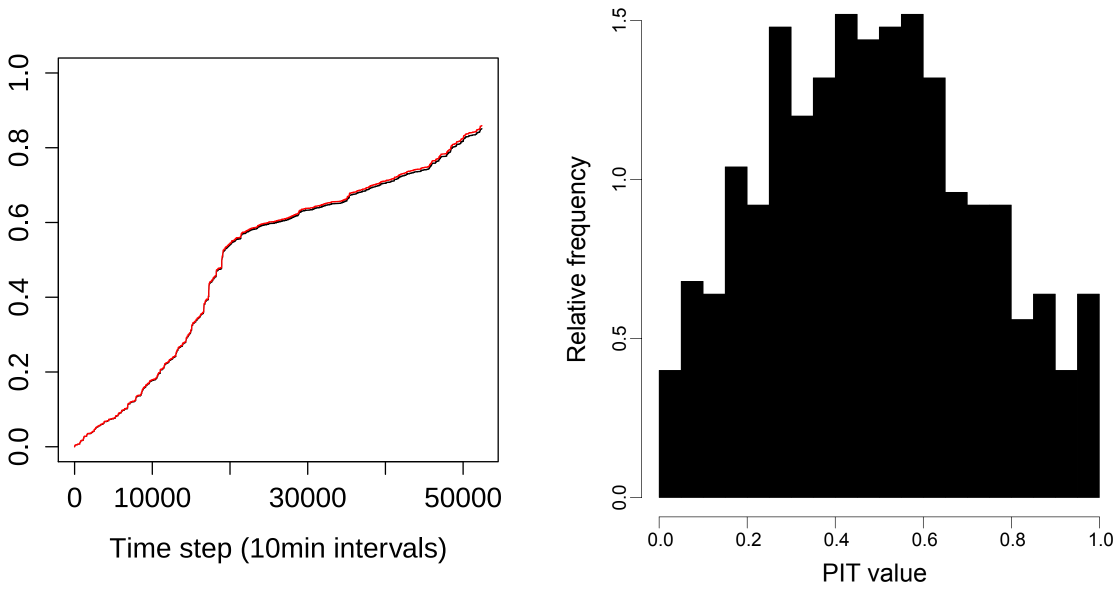

Figure 6: Left: Cumulative power output against optimal forecasts with AT in red and PER in black. Right: PIT histogram output for probabilistic output of TDD-based AT model at Hakuba.

reliable, suggesting that in-depth extensions of the proposed approach in a site-specific context would have a substantive impact on power output, and wind-based power forecast reliability.

In our analysis 30-minute, 1-hour, and 2-hour forecast horizons were also examined, and as would be expected seeing results at the 10-minute horizon, the TDD-based AT approach outperformed references by an increasingly large margin. Since the shortest time scale is most relevant to the control problem of interest here and the results are redundant, our discussion concludes at the 10-minute time scale. The inherent short-term noise present in most wind speed systems make it such that time scales on the order of one minute or less present unique challenges. Representative successful research for very short-term forecasts remains extremely sparse, though a general approach for 10 second forecast horizons using a standard Bayesian treatment has been proposed by Jiang et al. [7].

6 Concluding remarks

Our observations suggested that within the context of an anticipatory rotor speed control problem in which power output from a single turbine is optimized given blade pitch and a forecast of future wind speed, a CRPS-minimizing probabilistic forecasting approach, even with the limiting assumption of Gaussianity, can lead to superior long-term power output when compared with more naive approaches. Moreover, since the observations were for multiple locations spanning diverse segments of the AMeDAS weather observation network, these results can be interpreted as a verification of the utility of this data resource. As a general take-away, the operational decision to make use of modern density forecasting techniques in the vein of improving wind farm output at the single-turbine level would be justified not just for optimizing rotor speed and directly dependent controllable variables, but also within the context of more general power output forecasting tasks, and naturally optimal power supply planning.

Having seen the potential for probabilistic forecast-based control utilizing AMeDAS observations, there are a number of natural directions for potential new lines of work. To begin with, a site-adaptive model which takes into account spatial information about arbitrary sites of interest in a more meaningful way than the simplified approach taken here can be thought to be of value. Density forecasts within the CRPS framework are natural, though extensions beyond the Gaussian case also are natural improvements worth making in the pursuit of additional generality of the technology. A shorter time horizon, on the order of 0.1–1Hz will assuredly be of interest to the practitioner, and with the exception of the Bayesian approach taken in the work cited in the previous section, representative probabilistic forecasters focused on "short term" forecasting still tend to go no finer than 10 minutes [14]. At very short time intervals, small amplitude noise need not be forecast, and as such high-reliability forecasts are only important when wind velocity is extremely volatile. A natural next line of work then, may have the analyst change the formulation of the problem being considered to that of predicting the (probabilistic) risk of extreme weather, in this case wind gusts. For non-volatile sites, this will be a non-issue, though for sites where such forecasts are of value, an approach focusing on forecasting extreme values should be approachable even at very short time scales, so long as the data is available.

References

[1] Boukhezzar, B. and Siguerdidjane, H. (2009). Nonlinear control with wind estimation of a DFIG variable speed wind turbine for power capture optimization. *Energy Conversion and Management*, 50(4):885–892.

[2] Dawid, A. P. (1984). Statistical theory: the prequential approach. *Journal of the Royal Statistical Society: Series A (General)*, pages 278–292.

[3] Gneiting, T., Larson, K., Westrick, K., Genton, M. G., and Aldrich, E. (2006). Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space-time method. *Journal of the American Statistical Association*, 101(475):968–979.

[4] Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.

[5] Gneiting, T., Raftery, A. E., Westveld III, A. H., and Goldman, T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review*, 133(5):1098–1118.

[6] Hering, A. S. and Genton, M. G. (2010). Powering up with space-time wind forecasting. *Journal of the American Statistical Association*, 105(489):92–104.

[7] Jiang, Y., Song, Z., and Kusiak, A. (2013). Very short-term wind speed forecasting with Bayesian structural break model. *Renewable Energy*, 50:637–647.

[8] Kusiak, A., Li, W., and Song, Z. (2010). Dynamic control of wind turbines. *Renewable Energy*, 35(2):456–463.

[9] Kusiak, A. and Zhang, Z. (2012). Control of wind turbine power and vibration with a data-driven approach. *Renewable Energy*, 43:73–82.

[10] Lynch, P. (2008). The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, 227(7):3431–3444.

[11] Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., and Conzelmann, G. (2009). Wind power forecasting: state-of-the-art 2009. Technical report, Argonne National Laboratory (ANL).

[12] Moyano, C. F. and Peças Lopes, J. A. (2009). An optimization approach for wind turbine commitment and dispatch in a wind park. *Electric Power Systems Research*, 79(1):71–79.

[13] Pinson, P. (2006). *Estimation of the uncertainty in wind power forecasting.* PhD thesis, École des Mines de Paris.

[14] Pinson, P. (2012). Very-short-term probabilistic forecasting of wind power with generalized logit–normal distributions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 61(4):555–576.

[15] Pinson, P., Chevallier, C., and Kariniotakis, G. N. (2007). Trading wind generation from short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems*, 22(3):1148–1156.

[16] R Core Team (2014). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

[17] Slootweg, J., De Haan, S., Polinder, H., and Kling, W. (2003). General model for representing variable speed wind turbines in power system dynamics simulations. *IEEE Transactions on Power Systems*, 18(1):144–151.

[18] Wasynczuk, O., Man, D., and Sullivan, J. (1981). Dynamic behavior of a class of wind turbine generators during randon wind fluctuations. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(6):2837–2845.