Prediction intervals for power generation from multiple wind farms

B. W. McDonald[1]

(Received 17 December 2013; revised 23 October 2014)

Abstract

New Zealand has several existing or proposed wind farms to generate electric power. Small variations in wind speed can cause large variations in power output, due to a nonlinear relationship between wind speed and power. These variations may be correlated across different wind farms. There is interest in finding prediction intervals to quantify the risk of extreme changes in total wind power generation. At the individual wind farm level, an ad hoc method is proposed for modelling the probability distribution for wind some minutes in the future. This is used to estimate the conditional cumulative distribution function for future power output at each farm, given the regression model. A discrete approximation is used for the power output random variable, which reduces the problem to a set of conditional probabilities that can

be calculated. A recursive algorithm is used to combine the discrete cumulative distribution functions to find the conditional distribution of total power across all the wind farms given the most recent winds, and hence to find a prediction interval.

Contents

1 Introduction

Wind farms generate about 3–4% of New Zealand's electricity, but this may grow to 15–20%. Wind power generation has the advantages of less environmental impact and less cost than hydroelectric stations, gas/coal-fired power stations or geothermal generation. However, wind power is highly variable and its output needs to be continuously monitored to avoid the risk of sudden large changes in voltage damaging electrical equipment. To prevent such risks, up to 100 MW of hydroelectric generation capacity is currently reserved to instantaneously offset changes in wind. More reserve capacity is likely to be needed in future. Transpower[1], the operator of New Zealand's

---

[1]Transpower, https://www.transpower.co.nz/

electricity grid, is interested in quantifying the need for extra capacity and, in particular, predicting occasions when it is economical to reserve extra hydro, gas or geothermal capacity. These alternative sources have different costs and need different amounts of lead-time to ramp up generation. For this reason, and for efficient operation of the competitive electricity market, power prediction needs to be made for different time scales: 10, 20, 30, 60, 120 and 300 minutes into the future.

In February 2011, Transpower presented the problem of predicting power output for multiple wind farms to the Mathematics and Statistics in Industry Study Group (MISG) held at RMIT University, Melbourne. Transpower was especially interested in winds at nearby farms being correlated, and hence an increased risk of extreme power fluctuations. Synthetic data based on weather forecasts was provided for wind speeds at 15 current or planned wind farm sites throughout New Zealand. Some analysis of this problem was given by Whiten, McDonald and Bedford [1].

Section 2 suggests an empirical distribution for modelling the wind speed given recent winds. Section 3 approximates the cumulative distribution function (CDF) of the power output from a wind farm by a discretized version with certain probabilities. These probabilities are estimated via a regression of the transformed wind on past observations. Section 4 uses a recursive algorithm to find the CDF of the total power added across all the wind farms, and hence to find a prediction interval for future power output.

Figure 1 shows histograms of the synthetic wind distribution for each of the 15 actual or planned wind farms, and Figure 2 shows the corresponding histograms of power generated. Here, for simplicity it is assumed each farm generates a maximum of 100 MW. The reason for the bimodal power distribution is that the wind-to-power relationship is highly nonlinear. The relationship is illustrated in Figure 3. For a single turbine (Figure 3(a)), no power is generated for winds $< 5\,\mathrm{m\,s^{-1}}$, but then the power increases with the cube of the wind speed until it reaches a maximum at $15\,\mathrm{m\,s^{-1}}$, at which point a braking mechanism prevents the turbine going faster. The
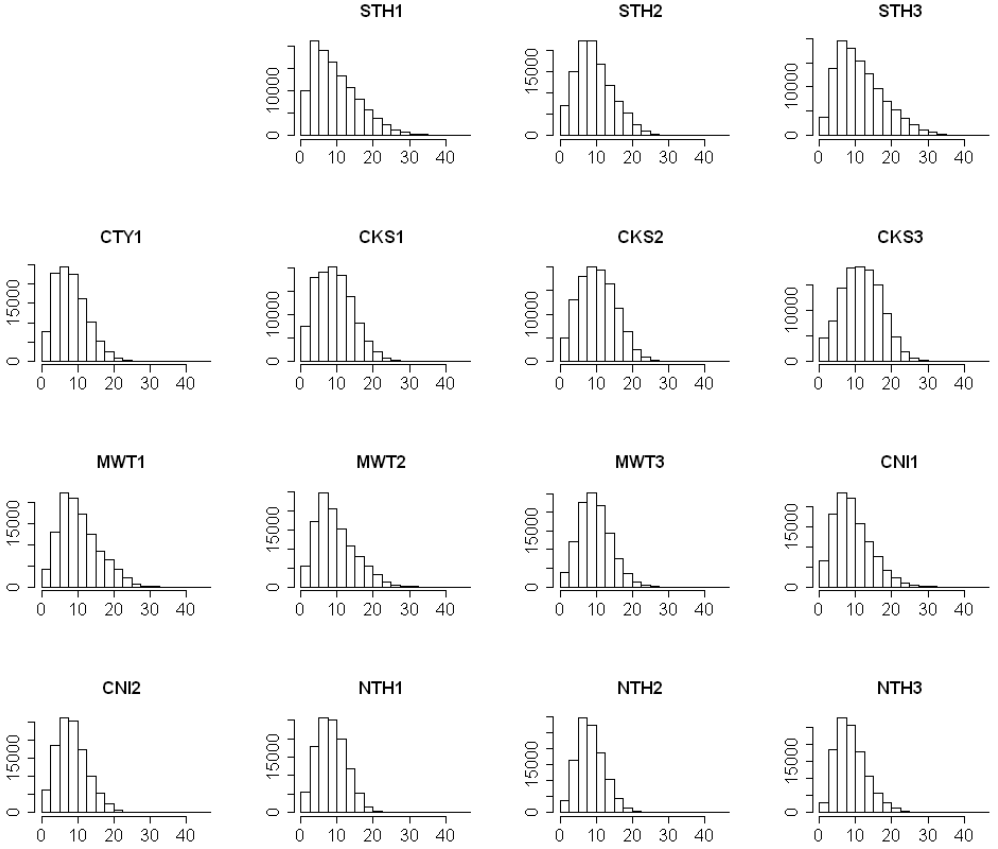
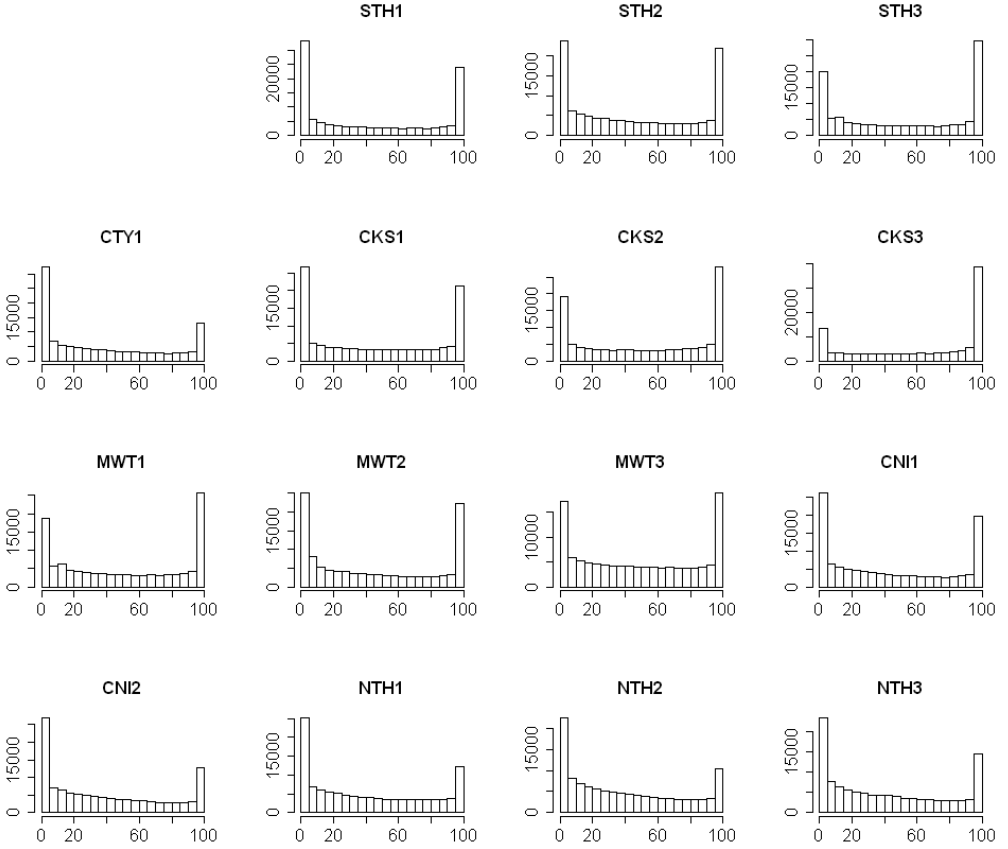Figure 1: Histograms of synthetic wind data for 15 wind farms ($\mathrm{m\,s^{-1}}$).

Figure 2: Histograms of power generation (MW), corresponding to Figure 1.

consequences of brake loss were demonstrated on video [3]. Once the wind reaches $25\,\mathrm{m\,s^{-1}}$ power generation cuts out altogether, again as a safety mechanism. Figure 3(b) shows the total power generation over several turbines in one farm, based on interpolation from a graph in an industry publication. Let X_{it} denote the wind speed for the ith farm at time t, and $Y_{it} = h(X_{it})$ the corresponding power, then the interpolated curve shown in Figure 3(b) is: $h(X) = 0$ for $X \leqslant 3.6$ or $X \geqslant 30.4$; $h(X) = 100$ for $15 \leqslant X \leqslant 23.1$; $h(X) = 33.679 - 20.264X + 3.4699X^2 - 0.12159X^3$ for $3.6 < X < 15$; and $h(X) = -6472.8 + 776.96X - 30.061X^2 + 0.37853X^3$ for $23.1 < X < 30.5$. This function was used to convert the winds in Figure 1 to the power output in Figure 2. The potential change in total wind power generation (across the 15 wind farm sites) in just 10 minutes is illustrated in Figure 4. Power rises or falls by more than $100\,\mathrm{MW}$ on 0.7% of occasions, or around 370 times over a year.

2   Wind distribution and transformation

Figure 5(a) shows 99% prediction bounds for future wind X_{it} at time t for the MWT2 site ($i = 9$), regressed on the observed wind x_{it-3} three ten minute lags earlier, using the standard formula

$$\hat{\beta}_0 + \hat{\beta}_1 x_{it-3} \pm t^* S \sqrt{1 + \frac{1}{t-3} + \frac{(x_{it-3} - \bar{x}_i)^2}{\sum_{j=1}^{t-3} (x_{ij} - \bar{x}_i)^2}}. \tag{1}$$

The $\hat{\beta}_0$ and $\hat{\beta}_1$ are estimated regression coefficients, x_{ij} refers to the wind history $j = 1, \ldots, t-3$, the sample mean wind is $\bar{x}_i$, and the sample standard deviation is S. If one ignores the wind autocorrelation, then the multiplier is $t^* = 2.58$ since the number of observations $t - 3$ is large. Figure 5(a) shows winds are underdispersed for initial winds $\leqslant 5$ $\mathrm{ms^{-1}}$ and overdispersed for initial winds between 10 and $25\,\mathrm{m\,s^{-1}}$. The upper (supposedly 99.5%) prediction bound is exceeded by 1.08% of points, while 0.61% of points are
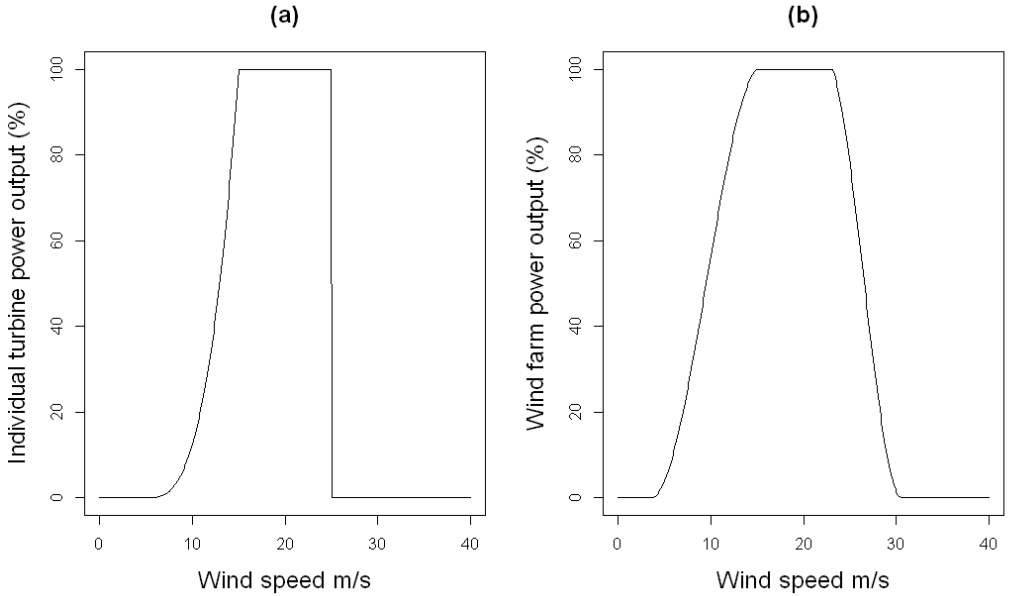
Figure 3: (a) Function relating wind to power for a single turbine. (b) Function relating wind to power across a wind farm. In both graphs power is expressed as a percentage of maximum power output.

below the lower (supposedly 0.5%) prediction bound. The probability plot Figure 5(b) shows the residuals are overdispersed relative to the normal distribution. To fix these problems, two ad hoc methods are proposed to rectify first the asymmetry and then the overdispersion.

Consider a power transformation of the wind speeds, X_{it}^p for farm site i and time t, where the exponent p is chosen empirically. Let $X_{i[q]}^p$ be the qth quantile of the distribution of all X_{it}^p values (aggregated over t). For symmetry, we want the sample 0.5% quantile $X_{i[0.005]}^p$ and 99.5% quantile $X_{i[0.995]}^p$ of transformed wind speeds to be equidistant from the median $X_{i[0.5]}^p$. That is,
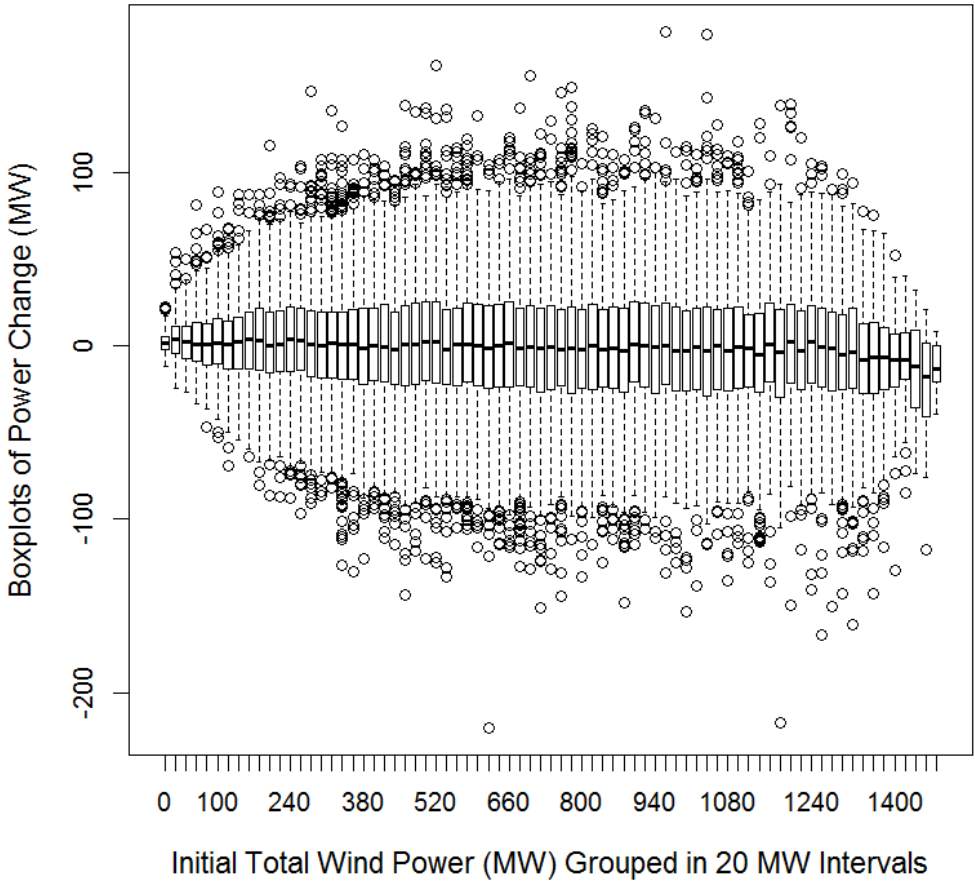
Figure 4: Boxplots of ten minute change in wind power generation versus initial power (grouped in 20 MW intervals).
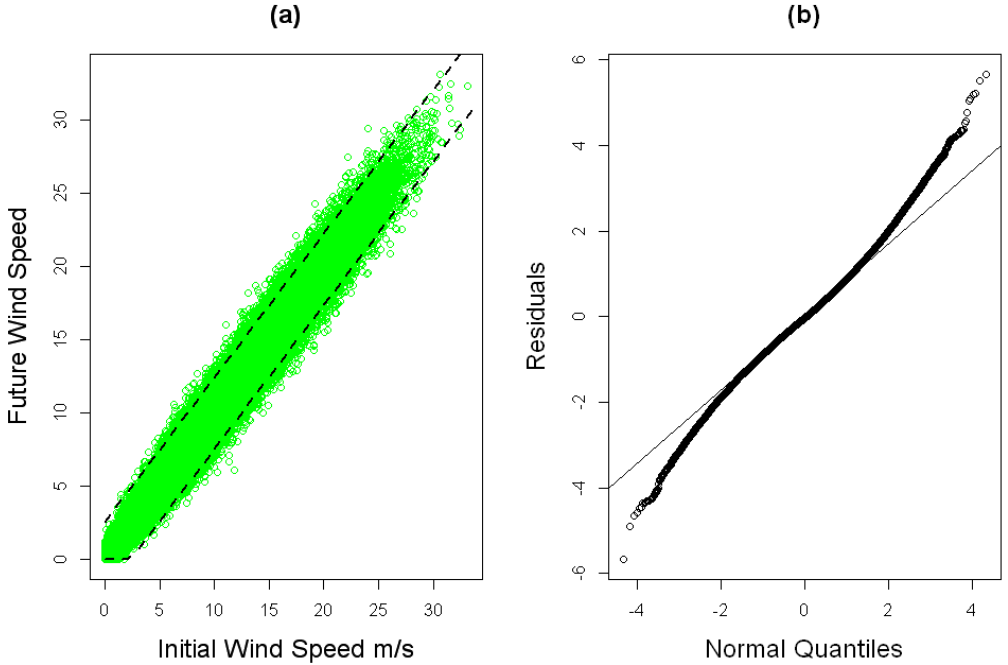
Figure 5: (a) The 99% prediction bounds (black lines) for a simple linear regression of the wind X_{it} at the MWT2 wind farm on the wind at the same farm 30 minutes earlier, X_{it-3}. (b) Normal probability plot of residuals from the regression.

$X_{i[0.995]}^p - X_{i[0.5]}^p = X_{i[0.5]}^p - X_{i[0.005]}^p$. Rearranging, we obtain the function

$$g(p) = \left(\frac{X_{i[0.995]}}{X_{i[0.5]}}\right)^p + \left(\frac{X_{i[0.005]}}{X_{i[0.5]}}\right)^p - 2 = 0, \tag{2}$$

which we solve for p using Newton's method $p = p_0 - g(p_0)/g'(p_0)$. A one-step solution with starting value $p_0 = 1$ gives good results for most of the wind farm data.

To correct the overdispersion we replace t^* in (1) with t_d, where the degrees of

freedom d is chosen so that the difference in quantiles $x^p_{i[0.995]} - x^p_{i[0.005]} = 2t_d S$. After solving for t_d an approximate d is found using

$$\frac{1}{d} = 0.581 - 1.647\frac{1}{t_d} + 0.388\left(\frac{1}{t_d}\right)^2, \tag{3}$$

which is obtained by regression. The error for d in using the approximation (3) is < 0.5 for $d < 200$. Figure 6 illustrates that these empirical methods reduce the heteroscedasticity and overdispersion. Here the estimates for farm MWT2 ($i = 9$) are $\hat{p}_i = 0.661$, $\hat{d}_i = 65$, and the percentage of points above and below the 99% prediction bounds are 0.55% and 0.43%, respectively.

The p_i and d_i need to be chosen separately for each wind farm, $i = 1, \ldots, 15$, and the transformed winds $X^{p_i}_{it}$ are then regressed on the lagged winds $X^{p_1}_{1t-l}$, $X^{p_2}_{2t-l}, \ldots, X^{p_{15}}_{15t-l}$ where l is the time lag. Whiten et al. [1] suggested also including time of day (via sinusoidal curves), indicator variables for month, and interactions between month and time, in the regression model as this will improve the fit, especially for predictions more than 30 minutes into the future, and also reduce the autocorrelation of errors.

3   Regression

After fitting the regression model, the residuals from the regression appear to be normal and have small autocorrelations. That is, the overdispersion and autocorrelation in X^p_{it} is largely removed by regression on past data. This implies that, conditional on suitable predictor variables (including past winds), the regression model along with normal errors fully describes the distribution of the random variable X^p_{it} for future observation times t. So, after back-transformation the model also gives the conditional distribution of the actual winds X_{it}, and from these the conditional distribution of the power output random variable $Y_{it} = h(X_{it})$, given the predictors.

The correlation r_{ij} between the regression residuals from different wind farms
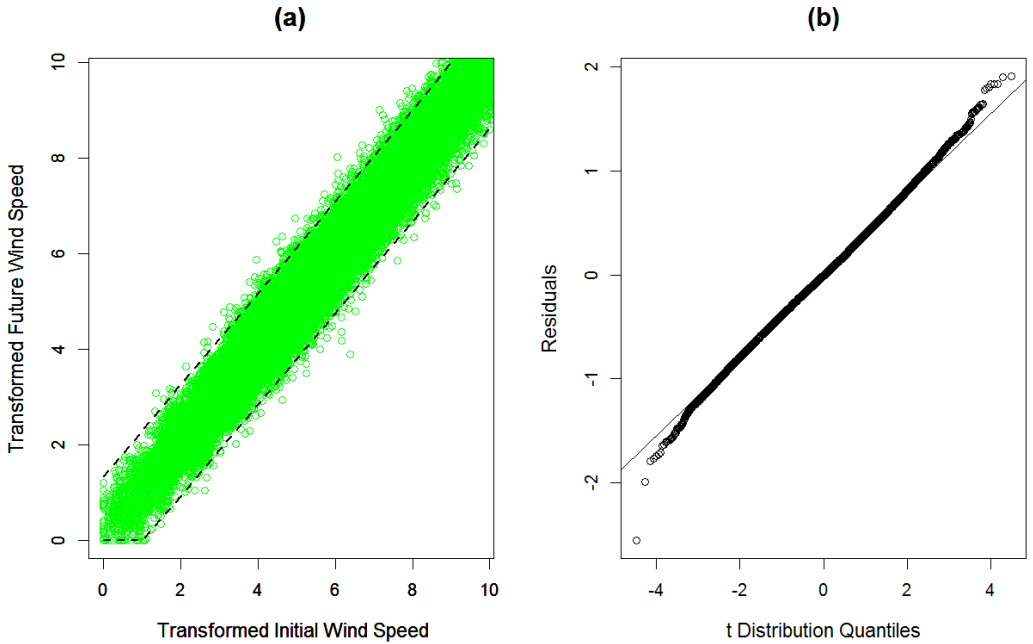
(a)   **(b)**

Figure 6:   (a) The 99% prediction bounds (black lines) for a simple linear regression of the X_{it}^p at MWT2 wind farm on the X_{it-3}^p at the same farm 30 minutes earlier.   (b) Quantile-quantile plot of the residuals against t_d, where p and d were chosen using the ad hoc methods.

i and j was also observed to be small, $r_{ij} \approx 0.03$ for all $i \neq j$. This allows the simplifying assumption that the Y_{it} and Y_{jt} are conditionally independent for $i \neq j$ (given the predictor variables). Thus the total power $S_t = \sum_{i=1}^{15} Y_{it}$ is approximately the sum of 15 (conditionally) independent bounded random variables. If one can estimate the conditional CDF of S_t, then one can find its 0.5% and 99.5% quantiles and this gives the 99% prediction interval. The difficulty is to calculate this conditional CDF in a computationally efficient way, since predictions for multiple time scales are needed every 10 minutes.

4   CDF for sum of bounded random variables

This section assumes the maximum power output for each wind farm is 100 MW. This is computationally convenient but the method can be generalized to unequal power maxima for each farm, and also to unequal wind-to-power functions $h(\cdot)$. Let $W_{it} = \mathrm{round}(Y_{it})$ be a discretized version of the power output at farm i at time t. That is, W_{it} takes values $w = 0, 1, 2, \ldots, 100$, with conditional probabilities $\pi_{it0}, \pi_{it1}, \ldots, \pi_{it100}$, say, given the predictor variables. Let ξ_{it-l} denote these predictor variables for $X_{it}^{p_i}$ based on data at time lag l. Now, consider the two sides of the wind-to-power conversion curve in Figure 3(b), say function $h_1(\cdot)$ for winds $\leqslant 20 \, \mathrm{m \, s}^{-1}$ and function $h_2(\cdot)$ for winds $> 20 \, \mathrm{m \, s}^{-1}$. Since we get a certain power output w from either the left or right side of the curve, the conditional probability

$$\pi_{itw} = \Pr(w - 0.5 < Y_{it} < w + 0.5 \mid \xi_{it-l}) = \pi_{1itw|\xi_{it-l}} + \pi_{2itw|\xi_{it-l}}, \quad (4)$$

say, where the subscripts 1 and 2 on π refer to the two sides of the curve. We expand these probabilities as

$$\pi_{itw} = \Pr\left(a_w < X_{itw} < a_{w+1} \mid \xi_{it-l}\right) + \Pr\left(b_w < X_{itw} < b_{w+1} \mid \xi_{it-l}\right), \quad (5)$$

where $a_w = h_1^{-1}(w-0.5)$ and $b_w = h_2^{-1}(w-0.5)$ for each $w = 0, 1, 2, \ldots, 100$. The a_w and b_w are known constants given by the wind-to-power curve h.

Now assume the regression of $X_{it}^{p_i}$ on ξ_{it-l} has normal errors and that p_i is known. Since n is large, a $100(1 - 2\pi)\%$ prediction interval for $X_{it}^{p_i}$ as in (1) is

$$\hat{X}_{it}^{p_i} \pm \Phi^{-1}(\pi) SR_{it-l}, \quad (6)$$

where $\hat{X}_{it}^{p_i}$ is the linear predictor, Φ^{-1} is the inverse of the normal CDF and SR_{it-l} is a term involving the lagged predictor variables. One does not need to calculate SR_{it-l} explicitly since it can be requested from a regression routine. Let C denote the lower value in (6). Then $\pi = \Phi\left[(C - \hat{X}_{it}^{p_i})/SR_{it-l}\right]$. Substituting $a_w^{p_i}$ for C, for $w = 0, 1, 2, \ldots, 100$, one obtains the first probability
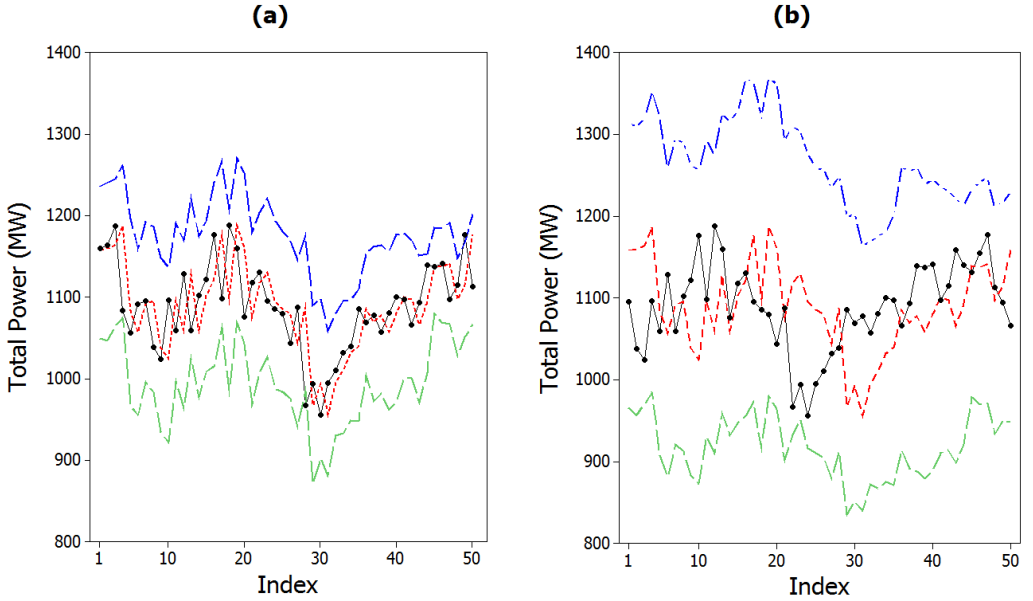
Figure 7: The 99% prediction intervals for total power, based on a simple linear model prediction of wind: (a) 10 minutes into the future; and (b) 60 minutes into the future. Black dots—realisation; red dashes—last value before prediction; blue and green dashed lines—99% prediction bounds.

in (5) as

$$\pi_{1itw|\xi_{it-l}} = \Phi\left(\frac{a_{w+1}^{p_i} - \hat{X}_{it}^{p_i}}{SR_{it-l}}\right) - \Phi\left(\frac{a_w^{p_i} - \hat{X}_{it}^{p_i}}{SR_{it-l}}\right), \qquad (7)$$

and similarly for the second probability π_{2itw} in (5). This is reasonably quick to compute since it only involves evaluating regressions and known functions, not simulation. The next step is to approximate the conditional CDF of S_t by the conditional CDF of the discrete sum $\tilde{S}_t = \sum_{i=1}^{15} W_{it}$.

The time subscript t is omitted in what follows. Consider the sequence of random variables $\tilde{S}^{(j)} = \sum_{i=1}^{j} W_i$ for $j = 1, 2, \ldots, 15$. Thus $\tilde{S}^{(j)}$ is the partial

sum of discretized power outputs over the first j wind farms. Denote the distribution function of $\tilde{S}^{(j)}$ as $F_s^{(j)} = \Pr(\tilde{S}^{(j)} \leqslant s)$. One finds $F_s^{(j)}$ recursively:

$$F_s^{(1)} = \Pr(W_1 \leqslant s) = \sum_{w=0}^{s} \pi_{1w} \quad \text{for} \quad s = 0, 1, 2, \ldots, 100\,,$$

$$F_s^{(2)} = \Pr(\tilde{S}^{(2)} \leqslant s) = \Pr(W_1 + W_2 \leqslant s)$$

$$= \sum_{w_2=0}^{s} \Pr(W_1 \leqslant s - w_2 \mid W_2 = w_2)\Pr(W_2 = w_2) = \sum_{w=0}^{s} F_{s-w}^{(1)}\pi_{2w}\,, \quad (8)$$

for $s = 0, 1, 2, \ldots, 200$. This involves inner products of vectors of probabilities padded with zeros and ones. Similarly,

$$F_s^{(j)} = \Pr(\tilde{S}^{(j)} \leqslant s) = \Pr(\tilde{S}^{(j-1)} + W_j \leqslant s) \qquad\qquad (9)$$

$$= \sum_{w_j=0}^{s} \Pr(\tilde{S}^{(j-1)} \leqslant s - w_j \mid W_j = w_j)\Pr(W_j = w_j) = \sum_{w=0}^{s} F_{s-w}^{(j-1)}\pi_{jw}\,,$$

for $s = 0, 1, 2, \ldots, (j \times 100)$. Thus the recursion involves convolutions which are quickly evaluated using fast Fourier transforms. Carrying the recursion to $j = 15$, one then identifies the s closest to the 0.5% and 99.5% quantiles of $\tilde{S}^{(15)}$ and these together give an approximate 99% prediction interval for future total power output. Conversely, since the whole approximate CDF is available by this method, one could use the method to estimate the probability that the change in power output will exceed $100\,\mathrm{MW}$, say, given the covariates.

As an illustration, Figure 7 shows 99% prediction intervals for simple linear regressions of wind on past wind at each individual site. The regression is for one day in May 2007, but based on all 2007 data. The red, green and blue lines refer to the power, and consequent 99% prediction bounds, at last observation ($t = 1, \ldots, 50$). The black dots refer to the realized 'future' value (they appear to precede the red line since the black values are later used as a predictor). In the 100 realizations, two are marginally outside the 99% prediction bounds, and visually the procedure appears to work well. A regression model with more predictors may have less autocorrelation in the errors, making the results robust over a range of conditions.

5   Conclusion and future work

This article presented a computationally feasible approach for calculating prediction intervals for total wind power across several wind farms. Further work needs to be done to validate these prediction bounds and to ensure the calculation is efficient. The regression needs to be sufficiently comprehensive (for example, including wind information from other farms) to ensure the residuals have minimal correlation. In practice the algorithm needs to allow unequal maximum output for the different wind farms and possibly different wind-to-power functions $h(\cdot)$, but this is a minor technicality. Industry can use these prediction bounds at various timescales to determine when additional hydroelectric generation capacity needs to be reserved to ensure safe, continuous supply in times of high variation. They can also be used to decide when other means of generation should be ramped up to overcome projected shortfalls in wind generation.

References

[1] B. Whiten, B. McDonald and R. Bedford. Modelling power output from correlated wind farms. *ANZIAM J. (MISG2011)*, 53:M104–M141, 2013. http://journal.austms.org.au/ojs/index.php/ANZIAMJ/article/view/5194/1759. C450, C457

[2] Electricity Authority. Wind integration analysis. *Technical Report*, 2010. ftp://ftp.emi.ea.govt.nz/Datasets/Supplementary_information/Archive/20101109_Wind_integration_analysis/

[3] The WindAction Group. *Wind turbine explodes.* February, 2008.
http://youtu.be/u14tBwO5QVQ C453

Author address

1. **B. W. McDonald**, Institute of Natural and Mathematical Sciences,
   Massey University, Auckland, New Zealand.
   mailto:b.mcdonald@massey.ac.nz