The Effect of Missing Data on Wind Resource Estimation

Aidan Coville
Department of Economics, University of Oxford, Oxford, UK
aidan.coville@qeh.ox.ac.uk

Afzal Siddiqui*
Department of Statistical Science, University College London, London, UK
afzal@stats.ucl.ac.uk

Klaus-Ole Vogstad
Energy Management and Trade, Agder Energi, Kristiansand, Norway
klaus-ole.vogstad@ae.no
BACKGROUND

★ Concerns with CO₂ emissions (IPCC (2007)) make renewable energy (RE) technologies attractive again
  ▶ Interest in biofuels, fuel cells, hydroelectric power, solar-based technologies, wave generation, and windmills
  ▶ Advantages: relatively low CO₂ emissions, less exposure to market or geopolitical risk
  ▶ Disadvantages: intermittency, integration, relatively high production cost (5.4 p/kWh for wind versus 2.5 p/kWh for coal)

★ Recently, adoption of RE technologies has been increasing exponentially albeit from a low base
  ▶ 92% of global energy needs are still met via dirty technologies (EIA (2007))
  ▶ Ongoing technological innovation and changes in policy provide scope for greater diffusion of RE sources
BACKGROUND

![Graph showing installed capacity (MW) from 1997 to 2007. The installed capacity increases significantly from 1997 to 2007.](image-url)
Revenue from a potential wind turbine is subject to risk due to:

- Intermittency from wind availability; and
- Volatility in the price of electricity

The former may be subject to measurement errors due to missing data at development sites

- Typically, there are higher wind speeds during winter months in Scandinavia, but anemometers are also more likely to freeze
- Take the perspective of a wind turbine developer in trying to impute the values of the missing wind speed data from an historical time series

Develop three *ad hoc* procedures for removing the bias in wind speed estimation due to missing data

- Find that the high volatility in the electricity price dominates in revenue density calculations
ASSUMPTIONS

★ No long-term patterns in electricity prices or wind speeds are considered
  ▶ Have only a one-year time horizon for evaluating wind resource and revenue distribution

★ Consider a now-or-never investment in only a single 850 kW turbine at the test site

★ Revenue is derived solely from electricity generated by the turbine and sold into the NordPool market
  ▶ Ignore government subsidies and CO₂ taxes

★ Investment and operating costs are not included in the calculation
Wind speeds typically follow a Weibull density:
\[ f_W(w) = \frac{\kappa}{\gamma} \left( \frac{w}{\gamma} \right)^{\kappa-1} e^{ \left( \frac{w}{\gamma} \right)^{\kappa}}, \quad w \geq 0 \]

There is also a dependency between wind speed and direction, which is dealt with by dividing the 360 degrees at the test site into twelve 30-degree sectors.

For each sector, separate Weibull densities are fitted using MLE.

Usually, investors are developing wind farms and want to know about the distribution of turbines across an area.

Use computational fluid dynamics to optimise the location, but we focus here on just a single turbine.
POWER CURVE

★ The theoretical power potential is $P(w) = \frac{\rho}{2} Aw^3$

★ Although the power output may be stochastic (see Anahua et al. (2007)), we assume a deterministic function with cut-in and cut-out points:

$$P(w) = \begin{cases} 
0.5w^3 - 21.4375 & \text{if } 3.5 \leq w < 12 \\
850 & \text{if } 12 \leq w < 25 \\
0 & \text{otherwise}
\end{cases}$$

★ The AEP sums up the product of the Weibull PDF of the wind speed for each sector and the power curve weighted by the frequency of the wind directions:

$$AEP = n \sum_{i=1}^{s} r_i \int f_{W_i}(w) P(w) dw$$
POWER CURVE

Vestas V52 Wind Turbine

Wind Speed (m/s)

Power Output (kW)

Cut-in

Rated Power

Cut-out

Simplified Estimate

Wind Speed (m/s)

Power Output (kW)
ELECTRICITY PRICE PROCESS

★ Typically, logarithms of electricity prices are modelled as being mean-reverting processes with possibly a long-term mean that follows a Brownian motion (Lucia and Schwartz (2002))

★ Also used are processes with mean reversion (Bierbrauer (2004)), stochastic volatility (Maribu et al. (2007)), and jump diffusion (Weron et al. (2004))

\[ dX_t = \eta(X - X_t)dt + \sigma dZ_t + J_t dq_t \]

- Work with deseasonalised data: \( P_t = S_t e^{X_t} \)
- Use Harvey (1989) to fit appropriate sine and cosine functions to capture the seasonality
DATA

★ Collect wind speed data for a two-year period (15 May 2006 – 14 May 2008) from wind masts installed by Agder Energi at Geitvassfjellet at a height of 48 m

★ About 1% of the data are missing, and we use the second year of data to do out-of-sample forecasts
  ▶ Other weather are available from NOAA and Norwegian meteorological stations, but with less resolution and reliability

★ Use electricity price data freely available from the Nord-Pool website from 1 January 1999 to 14 May 2008 (in SEK/MWh)
METHODOLOGY
IMPUTATION OF MISSING WIND SPEED DATA AND FITTING MRJD PROCESS

★ Three approaches for imputing missing wind speed data:

- Persistence: give the value of the most recent observation
- Regression: linear regression estimates wind speeds at Geitvassfjellet with 10% of observations removed, and MLE is used to find Weibull PDF parameters for AEP calculations via simulation
- Seasonality: remove the seasonality present in both the wind speed and direction before fitting the Weibull distributions to each directional sector

★ Approach for modelling the electricity spot price:

- Remove seasonality and identify spikes using Blanco and Soronow (2001) by flagging all spot price differences that exceed three SDs
- Repeat until all such spikes have been removed and fit a MR process to the residual data
- Model the spikes as occurring according to a Poisson process (independently of the MR process) and use lognormal distribution to model the magnitudes of the spikes
SEASONALITY METHODOLOGY

1. Fit wind direction using a Markov model with six transition matrices to account for seasonality
2. Deseasonalise wind speed using robust regression
3. Fit Weibull distributions to each of the twelve directional sectors
4. Simulate wind direction and wind speed on one-year time horizon
5. Convert to power and estimate AEP
6. Re-apply seasonality
7. Re-order wind speeds using sorting algorithm
EXAMPLE OF SORTING ALGORITHM

Wind speed  Directional sector

\[ f(x) = 1.78 \left( \frac{x}{0.59} \right)^{0.05} e^{(0.59)^{1.05}} \]

Generate random increment from:

Generate next optimal value
= 2 (wind speed) + 4 (random increment) = 6

Repeat the process to find an optimal replacement wind speed
for \( t_3 \), now restricting the sample to directional sector 7

Find value closest to 6 from the simulated data, restricting
the sample to wind speeds associated with a directional sector of 4
AEP DENSITY FUNCTION: EFFECT OF MISSING DATA
ELECTRICITY PRICE MODEL

\[ dX_t = 0.00048(5.3 - X_t)dt + 0.015dZ_t + J_t dq_t \]
REVENUE DENSITY: EFFECT OF MISSING DATA

![Graph showing the effect of missing data on revenue density.](image-url)
AEP DENSITY: PERSISTENCE AND REGRESSION METHODS

![Graph showing density distributions for different methods: Benchmark, Persistence, and Regression.](image-url)
AEP DENSITY: SEASONALITY METHOD

![Graph showing density distribution over MVWh/year with two curves representing full and 10% scenarios.](Image)
REVENUE DENSITY: NO SPIKES IN ELECTRICITY PRICE

![Graphs showing density distributions for Swedish Krona prices with full and 10% scenarios.](image-url)
CONCLUSIONS

★ Investors require accurate analysis of revenue and risk exposure
  ▶ Government agencies may also require such information in order to set regulation about subsidies or CO₂ taxes

★ Missing data from wind speed measurements can significantly bias downwards the AEP estimate

★ Develop seasonality method to restore the AEP density
  ▶ May not be relevant for the revenue density if selling generated electricity in a market with price spikes
  ▶ However, could be significant if selling electricity in a regulated or low-volatility market

★ Directions for future research
  ▶ Combine with optimal timing and technology choice
  ▶ Examine impact of location of turbines in a wind farm