Time series sampling rates for wind energy forecasting

Eckley, Idris A.
Department of Mathematics & Statistics
Lancaster University
Lancaster, LA1 4YF, UK
E-mail: i.eckley@lancaster.ac.uk

Nason, Guy P.
Department of Mathematics
University of Bristol
Bristol, BS8 1BX, UK
E-mail: g.p.nason@bristol.ac.uk

1. Introduction

The challenge of accurate wind power forecasting has received considerable attention in recent years as the prevalence of this variable energy source increases. As Giebel *et al.* (2003) describe, different forecast horizons are required for different purposes. For example long-term forecasts, in the region of 7–10 days, are used for maintenance scheduling. Conversely very short term forecasts, of the order of seconds ahead, can be used to aid turbine control. Several different methods exist for obtaining power forecasts. Some forecasting schemes model the power output directly, whilst others forecast meteorological variables such as wind speed and use a wind farm's power curve to convert wind speed forecasts into power forecasts. See Giebel *et al.* (2003) or Genton and Hering (2007) for an introduction to this field.

Our attention in this paper focuses on a question which arises from the second of these forecasting approaches. Specifically we consider the issue of appropriate time intervals for wind speed forecasting. Typically when forecasting wind speed, producers use either hourly or 10-minute average wind-speed observations. It is natural to enquire whether these are appropriate measurements for use within a time series analysis. It is this question which we consider in this paper.

Recall, from time series theory, that given a time series sampled at rate δt, the highest frequency that can be observed *undistorted* in the series spectrum is the Nyquist frequency $\pi/\delta t$, see Priestley (1983) for example. Consequently if, for example, the sampling rate, δt, is halved, i.e. the series is sampled twice as often, then one can observe *undistorted* frequencies in the series twice as high as before. Given a fixed sampling rate, if the highest frequencies in the series are higher than the Nyquist frequency then a phenomenon called *aliasing* occurs. Aliasing means that high frequency information reappears in the spectrum (or equivalently its estimate) at lower frequencies as the sampling rate is not sufficiently high to capture the higher frequencies. Hence, one cannot trust a spectral estimate of a series that is subject to aliasing. As one would expect, the spectral distortion can be drastic if there is an appreciable amount of spectral power in the series at frequencies higher than the Nyquist limit.

In practice aliasing can occur whenever a continuous-time process undergoes sampling. It can also occur when a discrete-time process undergoes sub-sampling. Many time series texts, such as Hannan (1960), Priestley (1983) and Chatfield (2003) describe the effect of aliasing but do not give any practical advice about detecting it, let alone locating it or removing its effects. Since the autocovariance and the spectrum of a series form a Fourier pair, series that are subject to aliasing often have their autocovariance, and related quantities, distorted also. Consequently estimates of time series model parameters and associated forecasts can also be distorted. In this paper we will use recent methodology proposed by Eckley and Nason (2011) to explore whether aggregating data to the standard 10-minute average wind speed level is appropriate for a proposed onshore wind farm. The approach proposed by Eckley and Nason (2011) is based on the locally stationary wavelet modelling framework proposed by Nason *et al.* (2000).

The paper is organised as follows: in Section 2 we briefly review the locally stationary time series modelling framework which underpins the test proposed by Eckley and Nason (2011). Section 3 describes key results presented by Eckley and Nason (2011) which permit the identification of aliasing (or white noise corruption) within a time series and describes the circumstances under which it can be applied. The paper concludes in Section 4 by describing the results of our investigation.

2. Locally stationary wavelet processes

It is well known that a second-order *stationary* process, $\{X_t\}_{t\in\mathbb{Z}}$, has spectral representation given by

$$X_t = \int_{-\pi}^{\pi} A(\omega) e^{i\omega t} d\xi(\omega),$$

where $d\xi(\omega)$ is an orthonormal increments process, see Priestley (1983). However it is arguable that the majority of observed time series, such as wind speed, do not meet the requirements of second-order stationarity. Given this inherent non-stationarity, our work adopts a framework which permits the second-order structure to *change over time*. Such localisation is often achieved by replacing the amplitude $A(\omega)$ by a time-varying version, $A_t(\omega)$, (see e.g. Priestley (1983) or Dahlhaus (1997)). However, as Eckley and Nason (2011) describe, the usual spectral quantities associated with estimating key quantities in these kinds of Fourier models are rendered unidentifiable under aliasing. Hence, we turn to a locally stationary *wavelet* model where not only are the key quantities identifiable but they are also able to be well-estimated.

The locally stationary wavelet (LSW) process model with spectrum $\{S_j(z)\}_{j=1}^{\infty}$ introduced by Nason *et al.* (2000) (NvSK00) can be briefly defined as follows:

$$(1) \quad X_t = \sum_{j=1}^{\infty} w_{j,k} \psi_{j,k}(t) \xi_{j,k},$$

for $t = 1, \ldots, T = 2^J$, for some $J > 0$. Here (i) $\{\xi_{j,k}\}$ are a collection of uncorrelated random variables with mean zero and variance one; (ii) the $\{\psi_{j,t}\}$ are a set of discrete nondecimated wavelets and (iii) the $\{w_{j,k}\}$ are a collection of amplitudes that are 'smooth' in a particular way as a function of k and $w_{j,k}^2 \approx S_j(k/T)$ (the evolutionary wavelet spectrum (EWS)). The speed of stochastic evolution of the second-order properties of the series X_t is controlled by the smoothness constraints on $S_j(z)$. The original NvSK00 work stipulated Lipschitz smoothness constraints, which was further extended by Van Bellegem and von Sachs (2008) to add the possibility of jumps and other inhomogeneities in the spectrum by extending S to functions of bounded variation, which we adopt here.

The EWS, $S_j(z)$, controls how much variance there is in the process at different scales or frequency bands at time z. Thus, for example, $S_1(z)$ comprises the contribution to variance from the highest frequency bands, and the frequency bands get lower as j increases. The original work of Nason *et al.* (2000) and subsequent works such as Van Bellegem and von Sachs (2008) proposed methods for estimating $\{S_j(z)\}$ using single realizations of X_t via a bias-corrected smoothed wavelet periodogram $< X_t, \psi_{j,k-t} >^2$ where $< \cdot, \cdot >$ denotes inner product. In the next section we summarise recent work by Eckley and Nason (2011) which considers the effect of aliasing on the EWS.

3. Identifying the presence of aliasing

Recent work by Eckley and Nason (2011) has investigated the behaviour of LSW processes which are over-sampled. The initial focus of their work was to consider the effect of aliasing induced by dyadic sampling on the raw EWS estimate. Below we briefly summarise their main findings.

To begin, let $\{X_t\}_{t\in\{1,2,\ldots\}}$ be a LSW process and let $Y_t = X_{2^r t}$. In other words, Y_t is a (dyadically) subsampled version of X_t. We will focus on the spectral behaviour of the observed Y_t. Following NvSK00 we define the non-decimated wavelet coefficients of the LSW process to be given by $d_{\ell,m} = \sum_t Y_t \psi_{\ell,m}(t)$ and the raw wavelet periodogram to be given by $I_{\ell,m} = d^2_{\ell,m}$ for $\ell = 1,\ldots,\infty$ and $m \in \mathbb{Z}$. The raw wavelet periodogram is a (biased) estimator of the EWS. In practice, for all but a finite number of m the value of $d_{\ell,m}$ will be zero due to the finite support of the wavelet.

Definition 1 *For $Y_t = X_{2^r t}$ we define $D^{(r)}_{\ell,m} := \mathbb{E}(d^2_{\ell,m})$. This is the expected raw wavelet periodogram of the subsampled series.*

With the above notation in place we now describe a key result which considers the effect of aliasing on the EWS:

Theorem 1 (Eckley and Nason, 2011) *Let X_t be a LSW process with EWS given by $\{S_j(z)\}_{j=1}^{\infty}$ which obeys the same smoothness conditions as in Van Bellegem and von Sachs (2008) for Daubechies compactly supported wavelets. The expectation of the raw periodogram of $X_{2^r t}$ for $r = 1,\ldots,2^{J-1}$ is given by*

$$(2) \quad D^{(r)}_{\ell,m} = \sum_{j=1}^{r} S_j(2^r m/T) + \sum_{j=r+1}^{\infty} A_{j-r,\ell} S_j(2^r m/T) + \mathcal{O}(T^{-1}),$$

where $\ell = 1,2,\ldots$ and $m \in \mathbb{Z}$ and $A_{j,\ell}$ is the bias-correction operator formed from the inner product of the autocorrelation wavelets of $\psi_{j,k}$ as defined in NvSK00. The result also holds true for Shannon wavelets if the EWS, $S_j(z)$, has continuous first derivative for each $j > 0$.

This result shows that when a realisation of a locally stationary wavelet process becomes aliased, power from the aliased scales contaminate all other scales. For example in formula (2), above, the first term $\sum_1^r S_j(2^r m/T)$, is independent of ℓ. As such this term appears in the periodogram at each remaining scale. Using this feature, it is possible to detect the presence of aliasing and/or white noise within a time series. However, as Eckley and Nason (2011) discuss, to achieve this we must make certain assumptions about the process, i.e. that it is bandlimited or satisfies a decay property. Typically such information will not be known *a priori*. All is not lost though, for a remarkable result can be derived for Shannon-based wavelet processes.

The Shannon case We recall from NvSK00 that the A matrix for Shannon wavelets is diagonal with entries $A_{j,j} = 2^j$ for $j = 1,2,\ldots$. As Eckley and Nason (2011) describe, we can simplify (2) and form the following system of linear equations:

$$
\begin{aligned}
v_m &= \sum_{j=1}^{\infty} S_j(z_m), \\
D_{1,m}^{(r)} &= \sum_{j=1}^{r} S_j(z_m) + A_{1,1} S_{r+1}(z_m), \\
D_{2,m}^{(r)} &= \sum_{j=1}^{r} S_j(z_m) + A_{2,2} S_{r+2}(z_m), \qquad (3) \\
\vdots \quad & \vdots \quad \vdots \\
D_{J,m}^{(r)} &= \sum_{j=1}^{r} S_j(z_m) + A_{J,J} S_{r+J}(z_m).
\end{aligned}
$$

The matrix representation of this system is

$$
(4) \quad \mathring{D}_m^{(r)} = \begin{pmatrix} v_m \\ D_{1,m}^{(r)} \\ \vdots \\ D_{J,m}^{(r)} \end{pmatrix} = \mathring{A}_J \mathring{S}_m = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 0 & \cdots & 0 \\ 1 & 0 & 4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 2^J \end{pmatrix} \begin{pmatrix} \sum_{j=1}^{r} S_j(z_m) \\ S_{r+1}(z_m) \\ S_{r+2}(z_m) \\ \vdots \\ S_{r+J}(z_m) \end{pmatrix},
$$

where $\mathring{A}_J$ is a $(J+1)$-dimensional square matrix. With the above notation in place, the following result can be established:

Theorem 2 (Eckley and Nason, 2011) *Let X_t be a Shannon LSW process, $Y_t = X_{2^r t}$ and let $\hat{\mathring{D}}_m^{(r)}$ be the wavelet periodogram of Y_t. Finally let $\hat{\mathring{S}}_m = \mathring{A}_J^{-1} \hat{\mathring{D}}_m^{(r)}$. Then $\hat{\mathring{S}}_m$ is an unbiased estimator of $\mathring{S}_m$.*

In essence, the above means that we can obtain an unbiased estimate of the (previously aliased) EWS. A number of other useful results follow from Theorem 2. In particular Eckley and Nason (2011) highlight that the Shannon wavelet formula means that we do not need to assume that $S_j(z) = 0$ for all $j > R$, for some R (see Eckley and Nason (2011) for a discussion of this subtle, but important point). Rather helpfully, the first entry in the 'dealiased vector', $\sum_{j=1}^{r} S_j(z_m)$, is the total amount of aliased and/or white noise content at the location $z_m = 2^r m/T$. In addition the remaining entries, $\{S_{r+1}(z_m), \ldots, S_{r+J}(z_m)\}$, are the uncontaminated estimates of power at the observed scales.

Eckley and Nason (2011) also show that when no aliasing is present, the top row of the dealiased vector, $\hat{\mathring{S}}_m$, is zero and the remaining elements are an unbiased estimate of the true EWS. If the process analysed is pure white noise, then the top row of the dealiased vector $\hat{\mathring{S}}_m$ is σ^2 at all time points m with all other vector entries zero (Proposition 4 of Eckley and Nason (2011)). Consequently if a series contains either white noise or aliased power, then using Theorem 2, the power migrates to the top row of the 'dealised vector' $\hat{\mathring{S}}_m$. In the next section we consider the application of the above results to wind speed data.

4. Application to wind speed time series

Recall from Section 1 that our attention in this paper focuses on whether the use of 10-minute average wind-speed data is appropriate for modelling and forecasting future wind speeds. The data which we consider arises from a proposed onshore wind farm in the Midwest of the United States during March 2011. Data were collected at a rate of approximately 1Hz and subsequently aggregated to provide 10-minute average wind speeds, as used by several commercial packages to forecast wind farm power output.

A plot of the 10-minute average wind speed data is displayed in Figure 1(a) below, together with a plot of the first differences (to remove trend) in Figure 1(b). Visually, it would appear that the differenced time series is non-stationary and a test of stationarity confirms this.

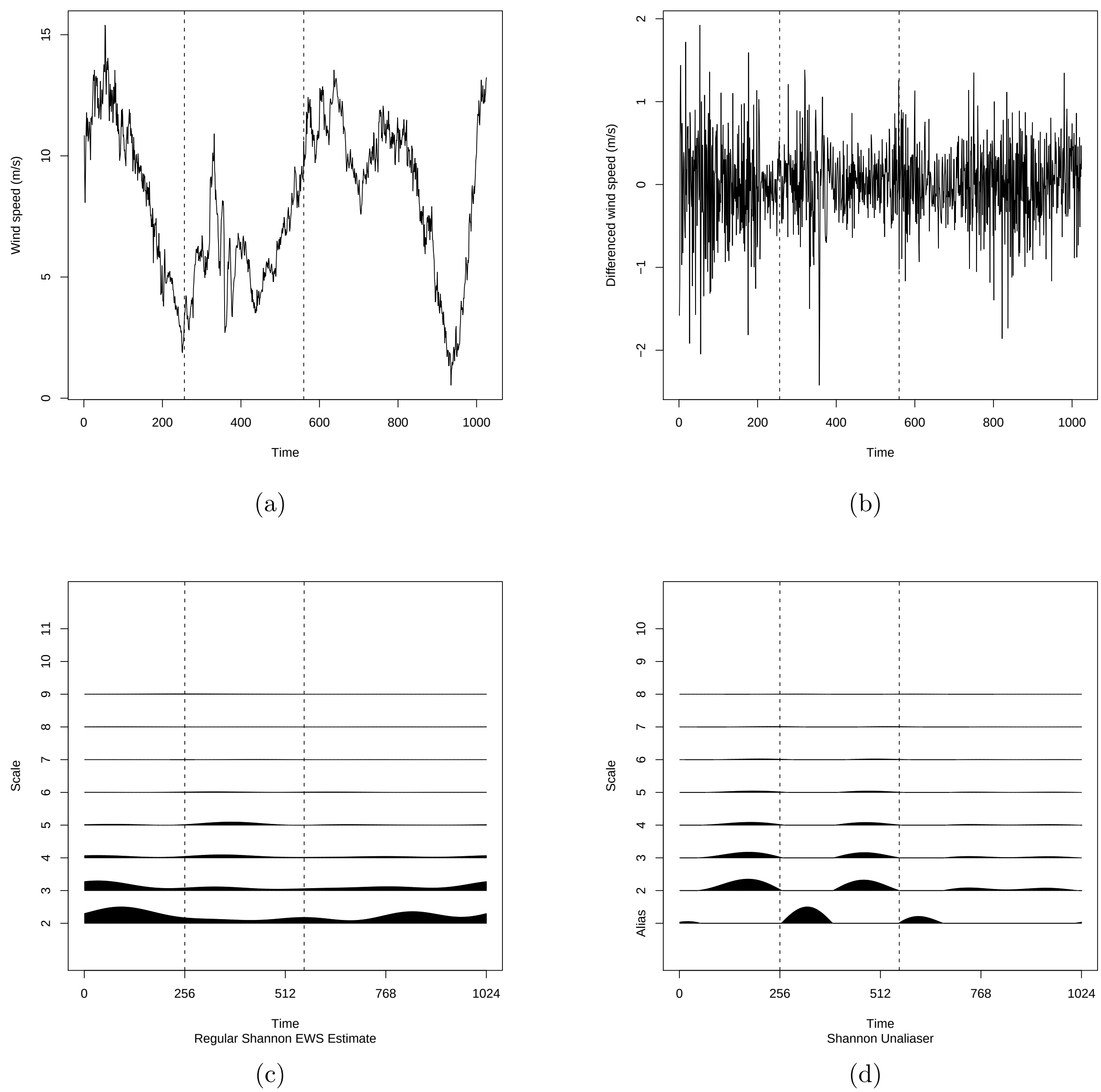

Figure 1: (a) 10-minute averaged wind speed recorded at a proposed wind farm in the US Midwest (ms^{-1}); (b) First differences of the 10-minute averaged wind speed (ms^{-1}); (c) Regular Shannon wavelet EWS estimate of the differenced data and (d) Dealiased Shannon wavelet EWS estimate of the differenced series. Vertical lines at $t = 256, 560$ are explained in the text.

To assess whether the observed wind speed data might be aliased, we apply the approach described by Eckley and Nason (2011). Figure 1(c) displays the standard Shannon wavelet EWS estimate of the differenced time series. Here the finest observed scale is labelled $j = 2$, as we are investigating whether power exists at finer (unobserved) scales. There is clearly oscillatory power at scales $j = 3$, 4 and 5 from time $t = 256$. At this time, power increases somewhat within scales 2–5 whilst power at the finest observed scale ($j = 2$) appears to tail off. Conversely, at around $t = 470$, the power in scale 2 increases whilst the power in the coarser scales diminishes. This pattern of power shifting between scales is reminiscent of a pulse changing in frequency over time. Something similar also appears to occur around $t = 560$.

Within Figure 1(d) the level labelled 'Alias' represents the sum of all potential aliases. It is particularly interesting to note that the visible content appears in two main bursts: the first occurs immediately after $t = 256$ whilst the second occurs around $t = 560$. Hence this analysis appears to confirm that, during the observed period power starts at lower frequencies, then increases through time to the extent that it becomes aliased and detectable by our method before returning to observable frequencies.

It is important to note that the results displayed in Figure 1(d) do not allow us to claim categorically that we are detecting aliasing within this time series. Recall that the methodology used cannot detect the difference between aliasing and white noise on its own. The power in the 'alias' level could therefore be due to a white noise component. However, given the above, we could be highly suspicious that aliasing is the underlying cause of the structure which appears on the 'Alias' level in Figure 1(d).

Acknowledgments

Eckley and Nason were supported by EPSRC Grants EP/I016368/1 and EP/I01697X/1 respectively. The authors also gratefully acknowledge the support of Chris Ziesler at Wind Capital Group who provided the data analysed in this paper.

REFERENCES (RÉFERENCES)

Chatfield, C. (2003) *The Analysis of Time Series: An Introduction*, Chapman and Hall/CRC, London.

Dahlhaus, R. (1997) Fitting time series models to nonstationary processes, *Annals of Statistics*, **25**, 1–37.

Eckley, I. A. and Nason, G. P. (2011) Alias detection and spectral correction for locally stationary time series, *in submission.*

Genton, M. and Hering, A. (2007), Blowing in the wind, *Significance*, **4**, 1114.

Giebel, G., Brownsword, R. and Kariniotakis, G. (2003) The state-of-the-art in short-term prediction of wind power, Project ANEMOS Report. http://anemos.cma.fr/.

Hannan, E. J. (1960) *Time series analysis*, Chapman and Hall, London.

Nason, G. P., von Sachs, R. and Kroisandt, G. (2000) Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum, *Journal of the Royal Statisical Society* B, **62**, 271–292.

Priestley, M. B. (1983) *Spectral Analysis and Time Series*, Academic Press, London.

Van Bellegem, S. and von Sachs, R. (2008) Locally adaptive estimation of evolutionary wavelet spectra, *Annals of Statistics*, **36**, 1879–1924.