

# Predicting sea waves in the presence of pink noise

Jacqueline Christmas

Computer Science Department, University of Exeter, Exeter, EX4 4QF, UK

Email: J.T.Christmas@exeter.ac.uk

**Abstract**—It has been shown that the power output of some wave energy converters can be greatly increased if they respond to very short-term predictions of the shapes of the waves. Observations of sea waves are traditionally made using buoys carrying GPS and accelerometers. The recent development of low-cost MEMS devices has led to cheaper devices, but also to a renewed interest in the effects of pink ( $1/f$ ) noise and signal processing methods for mitigating its effects. Bandpass filtering reduces the effects of this noise, but its remaining influence disrupts, in particular, the phase of the signal, which has significant consequences for prediction. We introduce a Bayesian model that promotes the smooth theoretical spectral shapes of the signal and the pink noise and estimates the true signal from one or more sets of observations recorded in parallel. The signal we are aiming to discover is the profile of sea waves at a fixed location; the spectral shape is determined by the Pierson-Moskowitz model. We demonstrate the model on synthetic data and give some preliminary results for the prediction of real sea waves.

## I. INTRODUCTION

It has been known for some time that a number of maritime operations can benefit from short-term knowledge of the sea surface shape. In particular, if wave energy converters of the point absorber type are able to alter their characteristics according to very short-term forecasts (a few tens of seconds) then their power output can be increased by a factor of the order of two [1]–[3]. Observations of wave profiles have been made using commercially available buoys that were originally designed to gather wave statistics, but we have noted that these can be prone to suffer from significant amplitudes of low-frequency noise, known variously as  $1/f$  noise (where  $f$  is the frequency), pink noise, or, in electrical systems, as flicker noise.

The simplest way to reduce the effects of this noise is to bandpass filter the signal, cutting off frequencies below and above the range of interest. However, where there is an overlap between  $1/f$  noise and signal, the remaining noise has a significant effect on the phase of the signal. In many applications this does not matter as only the amplitudes or the shape of the power spectrum are of interest. However, in our case we are interested in prediction and for this the correct estimation of the phase is vital [4].

In [5] we introduce a Bayesian spectral analysis model that assumes the noise to be white with occasional outliers. In this paper we extend this to explicitly model both white and  $1/f$  noise contributions, without outliers. The  $1/f$  noise is assumed to have a smooth spectral shape as described by (1) below. In the case of the measurements in which we are interested, the signal also has a smooth spectral shape,

allowing us to infer the amplitudes of the signal even where it is overpowered by the low frequency noise. On its own this does not enable us to extract the true phases from this overlap region. This is achieved by learning the signal from two (or more) parallel sets of observations where the noise is assumed to be independent. This shared signal, independent noise model is similar in concept to probabilistic Canonical Correlation Analysis (CCA) [6], though in this case we control the shape of both signal and  $1/f$  noise and the observations are ordered in time. The new model learns posterior distributions for the noise variables, as well as the amplitudes and phases of the signal's component sinusoids. The frequencies of the components are treated as parameters which we specify, as are the timestamps associated with each of the observations.

$1/f$  noise is a random process whose power spectrum is proportional to the inverse of the frequency,  $f$ , taking the form (for a one-dimensional signal)

$$S(f) \propto 1/f^\gamma \quad (1)$$

where  $0 < \gamma < 2$  and  $\gamma$  is usually close to 1 (a value of 0 gives white noise, while a value of 2 gives a random walk). This type of noise occurs widely in natural time series data found everywhere from climatology [7] to ecology [8], and in man-made systems ranging from music [9] to traffic [10]. It was first recognized in vacuum tubes by [11] and has since been recognised in all manner of electrical and electronic systems [12], including in graphene devices [13]. The recent development of low-cost MEMS devices for use in, for example, inertial navigation systems, has led to a renewed interest in the effects of this type of noise and signal processing methods for mitigating its effects on the true signals of interest (see, for example, [14]).

Figure 1 shows the power spectra from three different devices measuring the height of sea waves at a fixed location over time. The solid line is that from an expensive, low-noise, motion sensor and exhibits approximately the shape we expect of sea waves, while the black and grey dots represent the spectra of the altitude above mean sea level measured by Global Positioning System (GPS) devices. Notice that at low frequencies the GPS observations show high-amplitude noise, characterised by the dashed line. In this case the  $1/f$  noise is relatively benign; with some devices it swamps the signal. At the high frequency end of the spectrum both the signal and the  $1/f$  noise tend to be dominated by white noise, which has a flat spectrum.

Wavelet denoising [15] applies a basis function (a wavelet) to the signal at different scales and time offsets within the

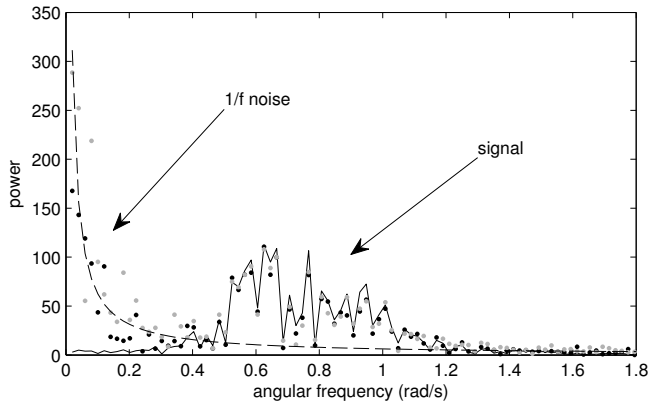


Figure 1 The power spectra from three different devices measuring the height of sea waves at a fixed location over time. The solid line is that from an expensive, low-noise, motion sensor and exhibits approximately the shape we expect of sea waves. The black and grey dots represent the spectra of the altitude above mean sea level measured by GPS devices. The dashed line characterises the shape of the estimated  $1/f$  noise in the GPS observations.

signal. The wavelet transform is a correlation analysis; the output is expected to be maximal where the noisy signal most resembles the wavelet. The Fourier transform uses a sinusoidal basis function which, conveniently for this application, represents the shape of a single frequency swell wave in the sea. In general, Wavelets with small coefficients are considered to be noise and removed, which can remove both high and low frequency noise, but, like bandpass filtering based on the Fourier transform, it can also remove components of the signal [14].

[16] use a fast orthogonal search algorithm [17] to model the low frequency range of the spectrum and conclude that it performs better than wavelet denoising. But they assume that the signal will have higher power than the noise at low frequencies and in our data this is not generally the case.

Allan variance [18], used by [19], provides a means of identifying and quantifying different noise contributions and depends on being able to record the noise where no signal is present. This method, along with autocorrelation and autoregression [20], is useful for determining which types of noise may be present, though for low frequency noise the number of autoregression coefficients required might be very large.

A Bayesian model allows us to incorporate any prior knowledge we may have regarding the likely distributions of the model variables and avoids the overfitting known to arise in maximum likelihood estimation by evaluating over all possible values of those variables. In this case the integrals required to perform exact Bayesian inference are intractable, so we use variational approximation to estimate the posterior distributions (for tutorials see [21,22] and [23, chapter 10]).

In section II we introduce the model and in section III demonstrate its utility on synthetic data. Some preliminary results with real sea data are shown in the conclusions section (IV).

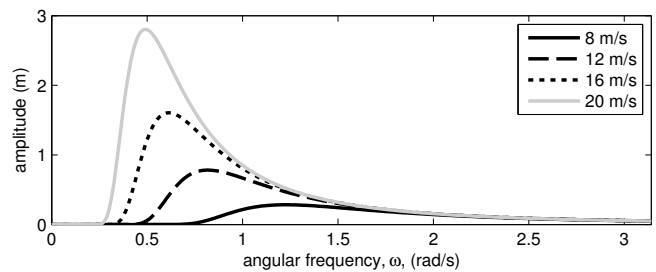


Figure 2 Examples of Pierson-Moskowitz amplitude spectra for a selection of wind speeds with 1,000 observations at 1 Hz.

## II. THE MODEL

The model assumes that each set of observations (for one or more sets) is made up of two parts: a signal that is common to all observation sets, and added noise that is unique to a particular set and independent between sets. For one set of observations consisting of  $N$ , one-dimensional measurements, we assume that each observation,  $y_n$ , is composed of signal and  $1/f$  noise contributions at  $C$  component frequencies with a white noise term  $\epsilon_n$ :

$$y_n = \sum_{c=1}^C \left[ A_c \cos(\phi_c - \omega_c t_n) + B_c \cos(\psi_c - \omega_c t_n) \right] + \epsilon_n \quad (2)$$

where  $\omega_c$  is the angular frequency of the  $c$ th component,  $A_c$  and  $\phi_c$  are the amplitude and phase of the signal at that frequency,  $B_c$  and  $\psi_c$  are the amplitude and phase of the  $1/f$  noise at that frequency, and  $t_n$  is the timestamp associated with the  $n$ th observation. In this paper the time intervals between observations are assumed to be equal, but this does not need to be the case [24].

We assume that the white noise is Gaussian distributed with zero mean and precision (inverse variance)  $\lambda$ , which leads to the following Gaussian likelihood:

$$\begin{aligned} p(y_n | \mathbf{A}, \phi, \mathbf{B}, \psi, \lambda; \omega, t_n) \\ = \mathcal{N}(y_n | \mathbf{A}^T \cos(\phi - \omega t_n) + \mathbf{B} \cos(\psi - \omega t_n), \lambda^{-1}) \end{aligned} \quad (3)$$

The  $\epsilon_n$  term not only accounts for the white noise, but when the model is trained on only a subset of the Fourier frequencies it also absorbs the effects of the missing frequencies [24].

### A. Smooth spectra

As previously mentioned, the theoretical spectra for the signal and the  $1/f$  are both smooth and of defined shapes.

1) *Signal*: The sea-shape of the signal spectrum is controlled by a single variable,  $U$ . Sea swell is caused by wind constantly blowing over the surface of the water for a period of several hours. During that time the waves build until an equilibrium is reached between the energy being input by the wind and that being dissipated by breaking and non-linear wave-wave interactions. A fully developed sea is one in which this equilibrium has been reached [25]. Based on observations of the North Atlantic, Pierson and Moskowitz

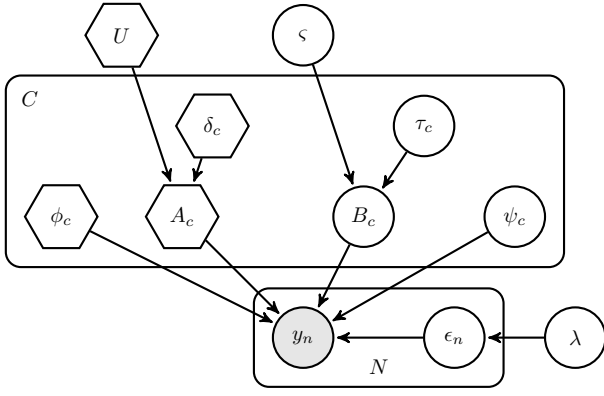


Figure 3 A graphical representation of the model variables and their dependencies. The variable  $y_n$  denotes the  $n$ th observation from one of the observation sets. All variables to the left of  $y_n$  (hexagons) are components of the signal that are shared across all observation sets. All other variables are components of the noise that is specific to one particular observation set.

[26] formulate a mathematical model for the power spectrum of a fully developed sea:

$$S(\omega_c) = \sqrt{\frac{0.0081g^2}{\omega_c^5} \exp\left(-\frac{5}{4} \left(\frac{\omega_0}{\omega_c}\right)^4\right)} \quad (4)$$

where  $g$  is the gravitational constant and  $\omega_0$  is the modal frequency of the wind, given as a function of the wind speed at an altitude of 19.5 m,  $U$ :

$$\omega_0 = 0.87g/U \quad (5)$$

For a set of  $N$  observations we may express (4) in terms of the amplitudes of the wave components:

$$A_c = 4\sqrt{S(\omega_c)/N} \quad (6)$$

Figure 2 shows examples of these smooth Pierson-Moskowitz spectra for a selection of wind speeds.

2) *Pink noise*: The shape of the  $1/f$  noise power spectrum (1) is controlled by two variables: a scale, corresponding to the proportionality, and a shape, corresponding to  $\gamma$  (1). In this paper we will assume that  $\gamma = 1$ , leaving us with just a scale variable,  $\varsigma$ . Let us redefine the shape of the spectrum in terms of amplitude  $B_c$ , rather than power, at a particular angular frequency  $\omega_c$  and for a set of  $N$  observations:

$$B_c = \varsigma \frac{2\pi}{\omega_c \sqrt{N}} \quad (7)$$

## B. Priors

A graphical representation of the model variables and their dependencies is shown in figure 3.

1) *Signal priors*: We specify a Gaussian prior for the signal amplitudes  $A_c$  that is dependent on the wind speed  $U$ :

$$p(A_c | \delta_c) = \mathcal{N}(A_c | f(\omega_c, U), \delta_c^{-1}) \quad (8)$$

where  $f(\omega_c, U)$  is the right-hand side of the Pierson-Moskowitz equation (6), and specify a Gamma distribution<sup>1</sup> for  $U$ :

$$p(U) = \mathcal{G}(U | a_U, b_U) \quad (9)$$

and a conjugate Gamma prior for the precision  $\delta_c$ :

$$p(\delta_c) = \mathcal{G}(\delta_c | a_\delta, b_\delta) \quad (10)$$

With no prior information about the component phases  $\phi_c$ , we model each of them with a Uniform distribution over the full range of values, i.e. any  $2\pi$  range; we choose  $-\pi$  to  $\pi$ :

$$p(\phi_c) = \mathcal{U}(\phi_c | -\pi, \pi) \quad (11)$$

Were prior information available we would instead specify a von Mises distribution for each phase. The Uniform distribution is a special case of the von Mises and we use it here to maintain the clarity of the implementation of the model by minimising the number of parameters.

2) *Noise priors*: The precision of the white noise Gaussian distribution for each observation set,  $\lambda$ , is assigned a conjugate Gamma prior:

$$p(\lambda) = \mathcal{G}(\lambda | a_\lambda, b_\lambda) \quad (12)$$

As for the  $\phi_c$ , the phases  $\psi_c$  are modelled with a Uniform distribution over the full range of values:

$$p(\psi_c) = \mathcal{U}(\psi_c | -\pi, \pi) \quad (13)$$

For the  $1/f$  noise scale variable  $\varsigma$  (7) we assume a Gaussian prior:

$$p(\varsigma | m_\varsigma, p_\varsigma) = \mathcal{N}(\varsigma | m_\varsigma, p_\varsigma^{-1}) \quad (14)$$

Finally, we specify a Gaussian prior for the amplitudes  $B_c$  that is dependent on the  $1/f$  noise variable (from the right-hand side of (7)):

$$p(b_c | \varsigma, \tau_c) = \mathcal{N}(b_c | \varsigma 2\pi / (\omega_c \sqrt{N}), \tau_c^{-1}) \quad (15)$$

and specify a conjugate Gamma distribution for its precision:

$$p(\tau_c) = \mathcal{G}(\tau_c | a_\tau, b_\tau) \quad (16)$$

3) *Summary*: Thus the model consists of two parameters, the  $C$ -dimensional  $\omega$  and the  $N$ -dimensional  $\mathbf{t}$ , and the variables  $\mathbf{A}$ ,  $\delta$ ,  $\phi$ ,  $\mathbf{B}$  and  $\psi$  (all  $C$ -dimensional), and  $U$ ,  $\gamma$  and  $\lambda$ . For the variables we aim to calculate their posterior distributions, which provide us with both estimates of their values and a measure of the uncertainty in those estimates.

## C. Posteriors

In an exact Bayesian model we would need to calculate the evidence by integrating the product of the likelihood and all the priors by all of the variables. As is often the case, this is intractable, so we resort to an approximation scheme. Rather than using a computationally expensive Monte Carlo method, we use the factorised version [27] of variational approximation [21]–[23]. To maintain the clarity of this paper, we assume a

<sup>1</sup>defined as  $\mathcal{G}(x | a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx)$

full factorisation of the approximate posteriors. With  $S$  observation sets and  $(s)$  denoting a variable associated with the  $s$ th set, we define, for convenience:

$$u_{c,n} = \cos(\phi_c - \omega_c t_n) \quad (17)$$

$$v_{c,n}^{(s)} = \cos(\psi_c^{(s)} - \omega_c t_n) \quad (18)$$

and the residuals

$$\tilde{y}_n^{(s)} = y_n^{(s)} - \sum_{c=1}^C B_c^{(s)} \cos(\psi_c^{(s)} - \omega_c t_n) \quad (19)$$

$$\hat{y}_n^{(s)} = y_n^{(s)} - \sum_{c=1}^C A_c \cos(\phi_c - \omega_c t_n) \quad (20)$$

which are the observations minus the estimated  $1/f$  noise and the observations minus the estimated signal respectively.

1) *Signal posteriors:* We now use the factorised variational approximation technique [27] for each variable in turn, starting with each  $A_c$ , the amplitude of the  $c$ th component sinusoid. With  $q(\cdot)$  denoting an approximate posterior and  $\langle \cdot \rangle$  the posterior expectation, we end up with  $q(A_c)$  as the Gaussian  $q(A_c) = \mathcal{N}(A_c | \mu_c, \sigma_c^2)$ , where

$$\sigma_c^2 = \left( \langle \delta_c \rangle + \sum_{s=1}^S \langle \lambda^{(s)} \rangle \sum_{n=1}^N \langle u_{c,n}^2 \rangle \right)^{-1} \quad (21)$$

$$\mu_c = \sigma_c^2 \left( \langle \delta_c \rangle \langle f(\omega_c, U) \rangle + \sum_{s=1}^S \langle \lambda^{(s)} \rangle \sum_{n=1}^N \langle u_{c,n} \rangle \left( \sum_{\substack{i=1 \\ i \neq c}}^C \tilde{y}_n^{(s)} - \langle A_i \rangle \langle u_{i,n} \rangle \right) \right) \quad (22)$$

For the precision of each amplitude,  $\delta_c$ , we obtain the Gamma posterior  $q(\delta_c) = \mathcal{G}(\delta_c | \alpha_{\delta_c}, \beta_{\delta_c})$ , where

$$\alpha_{\delta_c} = a_{\delta} + 1/2 \quad (23)$$

$$\beta_{\delta_c} = b_{\delta} + \frac{1}{2} \left( \langle A_c^2 \rangle \langle f(\omega_c, U)^2 \rangle - 2 \langle A_c \rangle \langle f(\omega_c, U) \rangle \right) \quad (24)$$

For the wind speed  $U$  we end up with

$$\log(q(U)) \propto (a_U - 1) \log(U) - b_U U - \frac{1}{2} \sum_{c=1}^C \langle \delta_c \rangle \left( f(\omega_c, U)^2 - 2 \langle A_c \rangle f(\omega_c, U) \right) \quad (25)$$

This does not represent a standard probability distribution, but we may use numerical methods to estimate the expectation and variance of  $U$ . Based on the posterior  $q(U)$ , we approximate  $\langle f(\omega_c, U) \rangle$  and  $\langle f(\omega_c, U)^2 \rangle$  (required in (22) and (24)) by importance sampling (e.g. [28]).

For each phase we obtain a circular Von Mises distribution,  $q(\phi_c) = \mathcal{M}(\phi_c | \alpha_c, \beta_c)$  (this is an approximation to a

Generalised von Mises distribution of order 2; see [5] for details), where:

$$\alpha_{c,1} = \langle A_c \rangle \sum_{s=1}^S \langle \lambda^{(s)} \rangle \sum_{n=1}^N \cos(\omega_c t_n) \langle q_{c,n}^{(s)} \rangle \quad (26)$$

$$\beta_{c,1} = \langle A_c \rangle \sum_{s=1}^S \langle \lambda^{(s)} \rangle \sum_{n=1}^N \sin(\omega_c t_n) \langle q_{c,n}^{(s)} \rangle \quad (27)$$

$$\text{with } \langle q_{c,n}^{(s)} \rangle = \tilde{y}_n^{(s)} - \sum_{\substack{i=1 \\ i \neq c}}^C \langle A_i \rangle \langle u_{i,n} \rangle \quad (28)$$

2) *Noise posteriors:* For each white noise precision  $\lambda^{(s)}$  we get the Gamma  $q(\lambda^{(s)}) = \mathcal{G}(\lambda^{(s)} | \alpha_{\lambda}, \beta_{\lambda}^{(s)})$ , where

$$\alpha_{\lambda} = a_{\lambda} + N/2 \quad (29)$$

$$\beta_{\lambda}^{(s)} = b_{\lambda} + \frac{1}{2} \sum_{n=1}^N \left[ y_n^{(s)2} - 2 \sum_{c=1}^C \left( y_n^{(s)} \langle A_c \rangle \langle u_{c,n} \rangle + \hat{y}_n^{(s)} \langle B_c^{(s)} \rangle \langle v_{c,n}^{(s)} \rangle \right) + \sum_{c=1}^C \left( \langle A_c^2 \rangle \langle u_{c,n}^2 \rangle + \langle B_c^{(s)2} \rangle \langle v_{c,n}^{(s)2} \rangle + \left( \langle A_c \rangle \langle u_{c,n} \rangle + \langle B_c^{(s)} \rangle \langle v_{c,n}^{(s)} \rangle \right) \cdot \sum_{\substack{i=1 \\ i \neq c}}^C \left( \langle A_i \rangle \langle u_{i,n} \rangle + \langle B_i^{(s)} \rangle \langle v_{i,n}^{(s)} \rangle \right) \right] \quad (30)$$

For each  $1/f$  noise amplitude we end up with  $q(B_c^{(s)})$  as the Gaussian  $q(B_c^{(s)}) = \mathcal{N}(B_c^{(s)} | \mu_c^{(s)}, \sigma_c^{(s)2})$ , where

$$\sigma_c^{(s)2} = \left( \langle \tau_c^{(s)} \rangle + \langle \lambda^{(s)} \rangle \sum_{n=1}^N \langle v_{c,n}^{(s)2} \rangle \right)^{-1} \quad (31)$$

$$\mu_c^{(s)} = \sigma_c^{(s)2} \left( \langle \tau_c \rangle \langle \zeta^{(s)} \rangle \frac{2\pi}{\omega_c \sqrt{N}} + \langle \lambda^{(s)} \rangle \sum_{n=1}^N \langle v_{c,n}^{(s)} \rangle \left( \sum_{\substack{i=1 \\ i \neq c}}^C \hat{y}_n^{(s)} - \langle B_i^{(s)} \rangle \langle v_{i,n}^{(s)} \rangle \right) \right) \quad (32)$$

Following a similar procedure for the precision of each amplitude,  $\tau_c^{(s)}$ , we obtain the Gamma posterior  $q(\tau_c^{(s)}) = \mathcal{G}(\tau_c^{(s)} | \alpha_{\tau}, \beta_{\tau}^{(s)})$ , where

$$\alpha_{\tau} = a_{\tau} + 1/2 \quad (33)$$

$$\beta_{\tau}^{(s)} = b_{\tau} + \frac{1}{2} \left( \langle B_c^{(s)2} \rangle + \frac{4\pi^2 \langle \zeta^{(s)2} \rangle}{\omega_c^2 N} - \frac{4\pi \langle B_c^{(s)} \rangle \langle \zeta^{(s)} \rangle}{\omega_c \sqrt{N}} \right) \quad (34)$$

For each phase  $\psi_c^{(s)}$  we obtain a circular Von Mises distribution,  $q(\psi_c^{(s)}) = \mathcal{M}(\psi_c | \alpha_c^{(s)}, \beta_c^{(s)})$  (see comments regarding  $\phi_c$  above), where:

$$\alpha_c^{(s)} = \langle B_c^{(s)} \rangle \langle \lambda^{(s)} \rangle \sum_{n=1}^N \cos(\omega_c t_n) \langle q_{c,n}^{(s)} \rangle \quad (35)$$

$$\beta_c^{(s)} = \langle B_c^{(s)} \rangle \langle \lambda^{(s)} \rangle \sum_{n=1}^N \sin(\omega_c t_n) \langle q_{c,n}^{(s)} \rangle \quad (36)$$

$$\text{with } \langle q_{c,n}^{(s)} \rangle = \hat{y}_n^{(s)} - \sum_{\substack{i=1 \\ i \neq c}}^C \langle B_i^{(s)} \rangle \langle v_{i,n}^{(s)} \rangle \quad (37)$$

Finally, for each  $1/f$  noise scale variable  $\zeta^{(s)}$  we end up with  $q(\alpha_1^{(s)})$  as the Gaussian  $q(\zeta^{(s)}) = \mathcal{N}(\zeta^{(s)} | \mu_\zeta^{(s)}, \sigma_\zeta^{(s)2})$ , where

$$\sigma_\zeta^{(s)2} = \left( p_\zeta + \sum_{c=1}^C \langle \tau_c^{(s)} \rangle \frac{4\pi^2}{\omega_c^2 N} \right)^{-1} \quad (38)$$

$$\mu_\zeta^{(s)} = \sigma_\zeta^{(s)2} \left( m_\zeta p_\zeta + \sum_{c=1}^C \langle \tau_c^{(s)} \rangle \langle b_c^{(s)} \rangle \frac{2\pi}{\omega_c \sqrt{N}} \right) \quad (39)$$

Thus we end up with a set of expressions that define the approximate posteriors for each of the model variables, but each is dependent on the expected values of one or more of the other variables. Each of the expressions is reassessed iteratively until convergence.

By simple trigonometry, the expectation  $\langle u_{c,n}^2 \rangle$  (and similarly for the  $\langle v_{c,n}^{(s)2} \rangle$ ) may be written as

$$\langle u_{c,n}^2 \rangle = \frac{1}{2} \left( 1 + \langle \cos(2\phi_c - 2\omega_c t_n) \rangle \right) \quad (40)$$

For the iterative procedure we do not need the expectation  $\langle \phi_c \rangle$ , only  $\langle \cos(\phi_c - \omega_c t_n) \rangle$  and  $\langle \cos(2\phi_c - 2\omega_c t_n) \rangle$ , which we calculate numerically. Once the algorithm has converged we may calculate  $\langle \phi_c \rangle$  by rewriting the posterior von Mises in its alternative form:

$$\mathcal{M}(\phi_c | k_c, \mu_c) \propto \exp \left( k_c \cos(\phi_c - \mu_c) \right) \quad (41)$$

$$\text{where } \mu_c = \text{atan2}(\beta_c / \alpha_c) \quad (42)$$

$$k_c = \alpha_c / \cos(\mu_c) \quad (43)$$

In this form  $\langle \phi_c \rangle = \mu_c$ , and  $k_c$  is a concentration parameter that acts like a precision; when  $k_c$  is large the variable is tightly distributed around the expected value; as it becomes smaller the distribution becomes flatter, tending towards the Uniform distribution  $\mathcal{U}(-\pi, \pi)$  as  $k_c$  tends to zero. A similar process is followed for the noise phases in  $\psi_c$ .

#### D. Summary of algorithm

A summary of the processing is shown in algorithm 1. With  $\Delta t$  as the interval between each of the  $N$  observations, we calculate integer  $(N/2) + 1$  angular frequencies uniformly distributed across the range 0 to  $\pi/\Delta t$  (the Nyquist frequency) inclusive, and then remove the 0 value, giving the vector  $\omega$  and the number of components  $C$ . Theoretically the amplitude of  $1/f$  noise at a frequency of 0 is infinite, so we avoid the problem by mean-centring the observations and removing the 0 frequency. The time vector  $\mathbf{t}$  is the vector of elapsed time since the first observation. The expected values of the amplitudes are initialised to the theoretical smooth shapes obtained from the prior expectations  $\langle U \rangle$  and  $\langle \zeta \rangle$ , the phases are initialised to random values within appropriate ranges and the prior

---

#### Algorithm 1 Process for estimating the posterior distributions.

---

define the frequencies,  $\omega$ , and prior hyperparameters  
initialise the values of each variable

**while** not converged **do**

    update the signal posteriors using (21–27)

**for** each observation set,  $s$  **do**

        update the posteriors for  $B_c^{(s)}$  and  $\tau_c^{(s)}$  using (31–34)

        update the posterior for  $\zeta^{(s)}$  using (38–39)

        update the posteriors for each  $\psi_c^{(s)}$  using (35–36)

        update the posterior for  $\lambda^{(s)}$  using (29–30)

**end for**

**end while**

calculate means & concentrations of  $\phi_c$  and  $\psi_c$  (41–43)

---

hyperparameters for the observation noise are set. The process then iteratively recalculates the posterior distributions for each variable in turn, until they converge. Finally the mean and concentration parameters for each of the phases are calculated.

The amplitude distributions are Gaussian, which admits the possibility of negative values. This is not in itself a problem as a component with amplitude  $-a$  and phase  $\phi$  is equivalent to a component with amplitude  $+a$  and phase  $\phi + \pi$ . However, amplitudes with small magnitudes may oscillate between negative and positive values during the iterative process, causing big jumps in their associated phases and the model does not converge properly. To prevent this we use the absolute value of  $\langle A_c \rangle$  in place of  $\langle A_c \rangle$  (and similarly for the  $B_c^{(s)}$ ) in each of the posterior expressions.

We have chosen to initialise the variables randomly to demonstrate that the model converges to a good solution. We could converge more quickly if we chose to initialise the amplitudes and phases from the results of a Fast Fourier Transform (FFT).

### III. ILLUSTRATION: SYNTHETIC DATA

In this section we describe results obtained from synthetic data for which we know the actual values of each of the variables. In the first test the angular frequencies used to train the model are identical to those used to generate the test data, so that we may compare the estimated phases with the truth. The second test is more realistic, with the model frequencies being a small subset of the true frequencies.

#### A. Test 1

For the first test, in each case there are 200 observations, at 1 Hz, giving the set of 101 Fourier angular frequencies  $\omega$  in the range 0 to  $\pi$ . We start with 1,000 signals each generated from sea-shape spectra with wind speed values drawn from  $U \sim \mathcal{U}(0, 50)$  (0 to 112 mph) and phases from  $\phi_c \sim \mathcal{U}(-\pi, \pi)$ . Each signal is used to generate two sets of observations by adding  $1/f$  noise with scale value drawn from  $\zeta^{(*)} \sim \mathcal{U}(0, 2)$  and phase from  $\psi_c^{(*)} \sim \mathcal{U}(-\pi, \pi)$ , and white noise with variance drawn from  $1/\lambda^{(*)} \sim \mathcal{U}(0, 2)$ . Although the model assumes that the  $1/f$  noise shape value (the  $\gamma$  in (1)) is 1, the synthetic data are generated with  $\gamma$  drawn from  $\mathcal{U}(0.9, 1.1)$ .

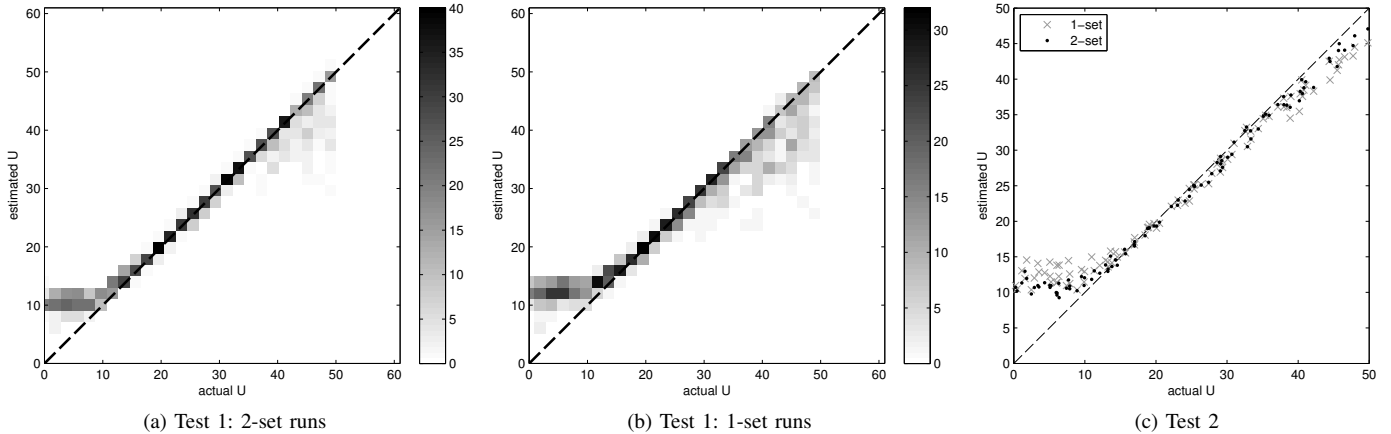


Figure 4 Wind speed. (a) and (b) show two-dimensional histograms from test 1, showing the actual *vs* estimated values of the wind speed variable  $U$  for the 2-set and 1-set runs respectively. (c) shows a plot of the actual *vs* estimated  $U$  from test 2, with black dots marking the 2-set runs and grey crosses the 1-set runs. The prior for  $U$  gives a mean value of 30.

Uninformative priors are used for the white noise precision and  $1/f$  noise scale variable, with

$$a_\lambda = b_\lambda = m_\zeta = p_\zeta = 10^{-6} \quad (44)$$

To promote adherence of the estimated signal and  $1/f$  noise spectra to the desired shapes, tighter, more informative priors are used for the amplitude precisions:

$$a_\delta = a_\tau = 1000 \quad b_\delta = b_\tau = 1 \quad (45)$$

and the wind speed is assigned a mildly informative prior:

$$a_U = 3 \quad b_U = 0.1 \quad (46)$$

For each test case the model is run twice, once with a single set of observations (“1-set”) and once with both sets of parallel observations (“2-set”). In figure 4, (a) and (b) show histograms of actual *vs* estimated values of the wind speed variable  $U$  across the 1,000 synthetic datasets. For both the 1-set and the 2-set runs, below a true value for  $U$  of about 10 m/s the noise is dominating the signal and the estimate is poor, but above this there is generally good correspondence. The effect of the prior over  $U$ , which has a mean of 30, can be seen in both runs, with  $U$  tending to be underestimated above this true value, but the 1-set runs are more prone to underestimating  $U$ .

Figure 5 shows histograms of actual *vs* estimated values of the signal phase  $\phi_c$ , comparing the 2-set and 1-set runs with the results from standard FFT. Only phases where the actual amplitude is greater than 0.1m are shown; if the amplitude is zero then the phase is arbitrary. The impact of noise is clearly seen in the FFT result, where there is no correspondence between actual and estimated values. The model is clearly capturing approximately the correct phases in both cases; the cleaner, more clearly-defined diagonal in (a) indicates that basing the model on two sets of parallel observations gives a better result than basing it on only one.

## B. Test 2

So far the model has been trained on the same set of frequencies that the observations were generated from. In the second, more realistic, sea-like test, 100 signals are generated, each of 100,000 data points, from 50,000 angular frequencies (not including 0), and each signal perturbed by two different sets of noise to generate pairs of observation sets. The model is trained on the first 1,000 observations, with just 500 frequencies, or 1% of the true number.

Figure 4c is a plot of actual *vs* estimated wind speed values  $U$ . As before,  $U$  is well estimated once the signal emerges from the noise at about 10 m/s. The 1-set model has a greater tendency to overestimate  $U$  when the signal is dominated by noise and to underestimate it where the wind speed is higher. This may be because the 1-set model has less evidence on which to base a deviation from the prior.

Figure 6 compares the true signal and noise spectra (grey lines; scaled to take account of the different number of frequencies) with the signal spectrum estimated by the model (black dots) and the amplitudes resulting from standard FFT (grey dots) for one of the 2-set test runs. Although the noise dominates the signal, the model is closely approximating the true signal.

## IV. REAL DATA

The model has been demonstrated on both single sets of observations and parallel pairs of observations, but there is nothing to prevent its use in the more general case where there are multiple parallel sets of observations. Clearly in its current form it will never be able to compete with FFT in terms of speed, but it provides a basis for estimating signals where the theoretical spectra are smooth and of a known shape, and the observations are contaminated with pink and/or white noise. The Bayesian method of combining an observation with prior information points to its potential for processing

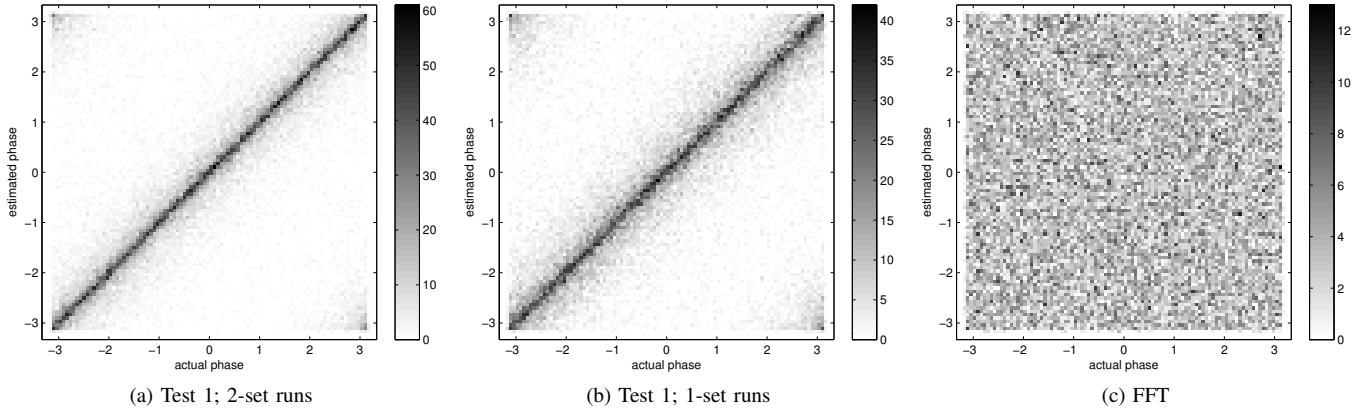


Figure 5 Test 1. Two-dimensional histograms showing the actual vs estimated values of the signal phase,  $\phi_c$ , for all contributions where the actual signal amplitude is greater than 0.1 m, from (a) the 2-set runs, (b) the 1-set runs, and (c) a standard FFT.

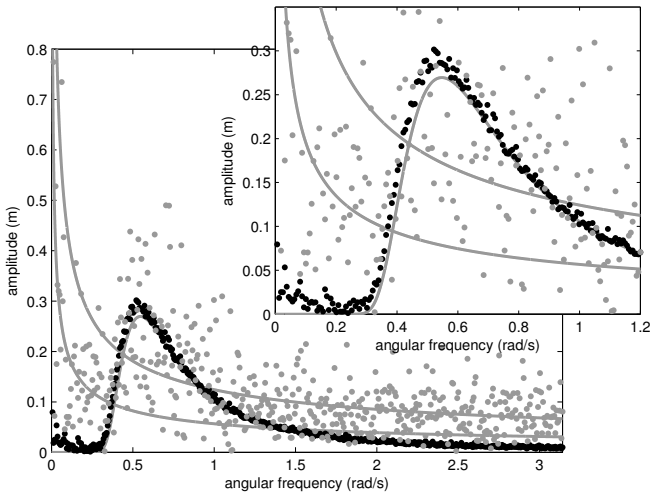


Figure 6 Test 2. A comparison of the true spectra (scaled to take account of the different number of frequencies) and those estimated by FFT and the model. The inset is a close-up of the critical region. The true signal and true noise spectra are shown as grey lines. Overlaid are the spectra obtained from FFT (grey dots) and the model (black dots). The true value of  $U$  is 15.5896; the value of  $\langle U \rangle$  estimated by the model is 16.0628.

observations one at a time, as they are received. For our needs an observation rate of 1 Hz is sufficient.

Testing this model in real sea conditions requires a single measurement device (for example, a buoy) containing three independent wave profile sensors, one low-noise (these tend to be expensive) and a pair exhibiting the effects of pink noise. Feeding the observations from the latter into the model results in an estimate of the underlying signal that can then be verified against the low-noise sensor's observations. An alternative method is to construct a buoy that contains the noisy sensors, train the model as before and use the estimated signal to predict the motion of the low-noise sensor using the method described in [4]. The danger with placing both noisy

sensors within the same device is that the noise may no longer be independent between the two.

We have a small amount of data (776 observations recorded at 2 Hz) recorded from two separate sensor buoys together in a small, rigid-hulled boat at some distance from a ship carrying a low-noise sensor. Apart from the small number of observations, there are two reasons why these data are not ideal. The first is that the waves were coming from two sources, travelling in two slightly different directions, each producing a small spread of wave directions. This means that the true wave spectrum is a mixture of slightly different Pierson-Moskowitz spectra while the model assumes only one. The second is that the boat was moving; this leads to encounter phase shifts in the component waves which are not represented in the model. The relatively small size of the boat and ship relative to the amplitude/period of the waves means that a third potential source of error, due to the assumption that they act as single points bobbing on the sea's surface, is unlikely to cause appreciable effects.

The model was trained on the parallel observations of noisy GPS altitudes recorded by the two buoys. The first 250 s of estimated signal were used to predict the motion at the ship's motion sensor for the subsequent 30 s. The 250 s window was then advanced by 0.5 s (one observation) and a second set of predictions obtained. This sliding window process resulted in 217 sets of predictions based on the model's estimate of the signal. The same process was followed using one buoy's bandpass filtered inertial sensor measurements (the accepted output from these buoys). Thus we may compare the quality of predictions from the model and the filtered observations. Figure 7 shows the linear correlations achieved by each method between the predictions and the ship's sensor observations. As in [4] the quality varies over time, but in general the model has provided the better predictive quality.

Figure 8 shows a period of truth from the low-noise sensor overlaid with the multiple predictions obtained from the moving-window method based on (top) the model's estimate of the signal and (bottom) the buoy's filtered observations.

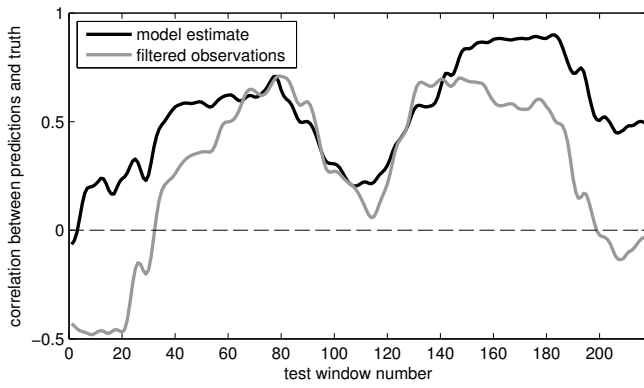


Figure 7 Real data. A comparison of the prediction quality of the model (black) and bandpass filtered observations (grey). These are the linear correlations between predictions and truth for the 217 predictions obtained using the moving-window method (see text).

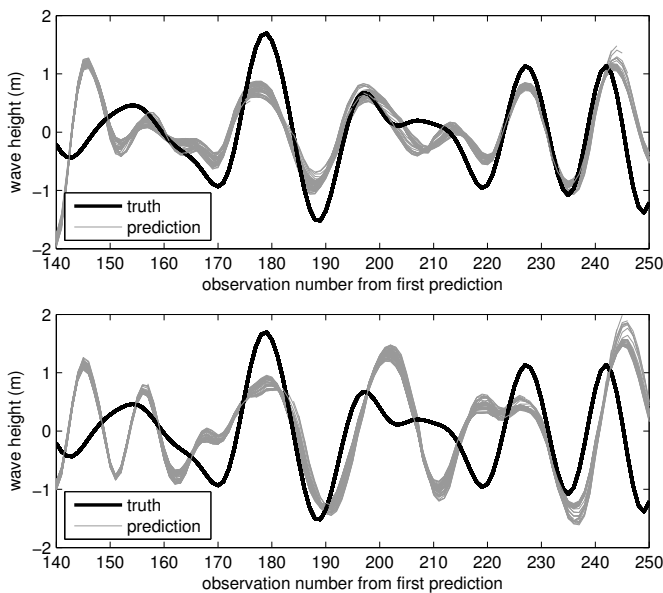


Figure 8 Real data. A comparison of the predictions obtained from (top) the model and (bottom) bandpass filtered observations for a section of the data, using the moving-window method (see text).

This is a region of particularly good prediction quality and we can see that the predictions based on the model's estimated signal is, in this case, better than those based on the filtered observations.

These results are not presented as a confirmation of the model as the prediction process is a current area of research, but they show the potential for the model to benefit the prediction of sea waves.

#### ACKNOWLEDGEMENT

The author wishes to thank Dr John Duncan (DE&S, Ministry of Defence) for providing data used in section IV which were recorded during an NSRS exercise Golden Arrow in November 2014.

#### REFERENCES

- [1] J. Falnes, *Ocean Waves and Oscillating Systems*. Cambridge University Press, 2002.
- [2] M. Belmont, "A lower bound estimate of the gains stemming from quiescent period predictive control using conventional sea state statistics," *Journal of Renewable and Sustainable Energy*, vol. 1, no. 6, 2009.
- [3] —, "Increases in the average power output of wave energy converters using quiescent period predictive control," *Renewable Energy*, vol. 35, no. 12, pp. 2812–2820, December 2010.
- [4] M. Belmont, J. Christmas, J. Dannenberg, T. Hilmer, J. Duncan, J. Duncan, and B. Ferrier, "An examination of the feasibility of linear deterministic sea wave prediction in multi-directional seas using wave profiling radar: Theory, simulation and sea trials," *Journal of Atmospheric and Oceanic Technology*, vol. 31, no. 7, pp. 1601–1614, 2014.
- [5] J. Christmas, "Bayesian spectral analysis with Student-t noise," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2871–2878, 2014.
- [6] F. Bach and M. Jordan, "A Probabilistic Interpretation of Canonical Correlation Analysis," Department of Statistics, University of California, Berkeley, California, Tech. Rep. 688, 2005.
- [7] B. Mandelbrot and J. Wallis, "Some long-run properties of geophysical records," *Water Resources Research*, vol. 5, no. 2, pp. 321–340, 1969.
- [8] J. Halley, "Ecology, evolution and 1/f noise," *Trends in Ecology and Evolution*, vol. 11, no. 1, pp. 33–37, 1996.
- [9] R. Voss and J. Clark, "1/f noise in music: music from 1/f noise," *Journal of the Acoustical Society of America*, vol. 63, pp. 258–263, 1976.
- [10] T. Nagatani, "Power-law distribution and 1/f noise of waiting time near traffic-jam threshold," *Journal of the Physical Society of Japan*, vol. 62, pp. 2533–2536, 1993.
- [11] J. Johnson, "The Schottky effect in low frequency circuits," *Physical Review*, vol. 26, pp. 71–85, 1925.
- [12] M. Keshner, "1/f noise," *Proceedings of the IEEE*, vol. 70, no. 3, pp. 212–218, 1982.
- [13] A. Balandin, "Low-frequency 1/f noise in graphene devices," *Nature Nanotechnology*, vol. 8, pp. 549–555, 2013.
- [14] A. Quinchia, G. Falco, E. Falletti, F. Dovis, and C. Ferrer, "A comparison between different error modeling of MEMS applied to GPS/INS integrated systems," *Sensors*, vol. 13, pp. 9549–9588, 2013.
- [15] G. Wornell and A. Oppenheim, "Estimation of fractal signals from noisy measurement using wavelets," *IEEE Transactions on Signal Processing*, vol. 40, no. 3, pp. 611–623, 1992.
- [16] A. Noureldin, J. Armstrong, A. El-Shafie, T. Karamat, D. McGaughey, M. Korenberg, and A. Hussain, "Accuracy enhancement of inertial sensors utilizing high resolution spectral analysis," *Sensors*, vol. 12, pp. 11 638–11 660, 2012.
- [17] M. Korenberg, "A robust orthogonal algorithm for system identification and time-series analysis," *Biological Cybernetics*, vol. 60, pp. 267–276, 1989.
- [18] D. Allan, "Statistics of atomic frequency standards," *Proceedings of the IEEE*, vol. 54, no. 2, p. 221230, 1966.
- [19] N. El-Sheimy, H. Hou, and X. Niu, "Analysis and modeling of inertial sensors using Allan variance," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 1, pp. 140–149, 2008.
- [20] B. Kaulakys, "Autoregressive model of 1/f noise," *Physics Letters A*, vol. 257, no. 1, pp. 37–42, 1999.
- [21] M. Jordan, Z. Ghahramani, T. Jaakkola, and L. Saul, "An introduction to variational methods for graphical models," *Machine Learning*, vol. 37, no. 2, pp. 183–233, 1999.
- [22] H. Lappalainen and J. Miskin, "Ensemble learning," in *Adv. in Independent Component Analysis*. Berlin: Springer-Verlag, 2000, pp. 75–92.
- [23] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- [24] J. Christmas, "The effect of missing data on robust Bayesian spectral analysis," in *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Southampton, UK, 2013.
- [25] M. Ochi, *Ocean Waves, the stochastic approach*, ser. Ocean Technology Series. Cambridge University Press, 2005, vol. 6.
- [26] W. Pierson Jr and L. Moskowitz, "A proposed spectral form for fully developed wind seas based on the similarity theory of S.A. Kitaigorodskii," *J. of Geophysical Research*, vol. 69, no. 24, pp. 5181–5190, 1964.
- [27] H. Attias, "A variational Bayesian framework for graphical models," *Adv. in Neural Info Processing Systems*, vol. 12, pp. 209–215, 2000.
- [28] J. Geweke, "Bayesian inference in econometric models using Monte Carlo integration," *Econometrica*, vol. 57, pp. 1317–1339, 1989.