

A Review of Lagrangian Time Series Models for Ocean Surface Drifter Trajectories (Sykulski et al. (2016))

Hannah Director

Department of Statistics, University of Washington Seattle, WA, 98195, USA

Abstract

This report reviews the spectral analysis method for modeling ocean surface drifters proposed by Sykulski et al. [2016]. Previous methods to model drifters are reviewed along with the authors' model. Where relevant to understanding the model of Sykulski et al. [2016], spectral analysis, Matérn covariances, the complex Ornstein-Uhlenbeck process, and Whittle likelihood are considered. We conclude by evaluating the strengths and weaknesses of this method.

1 Introduction

“Lagrangian Time Series Models for Ocean Surface Drifter Trajectories,” by Adam M. Sykulski, Sofia C. Olhede, Jonathan M. Lilly, and Eric Danioux presents a multi-component spectral model for analysis of data transmitted by ocean surface drifters. Drifters are free-floating instruments that transmit their location at regular time intervals, creating a *Lagrangian time series*, or a sequence of spatial locations over time. Oceanographers use these time series to increase understanding of ocean circulation patterns. Published in January 2016 in the *Journal of the Royal Statistical Society Series C*, this paper develops a model that can isolate a scientific quantity of interest, and, in so doing, introduces several new statistical techniques.

The motion of parcels of water moving over changing latitudes is known to have a rotational component. Because the Earth has different diameter at different latitudes, the speed of objects moving with Earth's rotation varies across latitudes. This induces a phenomena known as the Coriolis Effect, whereby very fast moving objects or objects being observed over long periods, such as drifters, end up moving at speeds that are different than the Earth beneath them as they change latitude. Viewed from a stationary reference frame, this

creates circular motions, or inertial oscillations. These motions have known frequencies, referred to as inertial frequencies. Oceanographers are interested in detecting deviations from these frequencies, as they are thought to indicate eddies, or persistent circular wave patterns [Kunze, 1985]. However, the ocean has a constant turbulent background flow which makes identifying eddies in real data difficult [Elipot et al., 2010]. Previous methods to model drifters are BLAH The methods in this paper overcome this challenge and successfully find shifts in the inertial frequency with a spectral time series model.

In achieving this result, several statistical advances are made. The authors introduce an additive spectral domain model that has components corresponding to both the turbulent background and the inertial oscillations. For the turbulent background model, the Matérn covariance commonly used in spatial statistics [Gneiting et al., 2012] is extended to the time-series context. This allows for a more flexible model for background behavior than previously obtained. Inertial oscillations are modeled with a complex Ornstein-Uhlenbeck (OU) [Arató, 1962, Jeffreys, 1968], which provides a stochastic equivalent to an accepted set of coupled differential equations describing inertial oscillations. Further statistical innovation is employed in fitting the model with a variation on standard Whittle likelihood to reduce bias and in allowing for non-stationarity through time-varying parameter estimates.

The remainder of this review proceeds as follows. After a succinct review of spectral analysis, Section 2 discusses the model for the drifters and the methods used to fit it. Section 3 presents the paper’s results including analysis of a long time series with time-varying parameters and an application from a simulated model. The paper concludes with a discussion of the value and limitations of the approach taken by Sykulski et al. [2016].

2 Methods

2.1 Spectral Analysis

We briefly review spectral analysis as these methods are central to understanding Sykulski et al. [2016]. More thorough coverage of this material can be found in [Percival and Walden, 1993], for example. From a statistical perspective, a time series, $x(t)$ can be understood as

a realization of a stochastic process over time, where $t = \{1, 2, \dots, n\}$ are the observed time points. Time series are often modeled as the sums of periodic functions, which, using Euler's formula, can be represented compactly as Fourier series. Moreover, data in the time domain can easily be related to data in the frequency domain and vice versa using the Fourier or inverse Fourier transformations

$$x(t) \stackrel{ms}{=} \int_{-\infty}^{\infty} f_x(\omega) e^{i\omega t} d\omega \quad f_x(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt. \quad (1)$$

[Percival and Walden, 1993]. To understand the distribution of frequencies, the *power spectral density*,

$$S_x(\omega) = \lim_{T \rightarrow \infty} \mathbb{E} \left(\frac{1}{2T} \left| \int_{-T}^T x(t) e^{-i\omega t} dt \right|^2 \right), \quad (2)$$

is often employed. This quantity represents the variance (or power) associated with each frequency [Percival and Walden, 1993]. Power spectral densities are useful for statistical modeling, since they can be related to the autocovariance sequence of a time series using Fourier transformation. Defining the autocovariance as $s_x(\tau) = \mathbb{E}[x(t)x(t - \tau)]$ where τ is the time lag, we have

$$S_x(\omega) = \int_{-\infty}^{\infty} s_x(\tau) e^{-i\omega \tau} d\tau \iff s_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) e^{i\omega \tau} d\omega$$

[Sykulski et al., 2013]. For an observed time series, the power spectral density is often estimated with the periodogram, $\hat{S}_x(\omega) = |J_x(\omega)|^2$ where

$$J_x(\omega) = \sqrt{\frac{t}{N}} \sum_{t=1}^N x(t) e^{-i\omega t} \quad (3)$$

[Sykulski et al., 2013]. Although intuitively appealing, this estimator has known problems that stem from the sampling of a periodic function at discrete intervals. In particular, high frequencies can capture periodic behavior of other lower frequencies if that happen to align in period. Additionally, frequencies above the highest observable frequency cannot be captured, and, are instead, recorded as other frequencies in the spectrum. These effects, referred to as *aliasing* and *leakage* respectively, are typically reduced by *tapering*. This method uses a window of frequencies for estimation rather than a single value. However, depending on how tapering is used, it can introduce its own biases in estimation.

2.2 Model

To model the drifter's motion over time, the eastward and northward components of the drifter's velocity, denoted $u(t)$ and $v(t)$ respectively, are converted to a complex-valued velocity, $z(t) = u(t) + iv(t)$. Stochastic models are then developed for both components of the drifter's behavior, inertial oscillations and movement due to the ocean's turbulent background, in the spectral domain. Finally, the two components of the model are added together, creating a six-parameter model. Parameter estimates for this model are obtained via maximizing the model's blurred Whittle likelihood. This is a relatively new approximation to the true spectral likelihood that balances computational efficiency with minimizing bias [Sykulski et al., 2013].

2.2.1 Inertial oscillations

We begin by discussing the model for the inertial oscillations. In a deterministic setting, inertial oscillations are often modeled with the following set of coupled differential equations [Pollard and Millard, 1970]

$$\begin{aligned}\frac{\partial u}{\partial t} + f_0 v &= F - cu \\ \frac{\partial v}{\partial t} - f_0 u &= G - cv\end{aligned}\tag{4}$$

where u and v again represent eastward and northward velocities, f_0 represents the inertial frequency in radians per unit time, and F and G are forces related to the wind. Suppressing the dependence on t and using the complex representation of the velocity, we show how these relationships can be equivalently expressed as follows

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial u}{dt} + i \frac{\partial v(t)}{dt} \\ dz &= (F - cu - f_0 v)dt + (iG - icv + if_0 u)dt \\ &= f_0(-v + iu)dt - c(u + iv)dt + (F + iG)dt \\ &= if_0(u + iv) - c(u + iv) + (F + iG) \\ &= (if_0 - c)zdt + (F + iG)dt.\end{aligned}$$

By further replacing the wind forcing term, $(F + iG)dt$, with complex-valued Brownian increments, Sykulski et al. [2016] obtain the stochastic model

$$dz(t) = (if_0 - c)z(t)dt + AdQ(t), \quad (5)$$

which defines equivalent behavior to the accepted deterministic formulation. Equation 5 is known as the complex-valued Ornstein-Uhlenbeck (OU) stochastic differential equation. It developed from brownian motion to OU to complex OU, mean-reversion, dampening and is a valid in this context by blah (look up and cite the Mandelrot and Van Ness justification) giving the This model is always Markovian. The OU model is stationary assuming the initial conditions of the stochastic differential equation are set as $z(t_0) \sim N(0, A^2/2|c|)$, and, if not, will revert to the stationary model quickly as the influence of the initial conditions decays exponentially (WHY? look this up).

Sykulski et al. [2016] further let f_0 be a free parameter, denoted ω_0 , to be estimated by data. In so doing, they enable identification and estimation of the shifts in the angular frequency due to eddies. If the estimated angular frequency differs from the inertial frequency, the angular frequency of the eddy can be calculated as $\omega_{eddy} = f_0 - \omega_0$, since ω_0 is a sum of the inertial and eddy frequencies.

Letting $z^*(t)$ represent the complex conjugate of $z(t)$ and using the previously-specified initial conditions, the complex OU process has the following autocovariance and power spectral density

$$s^{(0)}(\tau) = \mathbb{E}\{z(t)z^*(t + \tau)\} = \frac{A^2}{2c} \exp(i\omega_0\tau) \exp(-c|\tau|) \quad (6)$$

$$S^{(0)}(\omega) = \int_{-\infty}^{\infty} s^{(0)}(\tau) \exp(-i\omega\tau) d\tau = \frac{A^2}{(\omega - \omega_0) + c^2} \quad (7)$$

where τ is a time lag and the (o) subscripts denotes that these equations are for the OU component of the model. In addition to the inertial frequency, ω_0 , the parameter, $A > 0$ that defines the variability, and the parameter, $c > 0$ which controls the dampening, need to be fit from data.

2.2.2 Turbulent Background

The drifters' time series are also marked by large-scale turbulence (read and cite Rhines 1979). Many potential models have been considered for this background....Lit review here ... However, these models have been *integer order*, meaning the velocity, acceleration, and hyperacceleration must be Markovian, a constraint previous studies (READ AND CITE) indicate is not appropriate in all cases. Thus, the authors propose using the Matérn covariance structure [Gneiting et al., 2012], a model which has gained prominence in the spatial statistics literature due to its flexibility. In this context, the Matérn model usefully allows for markovianity when the data supports it while not universally enforcing it. The corresponding autocovariance and power spectral density are (read and cite Stein 1991 here)

$$s^{(m)}(\tau) = \frac{B^2}{2^{\alpha-1/2}\pi^{1/2}\Gamma(\alpha)h^{2\alpha-1}}(h|\tau|)^{\alpha-1}\kappa_{\alpha-1/2}(h|\tau|) \quad (8)$$

$$S^{(m)}(\omega) = \frac{B^2}{(\omega^2 + h^2)^\alpha} \quad (9)$$

where τ again denotes the time lag, $\Gamma(\alpha)$ is the gamma function, κ_η is a modified Bessel function of the second kind of order η . and the subscript (m) indicates that this is the Matérn part of the model. The parameters to be estimated are the amplitude, B , the dampening, $h > 0$, and a smoothness parameter $\alpha > 1/2$.

2.2.3 Overall Model

Combining the inertial oscillation and turbulent background parts of the model is straightforward. To find both the autocovariance and power spectral densities, we can simply add the corresponding components of the inertial oscillation and turbulent background. Using Equations 6 and 8 and Equations 7 and 9, we obtain

$$S(\omega; \boldsymbol{\theta}) = \frac{A^2}{(\omega - \omega_0)^2 + c^2} + \frac{B^2}{(\omega^2 + h^2)^\alpha} \quad (10)$$

$$s_\tau(\boldsymbol{\theta}) = \frac{A^2}{2c} \exp(i\omega_0\tau) \exp(-c|\tau|) + \frac{B^2}{2^{\alpha-1/2}\pi^{1/2}\Gamma(\alpha)h^{2\alpha-1}}(h|\tau|)^{\alpha-1}\kappa_{\alpha-1/2}(h|\tau|). \quad (11)$$

2.3 Fitting the Model

A relatively new technique referred to as blurred Whittle likelihood is employed to fit these models. Blurred Whittle likelihood builds on Whittle likelihood, a standard approximation to the true likelihood in the time-domain which improves computation from order $O(N^2)$ to $O(N)$. Whittle likelihood is a technique developed to maximize. Unfortunately, blurred whittle likelihood shows.

To allow for non-stationarity, the authors propose fitting the model independently at each observed time point using a rolling window around the time point of interest. The length of the window needs to be large enough to allow for a full sampling of the spectrum, but short enough to be treated as locally stationary. In this analysis, the authors select the window length empirically, but acknowledge that further work on this subject would be of value.

Sykulski et al. [2016] propose that a simple likelihood ratio test can be used at each time period to test whether or not a shift in the inertial oscillation has occurred. To do so, the authors suggest fitting a simpler model where the parameter ω_0 is fixed to be the inertial frequency f . Then this simpler model can be compared to the full six parameter model via a likelihood ratio test with test statistic

$$R(t) = 2[l_b\{\hat{\theta}_1(t)\} - l_b\{\hat{\theta}_0(t)\}] \quad (12)$$

where $\hat{\theta}_0(t)$ is the fitted parameter vector for the simpler, or null model, and $\hat{\theta}_1(t)$ is the fitted parameter vector for the full model. The authors claim $R(t)$ then follow a Chi-Squared distribution with 1 degree of freedom.

3 Results

We show the potential utility of this method by analyzing two non-trivial data sets. First, we consider the modeled trajectories of 200 near surface particles obtained from a computer simulation of ocean circulation (Set to be ‘very similar’ to Danioux et al. [2008] SPEM 5.2 set-up). In Figure 3, we display the fit of the aggregate spectral model to a single trajectory from the simulation and to an ensemble of trajectories. The resulting spectra has peaks at zero and at the inertial frequency, which indicates that the model captures the turbulent

background and the inertial oscillations, which are set not to have shifts. We also display the individual periodograms and model fits. Sykulski et al. [2016] note that this highlights the method’s robustness, since the individual periodograms show considerable variability, but the model fits are all quite consistent.

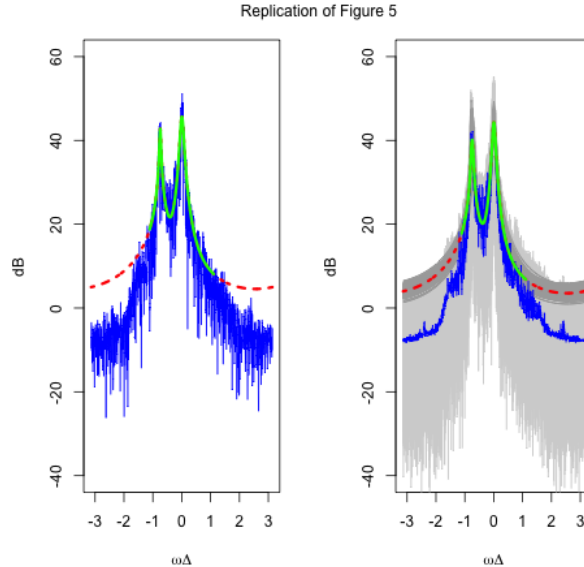


Figure 1: Left: The spectral density of one particle (blue) with the model fit overlaid for the portion of the spectra used to fit it (green) and extended to all frequencies (red). Right: The ensemble mean periodogram (blue) with the ensemble fit for the frequencies used to fit it (green) and extended to the entire spectrum (red) and individual trajectories’ periodograms (light grey) and model fits (dark fit)

We also consider the trajectory of a drifter in the South Pacific Ocean observed 12 times a day over a 1642 day period. Using a rolling window of 1000 observations and the semi-parametric set-up described in Section X, we obtain parameter estimates for the spectral model at each time point. In Figure 3, we compare the observed spectral density over time and the fitted spectral density. The fitted density appears to be a smoothed version of the observed density, indicating a reasonable model fit. We also display the fitted parameter values over time and their 95% confidence interval bands in Figure 8 (add fig and renumber) and report the correlation among variables in Figure 10 (and fig and renumber).

We also fit a 5-parameter where the frequency is set to the inertial frequency rather than

being a free-parameter. We then compare the fit to the full model using a likelihood ratio test as described in Section X. In Figure 9, we display the test statistic over time where a red line is used to indicate the level of statistical significance. We conclude for time periods where the test statistic is greater than this cut-off that there is a shift in the inertial frequency.

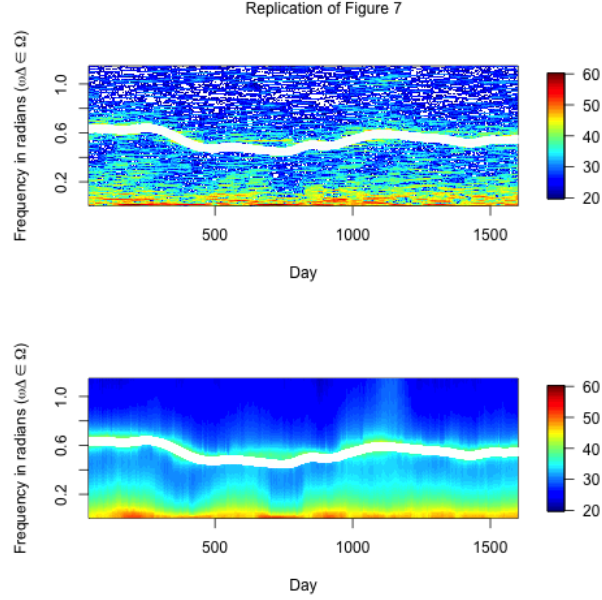


Figure 2: Top: Observed spectra of the drifter over time; Bottom: Modeled spectra over time. On both figures, the white line indicates the average inertial frequency within the rolling window at the current time period. Only frequencies used in the estimation are included.

4 Discussion

identifiability

References

- M. Arató. Estimation of the parameters of a stationary gaussian markov process. In *Doklady Akademii Nauk*, volume 145, pages 13–16, 1962.
- E. Danioux, P. Klein, and P. Rivière. Propagation of wind energy into the deep ocean through a fully turbulent mesoscale eddy field. *Journal of Physical Oceanography*, 38(10): 2224–2241, 2008.
- S. Elipot, R. Lumpkin, and G Prieto. Modification of inertial oscillations by the mesoscale eddy field. *Journal of Geophysical Research: Oceans*, 115(C9), 2010.
- T. Gneiting, W. Kleiber, and M. Schlather. Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 2012.
- H. Jeffreys. The variation of latitude. *Monthly Notices of the Royal Astronomical Society*, 141(2):255–268, 1968.
- E. Kunze. Near-inertial wave propagation in geostrophic shear. *Journal of Physical Oceanography*, 15(5):544–565, 1985.
- D.B. Percival and A.T. Walden. *Spectral analysis for physical applications*. 1993.
- R.T. Pollard and R.C. Millard. Comparison between observed and simulated wind-generated inertial oscillations. In *Deep Sea Research and Oceanographic Abstracts*, volume 17, pages 813–821, 1970.
- A.M. Sykulski, S.C. Olhede, J.M. Lilly, and J.J. Early. On parametric modeling and inference for complex-valued time series. *arXiv preprint arXiv:1306.5993*, 2013.
- A.M. Sykulski, S.C. Olhede, J.M. Lilly, and E. Danioux. Lagrangian time series models for ocean surface drifter trajectories. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 65(1):29–50, 2016.

5 Appendix

5.1 Errata

5.2 Additional Figures

5.3 Computation