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

Put your project summary here.

1 Introduction

1.1 Application

1.2 Data Description

Figure 1 included here

1.3 Oceanography background

- inertial oscillations
- turbulent background

1.4 Spectral analyses of time series

TO DO: vary who I cite here, unify notation across sources

Sykulski et al. [2016] exclusively employs spectral analysis to model the drifters. To provide context for the methods they propose, this section summarize the main concepts of this time series analysis method. More thorough coverage of this material can be found in [Percival and Walden, 1993], for example. From a statistical perspective, a time series, $X_t = X(t\Delta)$, where $\Delta > 0$ is the sampling interval and $t = \{1, 2, ..., n\}$, can be understood as a realization of a stochastic process over time. Spectral analysis models time series, not in their original time domain, but instead in a frequency domain. This is possible because time series can be represented as infinite sums of exponential functions, known as a spectral representation. Following the presentation of Percival and Walden [1993], the idea of spectral representation can be understood by initially considering the harmonic process,

$$X_t = \sum_{l=1}^{L} D_l cos(2\pi f_l t + \phi_l), t = 0, \pm 1, \pm 2..., \tag{1}$$

where L and D_l, and f_l are real-valued constants, ϕ_l are independent Uniform$(-\pi, \pi)$ random variables, and the frequencies, $0 < f_l < 1/2$, are taken to be increasing order. As the number of sinusoidal functions increases, more complex behavior can be represented by the sum in Equation 1. Applying Euler's formula, Equation 1, can be written as

$$X_t = \sum_{l=-L}^{L} C_l e^{i2\pi f_l t}. \tag{2}$$

which as $L \to \infty$ gives

$$X_t = \int_{-1/2}^{1/2} e^{i2\pi f t} dZ(f), \tag{3}$$

Equation 3 can be used to define any real or complex-valued discrete parameter stationary process, X_t, with mean zero. Specifically, there exists a process $\{Z(f)\}$ on $[-1/2, 1/2]$, where the intervals $dZ(f)$ and $dZ(f')$ are uncorrelated for all f, f', that satisfies Equation 3 for all t[Percival and Walden, 1993]. This result is known as the spectral representation theorem for discrete parameter stationary processes. An analogous result holds for continuous parameter stationary processes [Percival and Walden, 1993].

The relationship between spectral densities and autocovariances also contributes to the utility of analysis in this domain. In particular, the spectral densities of a time series and the autocovariance sequence are the Fourier transform of one another. In this paper, the focus is on second order stationary processes, or cases where the autocovariance, s_{xx}, is only a function of the distance between the increments, meaning $c_{xx}(t_1, t_2) = \mathbb{E}(X_{t1} X_{t2}) = s_{xx}(|t_1 - t_2|)$. (add derivation here?) In such cases, this relationship is written as

$$S_{xx}(\omega) = \Delta \sum_{\tau=-\infty}^{\infty} s_{xx}(\tau) e^{i\omega\tau\Delta} \tag{4}$$

and

$$s_{xx}(\tau) = \frac{1}{2\pi} \int_{-\pi/\Delta}^{\pi/\Delta} S_{xx}(\omega) e^{i\omega\Delta} d\omega. \tag{5}$$

where $\omega \in [-\pi/\Delta, \pi/\Delta]$ is the angular frequency measured in radians [?] . In the complex-valued case, s_{xx} and S_{xx} are replaced with s_{xy} and S_{xy} to represent the cross-covariance terms [?].

Given an observed time series X_t, $t = 1, 2, .., N$, the common estimator for the spectra is the periodogram, $\hat{S}_{xx}(\omega) = |J_x(\omega)|^2$ where

$$J_x(\omega) = \sqrt{\frac{\Delta}{N}} \sum_{t=1}^{N} X_t e^{-i\omega t\Delta}. \tag{6}$$

Although simple, this estimator has two known issues: aliasing and leakage. Aliasing is ... Leakage is... Tapering has been proposed to fix, but has problems..

2 Methods

2.1 Model

2.1.1 Inertial oscillations

- Ornstein-Uhlenbeck process
- frequency as a free parameter
- include figure 3

2.1.2 Turbulent background

- Matérn model
- comparison to other integer order processes (e.g. fractional brownian motion)

2.1.3 Aggregate model

State that you can add two component models together

2.2 Model fitting

2.2.1 Whittle likelihood

- explanation of original Whittle likelihood and its problems (aliasing, leakage)
- description of tapering 'solution' to Whittle and discussion of its imperfections
- blurred whittle likelihood
- allows for uncertainty estimates via asymptotics (Fisher information)

2.2.2 Model misspecification

- semi-parametric approach in both time and frequency

2.2.3 Time-varying parameters for non-stationarity

2.2.4 Model selection/likelihood ratio tests

3 Results

3.1 Simulated results

Include Figure 5

3.2 Real drifter data (with time-varying parameters)

Include Figures 6-10

4 Discussion

- powerful technique overall
- more work needed on selecting windows

5 Appendix

5.1 Errata

- Typo in equation 13

5.2 Optimization technique

- My approach transforming parameters to an unconstrained space gives slightly better (higher maximum likelihood) estimates than their use of Matlab's built-in box constraint approach

References

Donald B Percival and Andrew T Walden. *Spectral analysis for physical applications.* Cambridge University Press, 1993.

Adam M Sykulski, Sofia C Olhede, Jonathan M Lilly, and Eric 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.