UC San Diego
UC San Diego Electronic Theses and Dissertations

Title

Bayesian Harmonic Analysis of Tidal and Wind-Driven Currents in the California Current System

Permalink

https://escholarship.org/uc/item/3q94p7xw

Author

Kachelein, Luke

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Bayesian Harmonic Analysis of Tidal and Wind-Driven Currents in the California Current
System

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Oceanography

by

Luke Kachelein

Committee in charge:

        Sarah T. Gille, Co-Chair
        Matthew R. Mazloff, Co-Chair
        Matthew H. Alford
        William A. Coles
        Bruce D. Cornuelle

2023

The Dissertation of Luke Kachelein is approved, and it is acceptable in quality

and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

In memory of my 13-inch MacBook Pro (2017-2022), whose integral service to my doctoral studies was cut short in a tragic bicycle-related drowning accident.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Above all, I would like to acknowledge my three brilliant advisers, Sarah Gille, Matt Mazloff, and Bruce Cornuelle, whose individual and collective mentorship were instrumental to my training as a scientist and professional. All three are kind, humble, and humorous people who were encouraging at all times, providing the perfect balance of guidance and independence. While amazing in their own rights, the sum of the whole is even greater than its parts, as these three are the paragon of collegiality all should strive to maintain in- and outside of work.

Additionally, I would like to thank my committee members, Matthew Alford and Bill Coles, who have graciously provided comments and feedback over the years that have strengthened my work, the former by his expertise on internal tides, and the latter by his expertise on signal processing and analysis of noisy data.

I would also like to thank Myrl Hendershott, whose one-on-one talks with me since my first year have helped me gain insight into the complexities of oceanic tides, while also being the only person at Scripps whose passion for Baroque keyboard instruments matches or exceeds my own.

The contributions of those beyond the institution have also been key. Florent Lyard of LEGOS provided guidance for the use of the FES2014 barotropic tide model in person at the 2018 SWOT Science Team Meeting. Ed Zaron of Oregon State University has lent his expertise on non-phase-locked tides on many occasions, including guidance on nomenclature in this relatively new topic. Jinbo Wang of JPL has been critical for the postlaunch planning of the SWOT satellite, the existence of which funded much of my time at Scripps, and he also encouraged me to apply myself in ways that I hope to continue going forward.

Finally, I would like to thank those in my personal life for their support, which has been equally as critical to my work as the aforementioned professional guidance. My family has been lending their unwavering encouragement and support since day 0. Near-nightly Skype calls with my parents, Mark and Sabra, and brother and fellow San Diegan Strider have kept me connected in ways previous generations could only imagine. Additionally, my friends within the Scripps

Chapter One, in full, is a reprint of the material as it appears in the Journal of Atmospheric and Oceanic Technology, 2022. Kachelein, L., B. D. Cornuelle, S. T. Gille, and M. R. Mazloff (2022) Harmonic Analysis of Non-Phase-Locked Tides with Red Noise Using the red_tide Package. *Journal of Atmospheric and Oceanic Technology*, 39(7):1031 − 1051, https://doi.org/10.1175/JTECH-D-21-0034.1. The dissertation author was the primary investigator and author of this paper. © American Meteorological Society. Used with permission.

Chapter Two, in part, is currently being prepared for submission for publication of the material in the Journal of Geophysical Research: Oceans. Kachelein, L., S. T. Gille, M. R. Mazloff, and B. D. Cornuelle. Characterizing Non-phase-locked Tidal Currents in the California Current System using High-frequency Radar (In Preparation). The dissertation author was the primary investigator and author of this material.

Chapter Three, in part, is currently being prepared for submission for publication of the material. Kachelein, L., S. T. Gille, M. R. Mazloff, and B. D. Cornuelle. The Diurnal Cycle in the California Current System: Currents from High-Frequency Radar and Winds from ERA5 (In Preparation, Preliminary Title). The dissertation author was the primary investigator and author of this material.

VITA

2015   Bachelor of Arts in Physics, Vassar College

2017   Master of Science in Oceanography, University of California San Diego

2023   Doctor of Philosophy in Oceanography, University of California San Diego

PUBLICATIONS

Luke Kachelein, Bruce D. Cornuelle, Sarah T. Gille, and Matthew R. Mazloff, "Harmonic analysis of non-phase-locked tides with red noise using the red_tide package" Journal of Atmospheric and Oceanic Technology, 39(7):1031-1051, 2022, https://doi.org/10.1175/JTECH-D-21-0034.1 © American Meteorological Society. Used with permission.

ABSTRACT OF THE DISSERTATION

Bayesian Harmonic Analysis of Tidal and Wind-Driven Currents in the California Current
System

by

Luke Kachelein

Doctor of Philosophy in Oceanography

University of California San Diego, 2023

Sarah T. Gille, Co-Chair
Matthew R. Mazloff, Co-Chair

The surface layer of the ocean is a critical component of the ocean-atmosphere system.
The momentum and heat transfer that occur here, as well as tidal forcing, result in energetic
currents across a wide range of length and time scales. Because the surface is observable from
land and space, it is the subject of extensive study from land-based and satellite observation
systems. Only recently, however, have spatial resolutions from satellite instruments approached
the submesoscale, which evolves at time scales comparable to or faster than repeat orbit sampling
rates. At the same time, land-based radar networks have matured such that more than a decade of
surface current observations at kilometer scale resolution are available from many regions around

the world, unconstrained by the timing of orbits that affect satellites but are limited to coastlines. This leads to the central motivating question that this dissertation seeks to address: what can coastal radar data reveal about short time scale processes that could be useful to the goals of recent and upcoming Earth observation satellite missions? This work is limited to examining the surface expression of internal tides and diurnal wind-driven currents, as both are energetic processes at known frequencies, lending themselves well to harmonic analysis.

In Chapter 1, a novel tidal harmonic analysis method is presented, following from previous tidal analysis and incorporating Bayesian principles and prior statistics to inform the least squares fitting procedure. This method is applied in Chapter 2 to high-frequency radar data from the California Current System in order to characterize tidal currents in that region, finding a phase structure indicative of internal tidal propagation and a complicated spatial distribution of the fraction of tidal energy deemed to result from modulation by other processes in the ocean. Finally, Chapter 3 uses the same method on currents and winds in order to characterize the diurnal wind-driven current, finding the times of day that these quantities peak, their level of coherence, and the angle between them as a function of frequency. These chapters provide a detailed view of these processes in a well-studied coastal region, with implications for satellite oceanography.

Introduction

The ocean is a complex system coupled to the atmosphere, cryosphere, solid Earth, and even the Sun and Moon via mechanical forces, friction, radiation, and gravitation. The energy budget within the ocean and between it and the other systems to which it is coupled is the subject of much investigation (W. H. Munk & Wunsch, 1998; Wunsch, 1998; Egbert & Ray, 2001; Watanabe & Hibiya, 2002; Alford, 2003; Scott & Xu, 2009; Trenberth et al., 2009; Yu et al., 2018) and is perhaps the core problem of oceanography. Fluid interactions within the ocean occur at scales spanning many orders of magnitude, from millimeter to 10,000 km scales (Ferrari & Wunsch, 2009). The range of time scales at which these processes occur is similarly large, from seconds to centuries, with longer time scales generally corresponding to larger length scales and vice versa, and a cascade of energy transferring from both high to low and low to high length and time scales (Dickey & Bidigare, 2005). A notable outlier to this scale relationship is the barotropic tide, the length scale of which is large (hundreds of kilometers) but the time scale of which is relatively short (12 to 24 hours for major constituents). Similarly, the diurnal winds, which force the ocean at the surface, are known to have spatial scales up to hundreds of kilometers (Gille et al., 2003) and by definition a time scale of 24 hours, while diurnal radiative forcing occurs at the same frequency on a planetary scale. These high-frequency, large time scale processes contribute significantly to the ocean's energy budget: for example, tidal energy dissipation is estimated to be 2.5 TW for the principal lunar semidiurnal (M_2) tide, 3.2 TW for all lunar tides, and 3.7 TW for all solar tides, of which a large majority is dissipated in the ocean (Egbert & Ray, 2000, 2001) after cascading from barotropic tides to internal tides to turbulence (Garrett & Kunze, 2007).

Different observation systems are suited to different time and length scales, as indicated by Dickey & Bidigare (2005). A high sampling rate and geographically wide coverage are therefore necessary for the full resolution of high-frequency, large time scale processes. Satellite-mounted instruments have been revolutionary for oceanography in recent decades, as their domain is planetary in scale. The major constraints facing them are therefore the spatial resolution of the instruments and the temporal sampling rate. Latest generation altimeters have begun to improve on the first constraint, with the recently-launched Surface Water and Ocean Topography (SWOT) mission aiming to resolve sea surface height at 15-30 km wavelengths (Morrow et al., 2019). However, the sampling rate constraint is an immutable reality of Earth-orbiting satellites, which renders impossible the direct resolution of high-frequency processes (here 1 cycle per day or more). The most predictable components of the tides alias to known frequencies, and can be removed for interpreting altimeter data (Parke et al., 1987; R. Ray, 1993). However, in addition to the regular, so-called "phase-locked" component of the tide there is a non-phase-locked component that can contaminate mesoscale and submesoscale signals and is difficult to remove from altimeter data, therefore posing a challenge to high-spatial-resolution satellite missions (Chavanne & Klein, 2010; Savage, Arbic, Alford, et al., 2017; Zaron & Ray, 2018). Characterizing this component observationally can help validate numerical studies and provide critical information needed to assess aliased tidal variability in newer high-resolution satellite missions. Similarly, diurnal wind-driven signals in upper ocean currents are aliased by scatterometry, the orbits of which are usually sun-synchronous (Rodríguez et al., 2019), and therefore similar considerations should be taken. Here, we consider tools that can be used to investigate high-frequency variability, including methods and presently-available data sets, and perform regional investigations into the high-frequency signals that need to be carefully considered for future satellite missions as outlined above.

This dissertation is organized into three chapters. Chapter 1 develops a novel tidal harmonic analysis package designed to extract signals with correlated noise, which we expect to be important to the analysis of tides with a non-phase-locked component. In the tradition of

previous tidal harmonic analysis packages like t_tide (Pawlowicz et al., 2002), UTide (Codiga, 2011), and ns_tide (Matte et al., 2013), this package is called "red_tide" after the spectrally "red" noise it considers, and so that it may memorably share its name with the unrelated toxic algal bloom (e.g. Fleming et al., 2011). Evaluating synthetic and real data, we recover tidal information by expanding conventional harmonic analysis to include prior information and assumptions about the statistics of the background noise. This is implemented using Bayesian maximum posterior estimation and assuming Gaussian prior distributions. When the background energy spectrum is nearly spectrally white, red_tide results replicate those from ordinary least squares (OLS) commonly used in other tidal packages, while with red background spectra (a spectral slope of -2 or steeper), our results represent a measurable improvement over OLS.

Chapter 2 is an investigation of the tidal surface currents off the United States West Coast as observed by high-frequency radar (HFR) and comparing it to the FES2014 barotropic tidal model. This data set, analyzed using red_tide, is able to resolve surface currents at the kilometer scale and provides years of hourly data, which fully captures all major tidal constituents. From this we characterize the amplitude and phase of tidal currents in this region, with emphasis on the prominent M_2 constituent, finding regional variability in tidal amplitudes and a phase structure suggestive of baroclinic propagation. We find that the spatial scale of the fraction of non-phase-locked variance is comparable in size to SWOT altimeter swaths, and we also find that the low-frequency amplitude modulation of the M_2 constituent is correlated in many places with that of the S_2 constituent, indicating a physical mechanism of modulation that can be characterized from harmonically decomposed HFR data.

Chapter 3 investigates the diurnal component of the same surface current data as well as that of the wind, which is given by the 10-meter winds of the ERA5 reanalysis product. We once again use red_tide but now on 6-month segments of winds and currents in order to evaluate the seasonal coherence of the two quantities. We find high coherence at the diurnal frequency, especially between the dominant clockwise-polarized components. The deflection angle between the wind vector and current vector is typically greater than that expected by Ekman-type models of

wind-driven current, suggesting deviations from those models and the involvement of additional physical processes, as has been observed previously. We also estimate the times at which currents and winds are at their local maximum during the diurnal cycle, providing insight into potential optimization for designing the orbit of a future scatterometer mission.

The unifying goal of these chapters is to provide a toolbox and apply it to relevant data, with two broad, related motivations. Firstly, we wish to gain insight into processes that will be relevant for the gathering and interpretation of data from present and future satellite missions. Secondly, for general oceanographic understanding we wish to gain greater insight into the surface processes in the California Current System, a highly-studied region with years of HFR data comprising some of the most extensive coverage worldwide using this technology. Chapter 1 presents the analysis method with several test cases demonstrating its utility. Chapter 2 applies this method to over 9 years of data in order to separate long-term phase-locked tides from their more unpredictably modulated components. Chapter 3 applies the method to winds and currents together on sub-annual time scales to evaluate their interaction. Together, they provide high spatial resolution information about the energetic, high-frequency variability in the California Current System.

Chapter 1

Harmonic Analysis of Non-Phase-Locked Tides with Red Noise using the red_tide Package

Abstract

A novel tidal analysis package (red_tide) has been developed to characterize low-amplitude non-phase-locked tidal energy and dominant tidal peaks in noisy, irregularly sampled, or gap-prone time series. We recover tidal information by expanding conventional harmonic analysis to include prior information and assumptions about the statistics of a process, such as the assumption of a spectrally colored background, treated as non-tidal noise. This is implemented using Bayesian maximum posterior estimation and assuming Gaussian prior distributions. We utilize a hierarchy of test cases, including synthetic data and observations, to evaluate this method and its relevance to analysis of data with a tidal component and an energetic non-tidal background. Analysis of synthetic test cases shows that the methodology provides robust tidal estimates. When the background energy spectrum is nearly spectrally white, red_tide results replicate results from ordinary least squares (OLS) commonly used in other tidal packages. When background spectra are red (a spectral slope of -2 or steeper), red_tide's estimates represent a measurable improvement over OLS. The approach highlights the presence of tidal variability and low-amplitude constituents in observations by allowing arbitrarily configurable fitted frequencies and prior statistics that constrain solutions. These techniques have been implemented in MATLAB in order to analyze tidal data with non-phase-locked components and an energetic background that pose challenges to the commonly used OLS approach.

6

1.1   Introduction

Tides are a major driver of oceanic variability. They are forced by the gravitational effects of the moon and sun and have a ubiquitous presence throughout the global ocean. The tide generating potential was described harmonically by Sir George Howard Darwin in 1883 (Doodson & Lamb, 1921) and was further developed by Doodson, with the Darwin symbols for tidal constituents (e.g. O_1 and M_2) still in widespread use. The ubiquity of tides in oceanic data has motivated the development of techniques for determining tidal parameters. Two types of methods are frequently used to analyze tides: discrete Fourier transform-based methods and least squares-based harmonic analysis. Discrete Fourier transform-based methods use the energy contained in discrete frequency bands to diagnose the amplitude of tidal constituents. Least squares harmonic analysis has been used for decades (W. Munk & Hasselman, 1964; Zetler et al., 1965) to estimate the amplitude and phase of tidal signals at known tidal frequencies.

Tidal signals can be separated into two components: the relatively predictable barotropic tide and the more variable baroclinic tide (R. D. Ray & Mitchum, 1996). The predictability of the barotropic tide is a consequence of its stable phase and amplitude due to its large scale, rapid propagation, and the regularity of the astronomical forcing. Classical harmonic analysis at tidal constituent frequencies is effective for analyzing time series of sea surface height or bottom pressure that are dominated by the barotropic tide and are characterized by sharp, narrow peaks in the frequency domain at tidal constituent frequencies. The interannual lunar nodal cycle (18.61 years) and lunar perigee (8.85 years), which are not directly resolvable in typically short tidal records, cause tidal modulations that affect the interpretation of tidal records (Haigh et al., 2011). The t_tide package, for example, accounts for these cycles using nodal corrections (Pawlowicz et al., 2002). However, other processes can modulate the tidal peaks.

In the case of the baroclinic tide, which is more variable in amplitude and phase, propagation through varying stratification or non-linear interaction with other waves leads to amplitude and/or phase modulation via the transfer of energy to the internal wave continuum

spectrum and a loss of coherence with the astronomical forcing (Chiswell, 2002; Rainville & Pinkel, 2006). Interaction with eddy fields, background currents, and the seasonal cycle in stratification all cause internal tides to vary in time (R. D. Ray & Zaron, 2011). This variability spreads the tidal energy across a band of frequencies centered at the tidal forcing frequency, forming a tidal cusp (W. H. Munk et al., 1965); the spreading of energy in the frequency domain can pose challenges to describing the predictable tidal component. The component of tidal energy resulting from interaction with other processes in the ocean appears has been referred to by different names in the literature, including the incoherent tide (e.g. Eich et al., 2004), the non-stationary tide (e.g. R. D. Ray & Zaron, 2011), and more recently the non-phase-locked tide (Zaron, 2019). We have chosen the term "non-phase-locked" because it is associated purely with tides and their generating potential, whereas the other terms overlap with wave and statistics terminology. Nevertheless, these other terms are reasonable, as this component of the tides is incoherent with astronomical forcing and is non-stationary in time.

Throughout this paper, *model* refers to the series of basis functions (harmonics) at frequencies chosen to approximate (or *to model*) a given time series, while the coefficients of these harmonics are referred to as model parameters. Solving for these parameters to find the best estimate of the underlying tidal component of the observations is the goal of least squares tidal harmonic analysis. The choice of basis functions is central to the technique, and the designation of these functions as a model for observations follows standard least squares terminology (e.g. Wunsch, 1996). Conventional least squares tidal harmonic analysis models a time series as a sum of sinusoids at tidal frequencies, with amplitudes and phases optimized to best fit observations. This contrasts with analysis via the discrete Fourier transform in several important ways: Fourier analysis requires evenly sampled time series and decomposes a signal into components at evenly spaced frequencies determined by the record length and sample rate. The Fourier transform is periodic at the record length ("fundamental") and is band-limited at the Nyquist frequency. Additionally, the Fourier transform does not allow for a separate component of noise. Harmonic techniques for analyzing tidal time series overcome these limitations of Fourier analysis. They

allow for estimates that do not exactly match all observations and that provide a unique solution in the presence of noise and potential nonorthogonality between basis functions, for example when the frequencies of basis sinusoids differ by less than the fundamental frequency. This is solved as an inverse problem in which the fit is expected to differ from the observations by some residual. Data may be irregularly spaced and fit to arbitrary basis functions, including sinusoids of any frequency, unconstrained by periodicity over the record length and not band-limited by the sampling rate. Additionally, harmonic analysis allows for a noise component with prior statistics, generally mean and autocovariance. We denote the autocovariance and its matrix representation with the standard terms "covariance" and "covariance matrix", respectively, as we do not discuss any cross-covariance quantities for which such shorthand might be confusing.

At frequencies outside the tidal bands, ocean data tend to be spectrally red, with greater power at lower frequencies (W. H. Munk et al., 1965). When finite duration records with steep spectra are analyzed, high-amplitude low-frequency processes with periods that do not match the series length can alter estimated spectral power at higher frequencies. Spectral leakage is problematic for characterizing internal tides, especially in regions where mesoscale variability is much stronger than the internal tide (R. D. Ray & Zaron, 2016). Low-frequency variability, tidal cusps, and high-frequency noise must all be accounted for either explicitly or as a residual term. In other words, total signal variability is modeled as sinusoids at given frequencies added to a residual broadband background that has power at all frequencies, which is characterized by a residual (or noise) covariance matrix. Our approach is to choose basis functions and prior statistical assumptions about the signal and noise components (quantified in covariance matrices) that match the expected variability in the observations as well as computational cost allows; such constraints bias the model parameters by reducing their variance, which is appropriate when tidal constituents and related components are estimated from limited sampled data with noise. We seek to avoid overfitting and we argue that this produces better results than estimators obtained from methods that use less prior knowledge.

Least squares tidal harmonic analysis has drawbacks. Pawlowicz et al. (2002) identify

some challenges including record length requirements for distinguishing some tidal frequencies, the lack of distinction between true tidal lines and background energy at tidal frequencies, and the broadening of spectral lines from estuarine tidal responses and stratification-dependent internal tides. Nevertheless, harmonic analysis has been widely adopted for tides because it is well-suited for signals with a weak noise component relative to the tidal signal. Pawlowicz et al. (2002), expanding upon earlier code for tidal harmonic analysis (Foreman, 1977; Foreman & Henry, 1989; Foreman et al., 2009) and employing MATLAB, created the widely-used t_tide package, incorporating methods to mitigate known drawbacks to classical harmonic analysis. These methods include nodal corrections and inference of unresolvable constituents to account for the long record lengths required for resolution under classical harmonic analysis, as well as three algorithms to provide confidence intervals to account for non-tidal energy at tidal frequencies. Other authors have expanded the t_tide procedure (Leffler & Jay, 2009; Codiga, 2011) or have modified it for specific dynamical regimes, such as tides in the presence of river outflow (Matte et al., 2013).

The two-fold problem of accurately estimating tidal variability with a component that is not phase-locked to astronomical forcing in the presence of spectrally-colored noise while minimizing spectral leakage has motivated us to develop a new tidal harmonic analysis package, red_tide, that accounts for tidal cusps and red background spectra. The appendix provides information on the access of the package, which may be modified to accommodate individual needs. We do not presently incorporate nodal corrections like those used in t_tide in order to emphasize features specific to our method, though these corrections may be incorporated in a future version. The primary scientific motivation for developing the package is to support a detailed level of tidal analysis of highly variable baroclinic tides. For example, such tides are expected to be an important part of the signal at spatial scales of $O(10)$ kilometers to be measured by the upcoming Surface Water and Ocean Topography (SWOT) swath altimeter, for which the interaction between internal tides and ocean mesoscale variability requires the most accurate predictions possible (Chavanne & Klein, 2010; R. D. Ray & Zaron, 2011). Analysis

related to such an application is likely to focus on in-situ data and ocean model output sampled at smaller time intervals than semidiurnal tidal periods. Therefore relevant test cases are presented here. Though harmonic analysis of tidal signals aliased in satellite data has been performed on the phase-locked tide (R. D. Ray & Mitchum, 1996; Zhao et al., 2011), this is not examined here. This method is compared with existing methods (ordinary least squares and t_tide, which implements ordinary least squares) using synthetic and observational time series to examine performance and practical limitations. The t_tide package requires fewer inputs and produces matching results for lower computational cost when analyzing time series that do not exhibit the complications that the red_tide method seeks to address. Therefore, our method does not supplant existing tidal analysis packages except in specific regimes where it is advantageous. Other tidal packages, including those that are built upon t_tide (e.g. Leffler & Jay, 2009; Codiga, 2011; Matte et al., 2013), are not examined in this study.

The rest of this study is divided into three sections. We begin by outlining the linear algebraic and statistical methods that underpin our tidal analysis in Section 1.2. We then apply these methods to synthetic time series, first by highlighting specific features of the method, including application to a step function to demonstrate aliased signals (Section 1.31.3.1), and then by analyzing tide-like synthetic series to show the effect of prior statistical assumptions on model parameters (Sections 1.31.3.2-1.3.4). Two examples of observational data follow in Section 1.4, with comparisons to t_tide. Finally, summary and discussion follow in Section 1.5.

1.2   Methods

The basic framework of weighted least squares estimation used in red_tide is outlined in a number of references (e.g. Wunsch, 1996; Menke, 2018). Here we provide a review of this framework, formulated for tidal harmonic analysis of records with arbitrarily structured noise, with notation following that of Ide et al. (1997) with a few modifications. Note that tidally-driven components of a time series are considered "signal" while non-tidal processes are referred to

11

as "noise" for the purpose of distinguishing them in the context of harmonic analysis. These so-called noise terms may include instrument error as well as non-tidal processes in the ocean, such as submesoscale eddies and the internal wave continuum.

A zero-mean, discretely sampled ocean time series written as a column vector $\mathbf{y}$ of length N is modeled as the sum of sinusoids of tidal and non-tidal frequencies

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{r}, \tag{1.1}$$

where $\mathbf{H}$ is an N by $2M$ regressor matrix (i.e. the model basis functions), $\mathbf{x}$ are the $2M$ model parameters, and $\mathbf{r}$ represents the N-element residual time series. The columns of $\mathbf{H}$ are sines and cosines of prescribed frequencies, ω_m, for $m = 1, 2, \dots M$, such that equation (1.1) can be expressed as

$$\begin{aligned}
\mathbf{y} &= \sum_{m=1}^{M} \left(a_m \sin(\omega_m \mathbf{t}) + b_m \cos(\omega_m \mathbf{t}) \right) + \mathbf{r}, \\
\mathbf{x} &= [a_1, b_1, a_2, b_2, \dots, a_M, b_M]^{\mathrm{T}} \\
\mathbf{t} &= [t_1, t_2, \dots, t_N]^{\mathrm{T}}
\end{aligned} \tag{1.2}$$

The unknown model parameters $\mathbf{x}$ are estimated by $\hat{\mathbf{x}}$ such that the trace of $\langle (\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^{\mathrm{T}} \rangle$, the expected value of the sum of the squares of the model parameter errors, is minimized. This is done by calculating $\hat{\mathbf{x}}$ as the Bayesian maximum a posteriori (MAP) estimate, by requiring that $\mathbf{x}$ and $\mathbf{r}$ have independent Gaussian distributions which satisfy the conditions under which the MAP estimate gives $\hat{\mathbf{x}}$ (Van Trees, 2001). By Bayes' theorem, the posterior probability distribution of $\mathbf{x}$ given observations $\mathbf{y}$ is proportional to the product of the prior probability distribution of $\mathbf{x}$ and the likelihood of $\mathbf{y}$ given $\mathbf{x}$:

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{x})P(\mathbf{y}|\mathbf{x})}{P(\mathbf{y})}. \tag{1.3}$$

Given the assumptions of Gaussian statistics and a linear model:

$$P(\mathbf{x}|\mathbf{y}) \propto \exp\left(-\mathbf{x}^T\mathbf{P}^{-1}\mathbf{x}\right)\exp\left(-(\mathbf{y}-\mathbf{Hx})^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{Hx})\right). \tag{1.4}$$

The denominator $P(\mathbf{y})$ is not a function of $\mathbf{x}$ and therefore can be omitted, since it is not relevant to the optimization. The matrices $\mathbf{R} = \langle\mathbf{rr}^T\rangle$ (size N by N) and $\mathbf{P} = \langle\mathbf{xx}^T\rangle$ (size $2M$ by $2M$) are the covariance matrices of $\mathbf{r}$ and $\mathbf{x}$ respectively ($\mathbf{P}$ is thus a hyperparameter of the Gaussian prior for $\mathbf{x}$). This expression as well as its logarithm,

$$\ln\left(P(\mathbf{x}|\mathbf{y})\right) = -\mathbf{x}^T\mathbf{P}^{-1}\mathbf{x} - (\mathbf{y}-\mathbf{Hx})^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{Hx}) + \text{constant}, \tag{1.5}$$

are at a maximum for some value $\mathbf{x} = \hat{\mathbf{x}}$. Since the posterior probability $P(\mathbf{x}|\mathbf{y})$ follows a Gaussian distribution, its mean (the Bayes estimator) equals its mode (the MAP estimator) and can therefore be solved as a maximization problem (Van Trees, 2001). At the mode, the partial derivative of equation (1.5) with respect to $\mathbf{x}$ vanishes:

$$0 = \frac{\partial}{\partial\mathbf{x}}\ln\left(P(\mathbf{x}|\mathbf{y})\right)\bigg|_{\mathbf{x}=\hat{\mathbf{x}}} = -2\mathbf{P}^{-1}\hat{\mathbf{x}} + 2\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} - 2\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\hat{\mathbf{x}}. \tag{1.6}$$

The vector $\hat{\mathbf{x}}$ contains the most probable model parameters given observations $\mathbf{y}$, and solving for $\hat{\mathbf{x}}$ gives the most probable solution to equation (1.1):

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{P}^{-1}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}. \tag{1.7}$$

An equivalent expression via the matrix inversion lemma (e.g. Wunsch, 1996) is not used here but appears in the appendix.

The posterior covariance matrix of the difference between the estimated and true model

parameters is

$$\langle (\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^{\mathrm{T}} \rangle = \left(\mathbf{H}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{H} + \mathbf{P}^{-1} \right)^{-1}. \tag{1.8}$$

In red_tide, model parameters $\mathbf{x}$ are assumed to have a Gaussian probability distribution function (PDF) resulting from the Gaussian distributions of the prior and likelihood function. Therefore, the posterior PDF of x_m, the m-th element of $\mathbf{x}$, is a Gaussian with a mean given by the m-th element of $\hat{\mathbf{x}}$ from equation (1.7) and variance given by the m-th element of the diagonal of the matrix in equation (1.8).

If $\mathbf{x}$ was not expected to have a Gaussian distribution, a different expression for $\hat{\mathbf{x}}$ would need to be derived by similar Bayesian principles starting from equation (1.3), which is beyond the scope of this study. Quantities defined as nonlinear functions of Gaussian-distributed quantities, however, may have non-Gaussian distributions that can be estimated. In this paper, plotted uncertainty bounds for derived quantities are estimated using a Monte Carlo approach, based on an ensemble population of $\mathbf{x}$ with the posterior Gaussian PDF. This approach is implemented due to its flexibility and simplicity compared with analytical solutions, which are not always in closed form, or with more complicated approximations such as piecewise linear discretization of PDFs (e.g. Lourens & van Geer, 2016). The tidal amplitude is one such quantity, where $\hat{A}_m = \sqrt{\hat{a}_m^2 + \hat{b}_m^2}$ is an estimator of the true amplitude A_m, which follows a noncentral χ distribution when the standard deviations of a_m and b_m are equal ($\sigma_{a_m} = \sigma_{b_m} = \sigma_m$). Similarly, the tidal phase $\hat{\phi}_m$ is an estimator for the true phase shift ϕ_m, defined by $\tan(\phi_m) = -b_m/a_m$, whose statistics are also not Gaussian. Whenever uncertainty bounds are given for quantities derived from the Bayesian methods outlined above or for quantities that are functions of them, we refer to them as *credible intervals* in accordance with Bayesian terminology for the interval in which an estimated parameter lies with the stated probability (Lee, 1997). These are analogous to *confidence intervals*, the uncertainty bounds on quantities derived from frequentist methods. Because Pawlowicz et al. (2002) uses the term *confidence intervals*, and because t_tide is not

derived using a Bayesian framework, we refer to t_tide output and other non-Bayesian quantities as having confidence intervals when comparing it to red_tide output and its credible intervals.

Because $\mathbf{y}$ is modeled as the sum of sinusoids and a noise component, its expected power spectrum $S_{\mathbf{yy}}(f)$ may be interpolated to fitted frequencies and used to construct $\mathbf{P}$. This approach, however, results in doubly counting energy, as $\mathbf{r}$ contributes to the variance of $\mathbf{y}$ at all frequencies, including frequencies modeled by $\mathbf{Hx}$. Because the energy of a tidal peak and cusp are typically much higher than the background noise and may in fact be underestimated in $S_{\mathbf{yy}}(f)$ due to peak-broadening from spectral averaging, the double counting of energy will be small around most prominent tidal peaks. For cases where tidal signals are comparable in energy to $\mathbf{r}$ at tidal frequencies, $\mathbf{P}$ can be reduced by the appropriate amount. In sections 1.31.3.3 and 1.31.3.4, the correct partition of energy into $\mathbf{P}$ versus $\mathbf{R}$ is possible because we have perfect knowledge of the underlying synthetic processes. For real data for which we lack perfect knowledge, such as those in section 1.4, the approach we use for convenience is to assume wide-sense stationary noise so that we can obtain the residual spectrum $S_{\mathbf{rr}}(f)$ from the Fourier transform of any column of $\mathbf{R}$. We then subtract it from $S_{\mathbf{yy}}(f)$ to construct $\mathbf{P}$ in order to obtain a more accurate partition of signal versus noise energy. We also assume throughout that $\mathbf{y}$ is a wide-sense stationary time series; this means that the elements of $\mathbf{x}$ are assumed to be uncorrelated and therefore that $\mathbf{P}$ is diagonal throughout (see Bendat & Piersol (2010) for nonstationary data analysis and double frequency spectra, the continuous analogue to a non-diagonal $\mathbf{P}$).

Values along the main diagonal of $\mathbf{R}$ represent the expected variance of the misfits between fitted time series and observations. Off-diagonal elements indicate the covariance at lagged times, with values farther off the main diagonal corresponding to larger time lags. Beyond some time lag, the covariance may be approximated as zero if long-period energy is sufficiently small or explicitly represented in $\mathbf{H}$. This approximation is useful, as it limits the memory requirement for large $\mathbf{R}$. A diagonal $\mathbf{R}$ is a special case of this, with non-zero elements only along its main diagonal. This corresponds to an assumption of zero lagged noise correlation or equivalently an assumption of spectrally white noise. This approximation is often made for

15

computational efficiency, as $\mathbf{R} = \sigma^2 \mathbf{I}$ can be replaced in equation (1.7) with the constant σ^2, which is estimated as $\sigma^2 = \langle \mathbf{r}^T \mathbf{r} \rangle / N$. In cases where the residual $\mathbf{r}$ may be better approximated as non-white noise, other procedures can be used to construct a non-diagonal $\mathbf{R}$, including those described in section 1.21.2.2 and the appendix. In the examples that follow, we assume that all residuals have the same variance. This assumption of stationarity corresponds to constant elements along the diagonals of $\mathbf{R}$, i.e. a Toeplitz matrix.

Time series with a nonstationary residual $\mathbf{r}$, for example due to time-varying instrumental noise, may have their residual covariance approximated by a non-Toeplitz matrix $\mathbf{R}$ in order to reduce the impact of the affected segments on the calculation of $\hat{\mathbf{x}}$. Additionally, observational gaps have a non-trivial effect when treating background noise as correlated. The matrix $\mathbf{R}$ is also not Toeplitz in this case, though for the direct inversion of relatively small $\mathbf{R}$, this does not seem to impact performance. Computation may be reduced when analyzing multiple time series if each time series $\mathbf{y}_n$ can be analyzed with the same $\mathbf{H}$, $\mathbf{R}$, and $\mathbf{P}$ such that all terms in equation (1.7) before $\mathbf{y}$ are evaluated once and multiplied by each $\mathbf{y}_n$.

1.2.1 Relation to Least Squares

Ordinary least squares (OLS) has the same form as weighted least squares if $\mathbf{R} = \sigma^2 \mathbf{I}$ and $\mathbf{P}^{-1} \rightarrow 0$ (Wunsch, 1996), and is used in the t_tide package (Pawlowicz et al., 2002). OLS seeks to minimize $(\mathbf{y} - \mathbf{Hx})^T (\mathbf{y} - \mathbf{Hx})$, the misfit between the fitted time series and observations. In this case, the solution is

$$\hat{\mathbf{x}}_{\text{OLS}} = \left(\mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{y}. \tag{1.9}$$

Observation error is assumed to be white noise, and OLS places no constraint on the magnitude of the components of $\hat{\mathbf{x}}_{\text{OLS}}$. Steep, non-constant background spectra would need to be modeled by including more parameters within this framework, which may be impractical. Without using a probabilistic framework, in order to constrain the magnitude of the model parameters $\hat{\mathbf{x}}$ to avoid

overfitting, the quantity to be minimized is often written as $(\mathbf{y} - \mathbf{Hx})^{\mathrm{T}}(\mathbf{y} - \mathbf{Hx}) + \mathbf{x}^{\mathrm{T}}\mathbf{W_x}\mathbf{x}$, where $\mathbf{W_x}$ is a matrix that weights the relative importance of minimizing model parameter magnitude over misfit. This technique is called ridge regression and reduces overfitting at the expense of bias (Wunsch, 1996). The resulting least squares equation is said to be *regularized*:

$$\hat{\mathbf{x}}_{\mathrm{TLS}} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H} + \mathbf{W_x}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{y}. \tag{1.10}$$

Similarly, the misfit $\mathbf{r}$ may be weighted by the matrix $\mathbf{W_r}$ such that $(\mathbf{y} - \mathbf{Hx})^{\mathrm{T}}\mathbf{W_r}(\mathbf{y} - \mathbf{Hx}) + \mathbf{x}^{\mathrm{T}}\mathbf{W_x}\mathbf{x}$ is minimized, with the corresponding *weighted* regularized least squares equation

$$\hat{\mathbf{x}}_{\mathrm{WTLS}} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{W_r}\mathbf{H} + \mathbf{W_x}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{W_r}\mathbf{y}. \tag{1.11}$$

This is equivalent to equation (1.7) when $\mathbf{W_x} = \mathbf{P}^{-1}$ and $\mathbf{W_r} = \mathbf{R}^{-1}$, though the derivation of that solution is distinct from these least squares approaches (see section 3.6.2 of Wunsch, 1996, for details). A weighted least squares estimate (not regularized) is used in the UTide package (Codiga, 2011), which implements an iteratively reweighted robust fit corresponding to equation (1.11) with $\mathbf{W_x} = \mathbf{0}$ and $\mathbf{W_r}$ as a diagonal weighting matrix that de-emphasizes outliers and is updated iteratively, the details of which are beyond the scope of this study.

1.2.2 Power Law Noise

Away from energetic processes at tidal, seasonal, and inertial frequencies, spectra S of ocean time series tend to follow a power law of the form $S \propto f^{\gamma}$, where $\gamma \leq 0$ is the spectral slope and the negative value indicates more energy at lower frequencies (Agnew, 1992). For computational efficiency in problems with many observations, we can assume wide-sense stationary noise and construct the residual covariance matrix $\mathbf{R}$ as a sparse, symmetric Toeplitz matrix with the diagonals calculated from the Fourier transform of $S \propto f^{\gamma}$ per the Wiener-Khinchin theorem, truncated at a user-defined time lag. Users of red_tide have the option of multiplying

the covariance by a window function in order to reduce the spectral ringing that results from an abrupt drop to zero in the frequency domain. A low-amplitude white spectrum (spectral slope of 0) may also be added to account for observational error, such that the spectrum of **r** has a noise floor at all frequencies. This residual spectrum is not altered within fitted tidal frequency bands due to the impossibility of distinguishing tidal and non-tidal energy at the same frequency, which results in model parameter uncertainty. For the typical case of tides that are much more energetic than the background, this effect is small, while in cases where tidal constituents have low energy or a broadband cusp of interest, the relatively larger uncertainty estimates on model parameters reflect the fact that non-tidal variance is comparable to tidal variance. The approximation of the residual time series following some modified spectral power law with $\gamma < 0$ will hereafter be referred to as a *red noise* assumption, even when γ is not exactly -2. In addition to the special cases considered above, the noise can exhibit more complicated structure that allows **R** to be tractable while still representing some forms of red noise, for example noise as an autoregressive process (see Appendix).

1.3   Application of Red_Tide to Illustrative Cases

In order to compare the performance of red_tide to other widely-used fitting procedures, we lay out several illustrative examples of its application to synthetic and physical data. Before more complicated cases are examined, it is worth reviewing a simplified case of tidal harmonic analysis that typifies the separate treatment of signal and noise and demonstrates the consequences of using an incomplete model. A typical ocean time series will have many tidal constituents present, with longer records better able to differentiate nearby frequencies. Because harmonic methods model the data using sinusoids of prescribed tidal frequencies, energy at unmodeled frequencies will remain in the residual time series. Figure 1.1 depicts the results of modeling a bottom pressure record (examined in greater detail in section 1.41.4.1) at only two tidal frequencies: the principal lunar semidiurnal (M_2) and the principal solar semidiurnal (S_2). Figure

1.1A, with the interval in gray expanded in panel B, illustrates the misfit between the complicated data and the simple model that arises due to substantial residual energy at unmodeled frequencies. In the frequency domain (figure 1.1C), this is evident by the energetic tidal lines at frequencies that are not included in **H**. For a data set with well-defined tidal lines with energy much greater than that of the background, conventional harmonic analysis like that employed by t_tide is well-suited to addressing this issue by simply including more tidal constituents. The red_tide package is designed for more complicated cases for which single tidal constituents are insufficient as well as cases where non-tidal variance is comparable to tidal variance.

1.3.1 Model Coefficient Covariance, the Gibbs Phenomenon, and Periodicity

One challenge to harmonic analysis stems from the accurate representation of model coefficient covariance matrix $\mathbf{P} = \langle \mathbf{x}\mathbf{x}^T \rangle$. Before further examining realistic cases, we start with a familiar example that demonstrates the effect of the covariance on computed model coefficients for a discrete step function. Modeling the step as a finite set of sinusoids leads to the Gibbs phenomenon, the tendency of the partial sum of a Fourier series to overshoot in the neighborhood of a discontinuity of the modeled function (Hewitt & Hewitt, 1979). This phenomenon persists even with the addition of more terms in the partial sum, though the magnitude is reduced. While a discretely sampled time series may be fit well at observation points, band-limited Fourier coefficients do not adequately fit the step function between sampling times. The step function and resulting Gibbs phenomenon serve as a simple but extreme example of a situation present in real data: a process with variance at unmodeled frequencies has a prior that does not adequately describe the process. The variance of a step function is distributed across all frequencies (see equation (1.12) below), but when it is sampled coarsely and reconstructed as a band-limited process, some variance is aliased and the true underlying process is poorly reconstructed despite good agreement at observation times. Therefore, we incorporate assumptions and prior knowledge of a process, including cases in which there are fewer data than parameters that are suspected to

be worth estimating.

Another challenge stemming from the limitations of finite sampling and fitting is the inherent periodicity of solutions when a finite record is modeled as the sum of periodic functions; this is not a problem for tidal processes, which have periods much shorter than those of typical observations, but it will affect estimates for explicitly modeled low-frequency processes that may also be of interest. To demonstrate the effect of the estimator on both of these related issues (Gibbs phenomenon and artificial periodicity), we analyze a finite, uniformly sampled record of a step function using the method described in Section 1.2.

Because the step function, here represented continuously by the sign (signum) function, has an analytic Fourier transform, the expected model parameter covariance matrix $\mathbf{P}_{\text{step}} = \langle \mathbf{x}_{\text{step}} \mathbf{x}_{\text{step}}^{\text{T}} \rangle$ may be constructed from the absolute value squared of the Fourier transform:

$$y_{\text{step}}(t) = \text{sgn}(t - t_0) = \begin{cases} -1, & t < t_0 \\ 0, & t = t_0 \\ 1, & t > t_0 \end{cases}$$

$$|\mathcal{F}\{\text{sgn}(t - t_0)\}(f)|^2 = \frac{1}{\pi^2 f^2}, \tag{1.12}$$

where $y(t)$ is the underlying continuous time series to be analyzed after discrete sampling, $\text{sgn}(t - t_0)$ is the sign function with jump discontinuity at t_0 and $\mathcal{F}\{\cdot\}(f)$ denotes the Fourier transform operation to the frequency domain. The square of the Fourier transform of this function is proportional to frequency^{-2}, therefore it has a spectral slope of -2.

In the example, a step function of record length T is sampled symmetrically about the jump discontinuity at 1000 evenly-spaced times such that all Fourier frequencies, from the fundamental frequency $\Delta f = 1/T$ to the Nyquist frequency f_{Ny}, could be computed. Different choices of basis functions and model parameter covariance are used to evaluate the sensitivity of red_tide output to these inputs. If the model includes only frequencies greater than or equal to Δf, spaced at increments of Δf, then the model will have a fundamental periodicity of T,

the record length, because this is the longest period represented in the model. To reduce this effect, we incorporate frequencies less than the fundamental frequency into **H**, starting at $\Delta f/2$ and increasing by intervals of $\Delta f/2$. This extends the model periodicity to $2T$, twice the record length (Figure 1.2A), and also reduces Gibbs-like behavior at the beginning and end of the fitted time series.

Separately, we examine the effects of different model parameter covariance matrices on the behavior of the fitted time series near a jump discontinuity. Two choices for the model parameter covariance matrix $\langle \mathbf{x}\mathbf{x}^{\mathrm{T}} \rangle$ are tested: (1) an assumption that the covariance is constant at all frequencies (a spectrally white process), and (2) an assumption that its power spectrum is proportional to f^{-2} (a spectrally red process representing the true spectrum in equation (1.12)). Two sets of basis functions are used: one that is comprised of sinusoids at the Nyquist frequency f_{Ny} and lower, and one that is comprised of sinusoids at frequency $2f_{\mathrm{Ny}}$ and lower. Both basis sets are spaced in frequency by $\Delta f = 1/T$. Discretely sampled sinusoids of frequencies greater than f_{Ny} are indistinguishable from sinusoids at the frequencies lower than f_{Ny} to which they are aliased. The larger but linearly dependent bases set is used in order to demonstrate the effect that **P** in equation (1.7) has in constraining an underdetermined system; in this case, describing the behavior of a time series near a step discontinuity at times between the sampling interval requires that frequencies greater than f_{Ny} be represented, constrained by the expected spectral power of the signal. The residuals are assumed to be uncorrelated (white noise), and the expected residual variance can be calculated from the total spectral power at frequencies above the highest-frequency basis. The expected fraction of residual variance is calculated from the integral of the true spectrum over frequencies not explicitly modeled:

$$\frac{\mathrm{var}(\mathbf{r})}{\mathrm{var}(\mathbf{y})} = \frac{\int_{f_{\mathrm{high}}}^{\infty} f^{-2}\mathrm{d}f}{\int_{f_{\mathrm{low}}}^{\infty} f^{-2}\mathrm{d}f} = \frac{f_{\mathrm{low}}}{f_{\mathrm{high}}}, \tag{1.13}$$

where f_{low} and f_{high} are the lowest and highest frequencies in **H**. Frequencies less than f_{low} are not included in the integral due to the singularity of f^{-2} at $f = 0$ and the approximation of

very low-frequency variance as a mean and trend. The total residual variance in this case is small, and hence the impact of the residual covariance matrix is expected to be negligible and is not examined in this example. The effects of different assumptions of residual covariance are examined later with data for which these effects are noticeable. Both fitted time series in figure 1.2A, periodic on T and on $2T$, incorporate a covariance matrix $\mathbf{P}$ constructed from a spectrum proportional to f^{-2}.

Near the jump discontinuity, there are two basis choices that we will consider: distinguishable basis sinusoids up to the Nyquist frequency and indistinguishable sinusoids up to twice the Nyquist frequency, which are aliased for $f > f_{\mathrm{Ny}}$. We also consider two choices for model parameter covariance: a constant value and a covariance proportional to the known spectral power of the data. Together, these form four regimes that can demonstrate the effects of basis choice and model parameter covariance on the resulting fitted approximations to the same step data. The OLS approach models the step function as a series of sinusoids up to the Nyquist frequency with no assumptions about the model parameter covariance. This results in the Gibbs phenomenon near the jump discontinuity (Figure 1.2B). This result is nearly identical to the partial sum of the Fourier series for this function (not shown), with a slight difference due to least squares' allowance of non-zero $\mathbf{r}$. Using the same basis functions but making the assumption that the main diagonal of the parameter covariance matrix $\mathbf{P}$ corresponds to a spectral slope of -2 as does the true spectrum, the red_tide procedure reduces the Gibbs phenomenon at the expense of greater misfit in the immediate vicinity of the jump discontinuity (Figure 1.2C).

The Gibbs phenomenon may be reduced by including additional frequencies. Sinusoids at frequencies greater than the Nyquist frequency will be aliased, however, so the effect of the model parameter covariance is more pronounced than in the former examples. Figure 1.2D shows the result of fitting a finite record to indistinguishable (resolved and aliased) frequencies while naively assuming a spectrally white process. Without the regularization provided by a sufficiently accurate $\mathbf{P}$, the duplicate bases in $\mathbf{H}$ render it rank deficient and therefore the interpolated fitted time series is unrealistically large in amplitude at unobserved times. Modeling

22

aliased frequencies with the assumption of an accurate (spectrally red) covariance, however, results in a fit that has further reduced the Gibbs phenomenon while also reducing the misfit at sampled times immediately before and after the jump discontinuity (Figure 1.2E). This approach (regularizing according to prior statistics) yields the smallest root-mean-square error of the four regimes, both at observation times and at the interpolated higher temporal resolution (Table 1.1). Though the solution in panel E does not pass through the observations like those in panels B and D, it nevertheless is the best representation of the underlying sampled process due to its reduction of the Gibbs phenomenon, despite fitting to the same data. Note that the accuracy of these interpolations is only quantifiable here due to knowledge of the true underlying function from which the data are perfectly drawn, which is not possible with real observations.

These refinements to the standard OLS approach demonstrate three advantages of the methods used in red_tide. First, the use of a fundamental frequency lower than that suggested by the record length reduces the effect of periodicity imposed by the model on the solution by lengthening the time scale of this periodicity. Second, the choice of model parameter covariance matrix $\mathbf{P}$ impacts the solution: a choice of $\mathbf{P}$ that is more representative of the true behavior of the data reduces the magnitude of the Gibbs phenomenon. Third, frequencies inaccessible to a discrete Fourier transform may be included to more realistically account for variance. The ambiguity of aliased frequencies (those exceeding the Nyquist frequency but at which variance is present) is reduced by using an accurate $\mathbf{P}$. We do not examine aliased signals elsewhere in this study.

1.3.2 Synthetic Time Series

Real ocean time series are substantially more complicated than the idealized step function above, with variance distributed across a range of frequencies, driven by a variety of physical processes. In order to simulate data with tidally-driven components, we generate synthetic time series from smooth spectra of predefined slope (the continuum), with sharp peaks in frequency space that simulate tides. In this way, we can exactly control and separate the signal and noise

components. In order to simulate tidal cusps, we consider a modulation of the tidal amplitude following the description of amplitude-modulated baroclinic tides off Hawai'i by Chiswell (2002). By the convolution theorem, multiplying (modulating) the tidal component in the time domain by a function is equivalent to convolving the Fourier transform of the tidal component with that of the modulating function in the frequency domain. For example, a purely sinusoidal high-frequency tide $\cos(\omega_T t + \phi)$ may be multiplied by an amplitude envelope $A(t)$ such that:

$$\eta(t) = A(t)\cos(\omega_T t + \phi), \tag{1.14}$$

the Fourier transform of which is

$$\begin{aligned} \mathcal{F}\{\eta(t)\}(\omega) &= \mathcal{F}\{A(t)\}(\omega) \otimes \mathcal{F}\{\cos(\omega_T t + \phi)\}(\omega) \\ &= \tilde{A}(\omega) \otimes \left(e^{-i\phi}\delta(\omega - \omega_T) + e^{i\phi}\delta(\omega + \omega_T) \right) \sqrt{\frac{\pi}{2}}, \end{aligned} \tag{1.15}$$

where $\mathcal{F}\{\cdot\}(\omega)$ is the Fourier transform to angular frequency (ω) domain, $\tilde{A}(\omega)$ is the Fourier transform of $A(t)$, and $\otimes$ denotes convolution. The constant $\sqrt{\pi/2}$ is a matter of convention such that the inverse Fourier transform is symmetric with the forward transform.

To generate the synthetic time series used in this study, a spectrally-red modulating function is used to simulate the cusps observed around tidal lines in the spectra of tidally-dominated ocean time series, and a random phase is assigned at each frequency before an inverse Fourier transform is applied to produce a synthetic time series with known spectral power. Samples from a red background are added to the time series to simulate broadband non-tidal ocean variability. Figure 1.3 shows an example of three synthetic spectra (upper panel) and corresponding hourly time series computed with random phase and truncated at 500 hours (bottom panel). The top panel also shows the power spectrum of observations in gray for comparison (these are discussed in Section 1.41.4.1). The phase at each frequency is identical for each time series, so only the spectral power differs. All three have peaks of equal

24

magnitude at the M_2 frequency. The red background spectrum is proportional to $1/\omega^2$ while the modulating spectrum is proportional to $1/(\omega^2 + \omega_0^2)$, where ω_0 is a small frequency introduced to eliminate the singularity at $\omega = 0$. In the red-background time series, there is more variability at low frequencies as in the real ocean, while the modulated one imitates the interaction of the tide with low-frequency processes. The time series with the spectrally white background has comparatively more power at super-tidal frequencies, which is evident in the time domain from the short time-scale noise that is less noticeable in the data with a red background. The data displayed are sampled hourly to be consistent with the choice of the Nyquist frequency, while the length of the time series is set by the choice of the fundamental frequency.

1.3.3   Effect of Noise Covariance on Amplitude, Phase, and Uncertainty

The estimated amplitude and phase of the tidal constituents are affected by the spectrum used to construct $\mathbf{R}$, also referred to here as the noise covariance or residual covariance. At each frequency represented in $\mathbf{H}$, the data are represented twice: explicitly in the frequency domain as model parameters ($\hat{\mathbf{x}}$) and implicitly in the residual ($\hat{\mathbf{r}}$) in the time domain, which has energy at all frequencies in general. Here, "explicit" versus "implicit" refers to whether energy is described by the model sines and cosines or by a residual time series with the fitted tides removed. Choosing the residual covariance matrix $\mathbf{R}$ to approximate the true covariance of the non-tidal component of the data improves the accuracy of the estimated coefficients. This is important because geophysical time series generally do not have flat spectra but rather spectra that decrease with increasing frequency. For example, if data with a red noise term ($S_{\text{noise}} \propto f^{-2}$, like the red and blue curves in Figure 1.3) were modeled using a spectrally white $\mathbf{R}$ (equivalent to the black curve), the variance would be treated disproportionately as signal at low frequencies (where the true noise is actually more energetic than what is given by $\mathbf{R}$) and disproportionately as noise at high frequencies (where the true noise is actually less energetic than what is given by $\mathbf{R}$). This is the case even if the spectrum corresponding to $\mathbf{R}$ has the same total energy as the true background.

25

Here we analyze a 1001-point synthetic record sampled hourly. If there is a low-amplitude, high-frequency signal present in the data, a high assumed noise level at that frequency would limit the detection of that constituent, as shown by the large blue intervals in Figure 1.4A at $f > 2$ cpd. A colored spectrum that matches the frequency dependence of the true noise component is an improvement over the simpler assumption of an uncorrelated (spectrally flat) white noise that does not match the true noise component. The former gives a relatively constant ratio of assumed spectrum to true spectrum across frequencies, while the latter gives a frequency-dependent ratio. This results in the variance of $\mathbf{Hx}$ being overrepresented at low frequencies, because the noise covariance is too low, and underrepresented at high frequencies, because the noise covariance is too high. Because variance is represented twice at modeled frequencies, the covariance matrices $\mathbf{R}$ and $\mathbf{P}$ serve as constraints on the partition of energy between $\mathbf{Hx}$ and $\mathbf{r}$ (recall that $\mathbf{R}$ represent covariance in time of a wide-sense stationary noise component and therefore has power at all frequencies, including those to which data are fitted).

Figure 1.4 shows the effects of these constraints in the frequency domain ("model space"): as described in section 1.31.3.2 and illustrated in Figure 1.3, a synthetic time series is constructed by adding a random background process with a spectral slope of -2 to a modulated semidiurnal (M_2) tidal process. In Figure 1.4A, the total spectrum (solid black) and its tidal component (signal, dashed black) and background (noise, dotted black) are plotted along with red_tide squared amplitude estimates ($\hat{a}^2 + \hat{b}^2$) calculated using a spectrally white noise covariance (blue), an approximate red noise covariance (red), and a "true" red noise covariance that matches the spectral power of the noise component (orange). All assume the same model parameter covariance matrix $\mathbf{P}$ constructed from the true tidal spectrum, not the total spectral density. Shading indicates 90% credible intervals of amplitude. These intervals are calculated using Monte Carlo sampling from a 10,000-member population obeying the posterior error of $\hat{\mathbf{x}}$ (equation (1.8)) to infer the distribution of $\hat{A}$. Figure 1.4B shows again the spectrum of the non-tidal (noise) component of the time series (dotted black line), the spectra corresponding to the noise covariance matrices (solid lines), and the posterior spectra of the residual time series (connected dots) calculated using

Welch's method for spectral estimation (overlapping windowed segments of length one fifth that of the record). For each noise covariance, the total noise energy is approximately 42% of the total signal energy integrated from $f = \Delta f$ to f_{Ny}. This matches the true noise-to-signal ratio. Both red noise analyses (red and orange lines in Figure 1.4B), which more closely approximate the true noise component, result in more accurate amplitude estimates at low frequencies where the signal-to-noise ratio is low (red or orange lines vs black dashed line in Figure 1.4A at frequencies less than 0.2 cpd), and greater precision at higher frequencies, where the signal-to-noise ratio is high. The calculated spectra of the residual time series are lower for increasingly accurate noise spectra, indicating that more variance is allocated to the model parameters.

Figure 1.5 compares the performance of the algorithm for both amplitude and phase under these three noise regimes. The ratio of median estimated amplitude $(\hat{a}^2 + \hat{b}^2)$ to the true signal amplitude is displayed in Figure 1.5A, which shows that all configurations produce accurate estimates near the tidal peak (2 cpd). At frequencies higher than about 5 cpd, the assumed variance of noise and model parameters are similar in magnitude (Figure 1.4B). In this frequency range, the amplitude ratios assuming red noise in Figure 1.5A show high variability with average values closer to one, indicating that they are less biased than the white-noise assumption, which is always greater than 1 for $f¿0.5$ cpd, but also less precise, because the amplitudes under the red noise assumption vary more above and below 1. Amplitude uncertainty is measured by the standard deviation of the estimated distribution of modeled amplitude, normalized by the true amplitude of the tidal component (Fig. 1.5B). The standard deviations of the amplitudes are smaller for the red noise assumptions than for the white noise assumption for all frequencies, and especially near the tidal peak and its cusp. The assumption of full, unapproximated red noise (orange dots in Fig. 1.5B) gives the lowest standard deviation. For the phase, all methods give comparable estimates (Fig. 1.5C). At higher frequencies, both red noise results are more precise than the white noise result. The uncertainty of the phase (Fig. 1.5D), here measured by the standard deviation of the estimated distribution of the centered phase, is lower near the tidal cusp and at higher frequencies for the red noise results. The standard deviation is lower for

the white noise result at frequencies lower than 0.4 cpd, but this implies unreasonable precision, as suggested by the inaccurate phase estimates at these frequencies in panel C resulting from a signal-to-noise ratio of less than 1; the corresponding precision is too high to be reasonable for such inaccurate estimates. At these low frequencies, the standard deviation of phase for the red noise analyses approaches $\pi/\sqrt{3}$, indicated by the horizontal dotted line, which is the standard deviation of a uniform distribution on the interval $(-\pi, \pi)$. This is expected for the phase of a low-amplitude signal in the presence of energetic noise that renders that signal's phase unrecoverable. At frequencies higher than the tidal cusp, the standard deviation of phase from the white noise analysis approaches this value because it assumes that noise is more energetic than the signal, while the red noise assumptions give results with lower variability, with the full red noise assumption giving the lowest variability.

The lessons learned from the idealized case of a short record with a tidal cusp and red noise also hold for a more complicated and realistic case in which a long record with a proportionately finer frequency resolution is modeled only at frequencies in limited bands. In Figures 1.4 and 1.5, all resolvable frequencies are modeled to illustrate the effect of noise representation on the partition of variance into signal and noise. In practice, however, tidal time series may be several years long, and so the number of data, and hence the number of resolvable frequencies, can be large enough such that modeling all frequencies explicitly is computationally prohibitive. Furthermore, in such cases the residual covariance matrix **R** cannot be practically constructed at all possible time lags, necessitating instead an approximation like that discussed earlier (truncated, in sparse form, and windowed to reduce spectral ringing). Figure 1.6 shows results analogous to those in Figure 1.4 for an hourly-sampled seven-year-long synthetic time series. In this case, only limited frequency bands are modeled, with the rest of the total variance accounted for by the residual. An accurate model parameter covariance enforces realistic amplitudes and uncertainty on the model parameters by setting a signal-to-noise ratio that approximates the true ratio of tidal energy to non-tidal energy as a function of frequency and not only the total, frequency-integrated ratio. This corresponds to the red curve in Figure 1.6. The

28

blue curve corresponds to an assumption of white noise: though the total frequency-integrated signal-to-noise ratio is the same for both estimates, the white noise assumption over-represents the variance of the non-tidal component in the semidiurnal band (around 2 cpd), giving larger credible intervals resulting from the unnecessary uncertainty introduced by overestimating the non-tidal variance. Though the red noise covariance is a windowed, truncated approximation of the true non-tidal covariance, its spectrum (dotted red) closely matches the true underlying spectrum of non-tidal variability (dashed black). With only 500 frequencies modeled out of the over 60,000 that are fully resolvable, the reconstructed time series accounts for more than 45% of the total variance and estimates tidal and off-tidal amplitudes within error. This shows that modeling all resolvable frequencies directly in $\hat{\mathbf{x}}$ is not only computationally impractical but also unnecessary, even for data with substantial non-tidal energy like those in Figure 1.4.

1.3.4 Effect of Noise Covariance and Record Length on Constituent Estimates

In order to evaluate the impact of the choice of noise covariance and record length on estimated tidal coefficients, we ran ten Monte Carlo experiments. These ten experiments used varying combinations of tidal energy (two regimes) and background structure (five regimes), which are described below. The six runs with more negative spectral slopes ($\gamma = -3, -2.5, -2$, corresponding to panels A-C and F-H in Figures 1.7 and 1.8) used 10,000 sample time series, while the four runs with less negative spectral slopes ($\gamma = -1.5, -1$, corresponding to panels D, E, I, and J) used 50,000 sample time series to obtain stable results. Each time series had 24,000 hourly samples but was only analyzed up to 1500 hours in order to investigate the effect of record length on estimated parameters, with the exception of the $\gamma = -1$ case, which was analyzed up to 5000 hours in order to examine convergence at longer record lengths. Three tidal constituents typically seen in observations (K_1, M_2, and S_2) are added onto a synthetic noise background. Analysis for only the M_2 constituent is shown, as other tidal constituents are qualitatively similar. Only phase varies from one time series to another within a run, with signal and noise amplitude

29

constant for all ensemble members. The phase is varied by randomly selecting from a uniform distribution on the interval $[0, 2\pi)$, and is varied separately for the noise and signal components. Two quantities derived from model parameters are examined to determine the impact of noise covariance and record length: the bias of tidal amplitude estimates as a measure of accuracy, and the variance of model parameters (harmonic coefficients) about the true values as a measure of precision.

The observed bias of estimated tidal amplitude $\hat{A}$ is affected by the variance of the tide relative to the background time series, the record length, and the choice of noise covariance. Figure 1.7 shows the mean ratio of the estimated M_2 constituent amplitude to true amplitude across the ten Monte Carlo experiments:

$$\text{Amplitude Ratio} = \frac{\left\langle \sqrt{\hat{a}^2 + \hat{b}^2} \right\rangle}{\sqrt{a^2 + b^2}}, \tag{1.16}$$

where brackets indicate averaging across Monte Carlo simulations, a and b are respectively the sine and cosine coefficients of the tidal constituent in question (here M_2), and hats denote estimates. Note that a and b vary with each Monte Carlo simulation, but the sum of their squares does not, hence $\sqrt{a^2 + b^2}$ can be pulled out of the brackets. The amplitude ratio is a measure of the accuracy of tidal harmonic estimates in terms of a simple quantity of physical relevance, tidal amplitude. Each panel shows this quantity as a function of the assumed noise spectral slope and record length. The panels are organized with steep spectral slopes at the top and successively flatter spectral slopes below. The left column (A-E) represents a strong tide case meant to simulate bottom pressure, with tidal spectral power 100 times that of the noise background at tidal frequencies, and the right column (F-J) shows a weak tide case meant to simulate tidal velocities or baroclinic tides, with tidal spectral power 10 times that of the local background. A perfect tidal amplitude estimate would have a normalized value of 1, and cases for which the interquartile range of simulations do not include 1 are hatched in Figure 1.7, as

they do not provide accurate estimates of the tidal amplitude. Records with steeper spectral slope background processes (Figure 1.7 A-B, F-G) have less accurately estimated amplitudes than those with less steep background spectral slope (Figure 1.7 D-E, I-J), while the latter required more Monte Carlo simulations to show convergence to a record length-dependent amplitude accuracy. Additionally, records with less spectrally-steep backgrounds are less sensitive to the choice of noise covariance. In all cases, the use of a noise spectral background that matches the spectral structure of the data yields a more accurate estimate of amplitude than the use of a noise spectral background with a substantially different slope. These results are in agreement with the expectation that a more energetic and structured background process is best described by a residual component with a covariance matrix $\mathbf{R}$ constructed from an appropriate power spectrum. Amplitude estimates also improve with increasing record length, though this effect is smaller than the effect of the noise covariance.

Estimated amplitudes are comparably accurate and in some cases more accurate for spectral slopes less than (steeper than) the true slope when compared to the amplitude accuracy when using the true slope to construct $\mathbf{R}$. This is, however, a small effect that appears to diminish with increasing record length. For example, in Figure 1.7 E and J, an assumed spectral slope of -1.5 achieves more accurate results than the true slope of -1, which gives lower amplitude ratios, an effect that is stronger at shorter record lengths. This may be due to the bias of the estimators $\hat{a}$ and $\hat{b}$: the squared error of $\hat{a}$ and $\hat{b}$ are minimized in the least squares procedure and these estimators are biased low due to regularization (Hoerl & Kennard, 1970; Wunsch, 1996). The reduction of parameter error at the expense of added bias inherent to this method is evident in the difference between amplitudes estimated from OLS (noise slope "None") and the regularized methods. In particular, the use of a white noise covariance (noise slope 0) differs from OLS only by its inclusion of the regularizing term $\mathbf{P}$ (noise slope of 0 means that $\mathbf{R}$ is a scaled identity matrix that can be treated as a constant and would be canceled out if $\mathbf{P}$ were 0). OLS overestimates tidal amplitudes consistently while the regularization of the otherwise identical white noise approach typically underestimates tidal amplitudes except for the case of

strong tides and a spectrally steep background (Figure 1.7 A-B) for which such a white noise assumption is not suitable.

Model parameters $\hat{\mathbf{x}} = [\hat{a}_1, \hat{b}_1, \hat{a}_2, \hat{b}_2 ...]$ are estimators for the true parameters $\mathbf{x}$ and, like amplitude $\hat{A}$ derived from them, better estimate truth with increasing record length, a more accurate noise covariance, and a higher signal-to-noise ratio. As a measure of precision of these parameter estimates, we calculate their variance about the true mean, normalized by the true amplitude:

$$\text{Normalized Parameter Variance} = \frac{1}{2\sqrt{a^2 + b^2}} \left(\left\langle (\hat{a} - a)^2 \right\rangle + \left\langle (\hat{b} - b)^2 \right\rangle \right). \qquad (1.17)$$

The base-10 logarithm of this quantity is shown in Figure 1.8 in order to better compare results across orders of magnitude. Due to the random phase, the statistics of sine and cosine coefficients are identical, constrained by $a^2 + b^2 = A^2$, where A is constant across all Monte Carlo simulations. Therefore, both $\hat{a}$ and $\hat{b}$ are used in calculating the normalized parameter variance, effectively doubling the number of realizations.

As with amplitude, when the random background has a more negative spectral slope, estimates of tidal coefficients are more sensitive to noise covariance choice (Figure 1.8 A-B, F-G) than they are when the background has a less negative spectral slope (Figure 1.8 D-E, I-J). Unlike amplitude, coefficients are primarily sensitive to record length when the background noise is relatively unstructured, with increasing sensitivity to the choice of noise covariance as the true noise spectrum becomes steeper. This can be seen in the gradient of the normalized parameter variance in each panel: in panels A-B and F-G, where γ is more negative, the variance decreases for more accurate noise slopes, whereas in panels D-E and I-J, where γ is less negative, the variance decreases more strongly with increasing record length. Filled dots are placed at elements corresponding to the lowest value (highest precision) in their respective rows. For all simulations with a background process of spectral slope $\gamma \geq -2$ (less steep), precision is highest for the noise assumption that matches the true spectral slope. For $\gamma < -2$ (steeper), this is also

the case for sufficiently long records, and for short records it is still better to match the noise spectrum than to use OLS.

In almost all cases, using equation (1.7) (with **P** and **R**) instead of equation (1.9) (OLS without an explicit noise assumption, labeled "none" in the rightmost column of every panel in Figures 1.7 and 1.8) resulted in more accurate estimates for A (ratios closer to 1 in Figure 1.7) and more precise estimates for **x** (lower parameter variance in Figure 1.8), at least for the cases examined here in which tidal energy is 10 to 100 times greater than that of noise at the same frequency. The OLS approach, which is widely implemented in tidal harmonic analysis, was comparable to the red_tide approach for $\gamma = -1.5$ and $\gamma = -1$, indicating that it is most suitable for tidal analysis of data with a spectral background that is nearly white. Data with steeper background spectra benefit from treating the residual **r** as a spectrally red process by way of the covariance matrix **R**.

1.4   Application to Oceanographic Data

1.4.1   Bottom Pressure

The methods implemented in red_tide seek to address several potential issues that can arise when harmonically analyzing real tidal time series, each of which are presented in isolation in section 1.3 by using synthetic data. The first demonstration of red_tide on observations uses a bottom pressure record, a data set with low background noise where record length is sufficient and non-phase-locked energy is much weaker than the phase-locked tide. Bottom pressure measurements from the NOAA Deep-ocean Assessment and Reporting of Tsunamis (DART) (NOAA, 2005) are dominated by the largely coherent barotropic tide. The time series examined here originates from site 51406 (8.48°S, 125.03°W) and spans 3 years and 6 months of observations, from February 12, 2011 to September 6, 2014. Many tidal constituents of amplitudes spanning several orders of magnitude are present in this record, as seen by tidal lines in Figure 1.9. Pressure measurements have a significantly lower noise level than coastal surface

height gauges and have accurate harmonic constituents even over short records (Le Provost, 2001). Bottom pressure is therefore useful for evaluating the accuracy of harmonic decomposition in the regime of tidally-dominated, low noise observations, which typically do not pose major problems when calculating constituents. Hourly averaging of 15-second sampled data further suppresses noise and instrumental artifacts, such as digitization, and does not alias major constituents.

The power spectrum of the bottom pressure time series (Figure 1.9) exhibits many prominent peaks, of which 22 are singled out for analysis. These 22 frequency bands together account for more than 99% of the time series' variance; therefore both the white noise assumption (not shown) and red noise assumption produce essentially identical results. Figure 1.10 shows the output from red_tide using only the red noise assumption for $\mathbf{R}$ with spectral slope $\gamma = -3/2$ alongside t_tide output. Where t_tide models only tidal frequencies associated with astronomical parameters, this analysis includes those same frequencies and 30 additional frequencies per constituent in a band of 10 year^{-1} centered on that tidal frequency in order to capture modulation at annual and semiannual cycles and cusp-like spreading of peaks. These harmonic amplitudes are spaced at $\Delta f = 1/3$ year^{-1}, a frequency step smaller (and hence of higher resolution) than that of the power spectrum, which is coarser in resolution due to segmenting. This corresponds to a 3-year period, shorter than the record length of 3 years and 206.5 days in order to ensure the resolution of annually periodic modulation of the main tidal constituents, regardless of the exact record length. The amplitudes are normalized to have units of spectral power for comparison. The results of analysis by t_tide are also normalized and plotted with 90% confidence intervals (the t_tide code that defaults to 95% confidence intervals is modified).

The tidal amplitudes given by t_tide largely match the red_tide results for high-amplitude constituents and fall within credible intervals at low-amplitude constituents, while providing analysis at more frequencies. Focusing on a single cluster of constituents shows this more clearly (Figure 1.11). The semidiurnal band, centered about the energetic M_2 constituent, contains several other well-resolved tidal constituents (Darwin symbols 2"N_2, μ_2, N_2, ν_2, λ_2, L_2, S_2, and K_2) that result from the complicated gravitational tidal forcing potential, many of which are

34

not resolvable when using segmenting methods or are not exactly mapped by a simple Fourier transform, which would result in spectral leakage. Additionally, the energy of the cusps is explicitly modeled, with intervals that reflect the uncertainty associated with a realistic noise level that is a function of frequency, which becomes important for these low amplitude components. The characteristics of the non-phase-locked component of the tides can therefore be diagnosed from these cusps. The high signal-to-noise ratio at energetic constituents, on the other hand, means that the noise level is less important at these frequencies, hence the OLS approach of t_tide works well for these constituents.

1.4.2   High-Frequency Radar

Like OLS, the red_tide package effectively models data with weak noise and tidal constituents that are highly coherent with astronomical forcing. The data regime for which red_tide is designed includes higher levels of structured noise and tidal energy that cannot be predominantly described by a small number of frequencies. Surface currents, which are driven by wind, tides, eddies, and mean flow, fall under this category.

Observations of surface currents are obtained from a high-frequency radar network (HFRnet, see Terrill et al., 2006). The California Current System (CCS), a region that is well-sampled by this network (Roarty et al., 2019), is used to evaluate the harmonic decomposition technique. Radial velocities measured by antenna stations are mapped to a Cartesian grid of zonal and meridional velocities using a least squares fit (Ohlmann et al., 2007). Surface currents are driven by a wide range of dynamics: direct wind forcing, near-inertial motions, interannual variability of the local current system, and tides, including tidal currents and the surface expression of internal tides. This contrasts with bottom pressure, which is dominated by tides, with many more prominent tidal frequencies than appear in surface currents.

Figure 1.12 shows the averaged rotary power spectrum (spectral power partitioned by rotational polarization) over the grid points in a region of the CCS ranging from 33.7561°N to 38.1252°N out to approximately 100 km from the coast (for the formulation of rotary spectra,

35

see Gonella, 1972). This formulation is convenient for visualizing the spectral power of surface currents due to the polarized flow resulting from near-inertial oscillations, which at these latitudes occur at frequencies between 1.12 to 1.26 cycles per day. Tidal peaks in HFR surface currents are pronounced and are comparable to or greater than the energy in the inertial and low-frequency (¡0.4 cpd) bands. These spectra are calculated using Welch's method, and the error bar denotes the ratio of high to low estimates for the power spectrum for a confidence level of 95%; this ratio is constant on logarithmic axes and does not vary with frequency. It is calculated from a chi-squared distribution for degrees of freedom equal to twice the number of windowed segments whose periodograms are averaged to calculate the spectrum (e.g. Bendat & Piersol, 2010), divided by 9 from the assumption that neighboring sites are correlated.

An hourly-sampled, 9 year and 3 month-long (January 1, 2012 to April 1, 2021) HFR surface current time series from 35.5361°N, 121.1776°W, roughly 7 kilometers off the coast of San Luis Obispo County, California, is analyzed in four frequency bands: low frequency (0.00732 cpd and less), S_1 solar diurnal (1 cpd), M_2 lunar semidiurnal (1.932 cpd), and S_2 solar semidiurnal (2 cpd). These frequency bands, shown in Figure 1.13A, account for roughly a fifth of the variance of the time series for both zonal velocity u and meridional velocity v. Therefore, 80% of the variance is included in the residual time series. Harmonic coefficients from red_tide are calculated using a model parameter covariance built from the domain-averaged power spectrum (solid black line) from 1191 grid points at latitudes ranging from 33.7561°N to 38.1252°N, out to about 100 kilometers offshore. The individual power spectrum calculated from the analyzed time series is shown in gray for comparison: smoothed peaks due to low frequency resolution and spectral leakage due to sampling and windowing result in an individual time series' spectrum that does not capture features that the least squares approaches can. The noise (residual) spectrum has a constant spectral slope of -1 and a corresponding covariance truncated at 300 hours lag, resulting in modest spectral ringing (dashed line). Amplitudes given by red_tide are calculated as half the sum of the sine- and cosine-coefficients squared, $\frac{1}{2}(a^2 + b^2)$, and are shown in red with 90% credible intervals, while equivalent amplitudes and 90% confidence

36

intervals given by t_tide are shown in blue. Filled blue circles are constituents that t_tide considers "significant", where the amplitude is greater than the amplitude error (Pawlowicz et al., 2002), while hollow circles are not considered significant. Only results for *u* are shown, as results for *v* are qualitatively similar; the full 2-dimensional character of time series in this region is shown in Figure 1.12 to illustrate the frequency-dependent polarization of surface currents, which may be modeled with red_tide. The calculation of rotary coefficients from red_tide output for *u* and *v* is straightforward, though results are not shown.

As with bottom pressure, harmonic decomposition reveals sharp peaks at tidal frequencies that are not resolved in power spectra with lower frequency resolution due to averaging over time intervals shorter than the record length. Because the choice of basis functions for red_tide is arbitrary, the fundamental frequency of the data set does not necessarily limit the spacing of modeled frequencies, though in practice the tolerance of basis nonorthogonality and resulting uncertainty will limit the choice of frequencies. The high noise level in these data results in large uncertainty at frequencies around tidal peaks as in Figure 1.13B, as indicated by shaded intervals. Despite this, annual modulation of the surface current is evident by the second and third most energetic peaks in the M_2 band appearing at $f_{M_2} \pm f_{\text{annual}}$, indicated by vertical dashed lines. The output of t_tide is also able to discern these modulation peaks, though it does not examine broadband modulation.

1.5   Summary and Discussion

The methods outlined here and implemented in red_tide are intended to provide best estimates of tidal amplitudes for data with red background spectra and significant tidal cusps. Red_tide incorporates a red noise covariance and includes additional frequencies beyond those of the astronomical forcing to accommodate data with highly energetic and correlated non-tidal components, a weak tidal signal relative to non-tidal processes, or a modulated tidal component with energy distributed across a band of frequencies. Short records, for which long-period

variance appears trend-like, also benefit from these methods because variance at fitted frequencies is allocated to model parameters according to prior statistical assumptions. The spectrally colored noise covariance is constructed to approximate the spectral properties of the non-tidal component of the data, and may be truncated and represented as a sparse matrix for computational efficiency with a window function applied to off-diagonal elements to suppress spectral artifacts that result from truncation. On the other hand, highly coherent tidal records with well-defined peaks and small cusps, such as bottom pressure, are well described by OLS, as the background noise is orders of magnitude lower in amplitude than the tidal signal.

These methods also address time series for which the choice of model covariance impacts results. We demonstrate this with a step function, a simple case that exhibits an extreme mismatch between the fit and the data when an inappropriate model parameter covariance is used, resulting in the well-known Gibbs phenomenon. We have found that when fitting a discrete step function across resolvable frequencies, the assumption of a realistic covariance reduces the magnitude of the Gibbs phenomenon near the jump discontinuity when compared to an assumption of constant covariance matrix $\langle \mathbf{x}\mathbf{x}^\mathrm{T} \rangle = \sigma_\mathbf{x}^2 \mathbf{I}$, and prevents dramatic overshoot when fitting to frequencies greater than the Nyquist frequency, which would otherwise be aliased (Figure 1.2 B-E).

The accuracy and precision of model parameters, given explicitly as a posterior covariance matrix, are impacted by the choice of residual (noise) covariance. Spectrally colored time series may have a residual background that varies over orders of magnitude across tidal bands, necessitating an appropriate noise covariance matrix if all constituents are to be estimated optimally. When the user-specified residual spectral power is significantly lower than the total spectral power of the time series, red_tide allocates more variance to the model parameters. If the energy of the non-tidal component is well understood, variance can be realistically allocated between the estimated signal $\mathbf{H}\hat{\mathbf{x}}$ and noise $\hat{\mathbf{r}}$ at the same frequencies. Estimating tidal amplitudes can be computationally expensive for long records ($O(10^5)$ or more data) because of the memory requirements for constructing and multiplying large matrices. In these cases, we have identified strategies for approximating the covariance matrix in order to reduce computational requirements,

including truncating and windowing the matrix in sparse form.

In summary, red_tide was designed to estimate tidal coefficients while incorporating prior assumptions that accurately account for the spectral structure of underlying noise and allow flexibility in the choice of modeled frequencies, which is important for data with a modulated, non-phase-locked tidal component. Longer records and less strongly correlated noise benefit less from this flexibility. Ordinary least squares is comparatively less suited to computing tidal harmonics from data with spectrally colored noise, especially red noise with a steep spectral slope. The code is available for use and modification, the details of which are in the appendix.

Acknowledgments

Data Statement

Relevant data and scripts, including those used to generate all figures and results in this paper, are publicly available through the University of California, San Diego library digital collections at `https://doi.org/10.6075/J080515G`. Readers interested in accessing observational data may visit `https://doi.org/10.7289/V5F18WNS` for bottom pressure data and `https://dods.ndbc.noaa.gov/thredds/hfradar.html` for high-frequency radar surface current data.

1.6   Appendix for Chapter 1

1.6.1   Downloading red_tide

The red_tide package is available for download as a GitHub repository at https://github.com/lkach/red_tide and in archived form (see Data Availability Statement). This package is written in the MATLAB language, but translation to other programming languages is welcome and encouraged. It has also been designed to work in the free software GNU Octave language and Octave-specific instructions are provided with the software release. Input for red_tide is flexible, with several options and default settings.

1.6.2   Noise as an Autoregressive Process

The residual time series $\hat{\mathbf{r}} = \mathbf{y} - \mathbf{H}\hat{\mathbf{x}}$ will have a power spectrum that is similar to that of $\mathbf{y}$, except in the bands of modeled frequencies. With the energetic tidal components removed, a spectrally-red residual can be modeled as an autoregressive process AR(p), where order p is the maximum number of time steps for which the system has memory (von Storch & Zwiers, 2003):

$$r_t = \alpha_0 + \sum_{k=1}^{p} \alpha_k r_{t-k} + \epsilon_t, \tag{1.18}$$

represents an order-p AR process r at time t with a white noise component ϵ. The AR parameters α_k can be estimated from the Yule-Walker equations (e.g. von Storch and Zwiers, 2003). From these, the spectral density of r can be estimated by

$$\hat{S}_{\mathbf{rr}}(f) = \frac{\hat{\sigma}_\epsilon^2}{\left| 1 - \sum_{\ell=1}^{p} \hat{\alpha}_\ell e^{-2\pi i \ell f} \right|^2}, \tag{1.19}$$

where hats indicate estimates of a true value, f is frequency, and σ_ϵ^2 is the variance of the white noise ϵ. The spectrum of the AR-modeled residual time series can be used as an estimate of the spectrum of underlying non-tidal, non-wind-driven intrinsic variability in the ocean. If $\hat{S}_{\mathbf{rr}}(f)$

accurately models the spectrum of **y** at frequencies outside those in **H** and assuming white noise ϵ, the coefficients $\{\hat{\alpha}_k\}$ may be used to construct **R** for a second iteration of fitting using this new estimate for the noise covariance matrix. This can be done by taking the inverse Fourier transform of $\hat{S}_{\mathbf{rr}}(f)$, which is the covariance of **r** via the Wiener-Khinchin theorem. **R** is then constructed from the covariance as outlined in section .

1.6.3 Alternative Form of Equation 1.7

An alternative equation that solves for $\hat{\mathbf{x}}$ can be achieved using the matrix inversion lemma (e.g. Wunsch, 1996):

$$\hat{\mathbf{x}} = \mathbf{PH}^{\mathrm{T}} \left(\mathbf{HPH}^{\mathrm{T}} + \mathbf{R} \right)^{-1} \mathbf{y}. \tag{1.20}$$

In practice, tidal records usually contain many more data points N than frequencies M that are of interest to model. Therefore the data-space inversion (equation (1.7)) is used in red_tide because $\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}$ has lower computation and memory requirements ($4M^2$ elements) than $\mathbf{HPH}^{\mathrm{T}}$ (N^2 elements). The $N \times N$ residual covariance matrix inverse, $\mathbf{R}^{-1}$, that appears in equation (1.7) does not need to be computed explicitly when using efficient linear system solution algorithms for matrix inversion, which instead directly calculate $\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}$, which is $2M \times N$. Further, a sparse representation of **R** minimizes memory requirements in equation (1.7) compared to the more challenging requirements of the dense $N \times N$ matrix in equation (1.20).

1.6.4 Non-dimensionalization

The assumption of uncorrelated noise is physically unrealistic when modeling processes with a correlated, energetic noise component, which would be better represented by a more complicated **R**. For example, Coles et al. (2011) use a non-diagonal matrix expression for **R**, the efficient inversion of which is outlined here. The computational resources to directly invert a non-diagonal **R** are too high to be practical: **R** is $N \times N$, where N is the length of **y**, which is

much longer than $\mathbf{x}$ in practice. Though red_tide uses MATLAB's default linear equation solving method, a Cholesky lower triangle factorization of the residual covariance matrix $\mathbf{R} = \langle \mathbf{r}\mathbf{r}^{\mathrm{T}} \rangle$ may also be used. The residual covariance can be factored as $\mathbf{R} = \mathbf{R}^{1/2}\mathbf{R}^{\mathrm{T}/2}$, where $\mathbf{R}^{1/2}$ is lower triangular with the inverse $\mathbf{R}^{-1/2}$ used as a (non-unique) whitening transform:

$$\mathbf{y_w} = \mathbf{R}^{-1/2}\mathbf{y}, \qquad \mathbf{r_w} = \mathbf{R}^{-1/2}\mathbf{r}, \qquad \mathbf{H_w} = \mathbf{R}^{-1/2}\mathbf{H}. \tag{1.21}$$

Noting that $\mathbf{R}^{\mathrm{T}/2}\mathbf{R}^{-1}\mathbf{R}^{1/2} = \mathbf{I}$, these can be substituted into equation (1.7), which simplifies to

$$\hat{\mathbf{x}} = \left(\mathbf{H_w}^{\mathrm{T}}\mathbf{H_w} + \mathbf{P}^{-1}\right)^{-1}\mathbf{H_w^T}\mathbf{y_w}. \tag{1.22}$$

The choice of $\mathbf{R}$, and thus $\mathbf{R}^{-1/2}$, relies on accurately estimating the residual covariance. This may be done by examining the calculated data covariance matrix $\mathbf{y}\mathbf{y}^{\mathrm{T}}$ or equivalently the power spectrum $S_{\mathbf{yy}}$ and estimating $\langle \mathbf{r}\mathbf{r}^{\mathrm{T}} \rangle$.

Analogously, the model parameter covariance matrix $\mathbf{P}$ can be expressed as $\mathbf{P} = \mathbf{P}^{1/2}\mathbf{P}^{\mathrm{T}/2}$. Define $\mathbf{H}' = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{P}^{1/2}$ such that equation (1.22) can be written as

$$\hat{\mathbf{x}} = \mathbf{P}^{1/2}\left(\mathbf{H}'^{\mathrm{T}}\mathbf{H}' + \mathbf{I}\right)^{-1}\mathbf{H}'^{\mathrm{T}}\mathbf{y_w}. \tag{1.23}$$

This is equivalent in form to the regularized least squares problem with a left-multiplied factor of $\mathbf{P}^{1/2}$ that allows $\hat{\mathbf{x}}$ to retain units of $\mathbf{y}$.

Chapter One, in full, is a reprint of the material as it appears in the Journal of Atmospheric and Oceanic Technology, 2022. Kachelein, L., B. D. Cornuelle, S. T. Gille, and M. R. Mazloff (2022) Harmonic Analysis of Non-Phase-Locked Tides with Red Noise Using the red_tide Package. *Journal of Atmospheric and Oceanic Technology*, 39(7):1031 – 1051, https://doi.org/10.1175/JTECH-D-21-0034.1. The dissertation author was the primary investigator and author of this paper. © American Meteorological Society. Used with permission.

Table 1.1. Root-mean-square error for the four step function analyses shown in Figure 1.2(B-E). RMSE is calculated at observation times (top row) and at a set of times sampled at 100 times the observational resolution (bottom row).

	White $\mathbf{P}$ $f_{max} = f_{Ny}$	True $\mathbf{P}$ $f_{max} = f_{Ny}$	White $\mathbf{P}$ $f_{max} = 2f_{Ny}$	True $\mathbf{P}$ $f_{max} = 2f_{Ny}$
RMSE at observation times	0.0080	0.0084	0.0122	0.0038
RMSE of interpolated time series	0.0211	0.0201	0.5014	0.0171
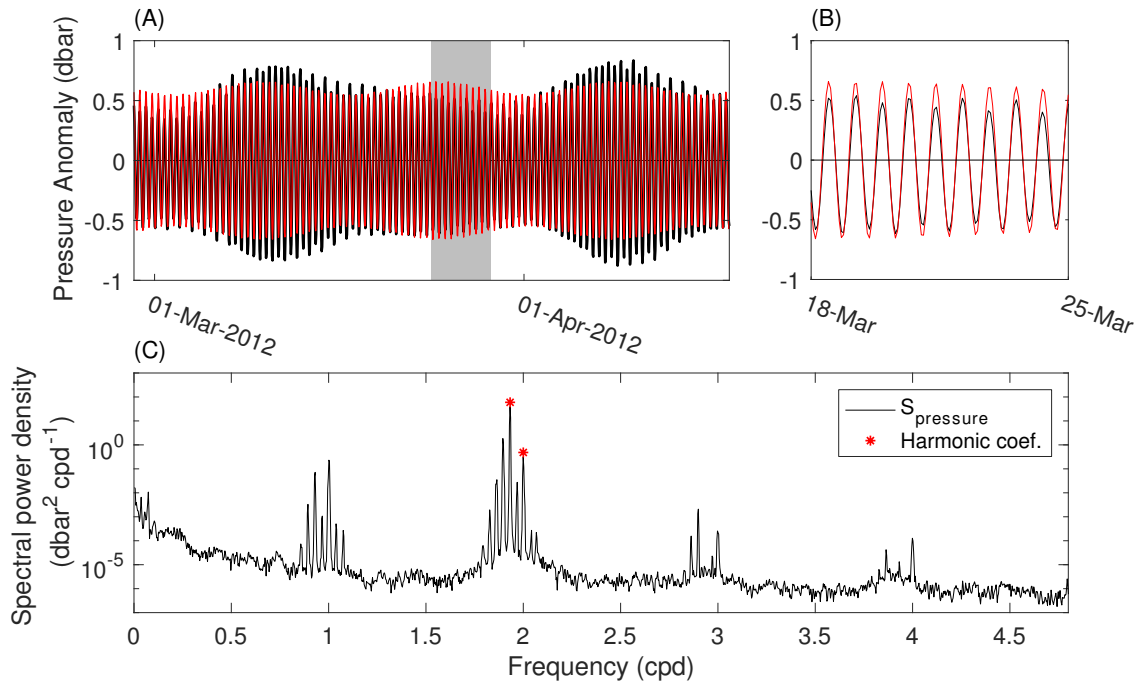
Figure 1.1. A simplified example of tidal harmonic analysis on bottom pressure data (3.5 year record at 8.48°S, 125.03°W, DART array site 51406, further analyzed in section 1.41.4.1). (A) A segment of the longer time series that was analyzed (black) and the resulting fit (red) to only two tidal constituent frequencies: M_2 and S_2. (B) A closeup of the time series over the gray interval in (A). (C) The power spectrum of the time series and the magnitude of the harmonic coefficients.
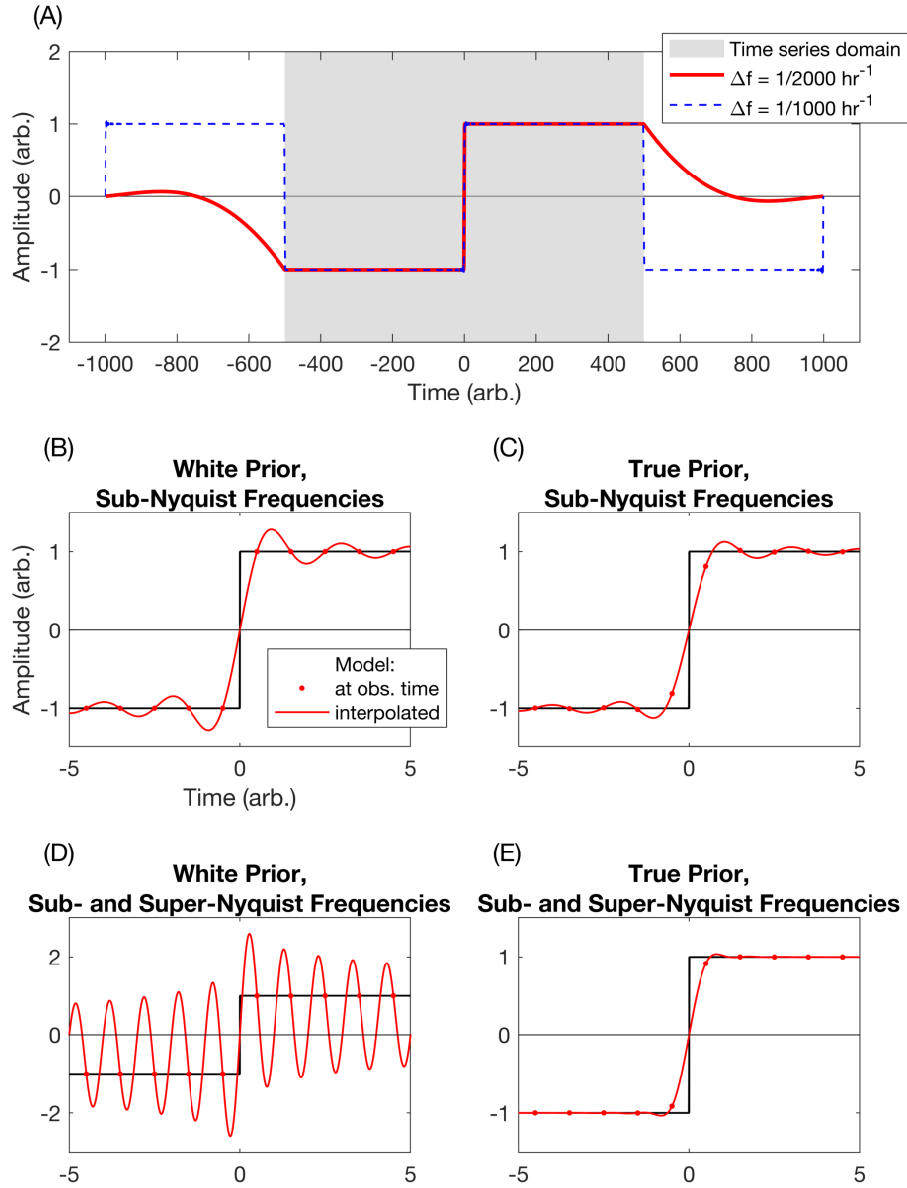
Figure 1.2. Step function discretely-sampled at 1000 points. (A) Full record modeled with orthogonal sinusoids (blue dashed) and with non-orthogonal sinusoids corresponding to a record twice as long (red solid). The gray shaded region indicates the range of data that are fit, outside of which the time series is extrapolated. (B) Close view of discontinuity and red_tide output given spectrally white (incorrect) model parameter covariance, $\mathbf{P}$, only modeling up to the Nyquist frequency f_{Ny}. (C) Discontinuity and output using true $\mathbf{P}$, only modeling up to f_{Ny}. (D) Same as panel B but including model frequencies above f_{Ny} up to $2f_{Ny}$. (E) Same as panel C but including model frequencies above f_{Ny} up to $2f_{Ny}$. The residual is treated as white (uncorrelated) noise in all panels, with minimal contribution to the total variance (0.2% for the blue dashed curve in panel A, 0.1% for the red curve in panel A, 0.1% for panels B-C, and 0.05% for panels D-E).

44
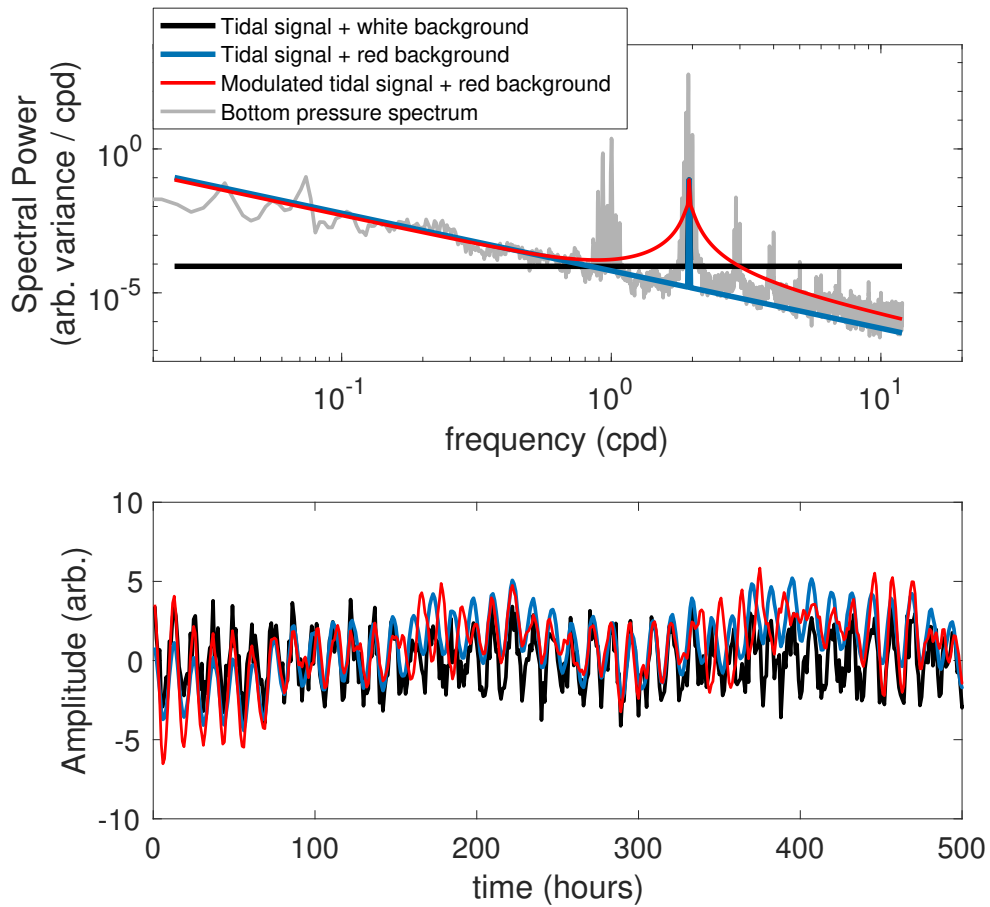
Figure 1.3. (top) Example of idealized power spectra with a single constituent, here M_2, with the computed power spectrum of the bottom pressure record from Figure 1.1 included for comparison, and (bottom) segments of corresponding time series computed from the power spectra with random phase at each frequency (500 hourly synthetic data points). All three spectra share the same tidal peak magnitude, and all three time series share the same phase across frequencies.
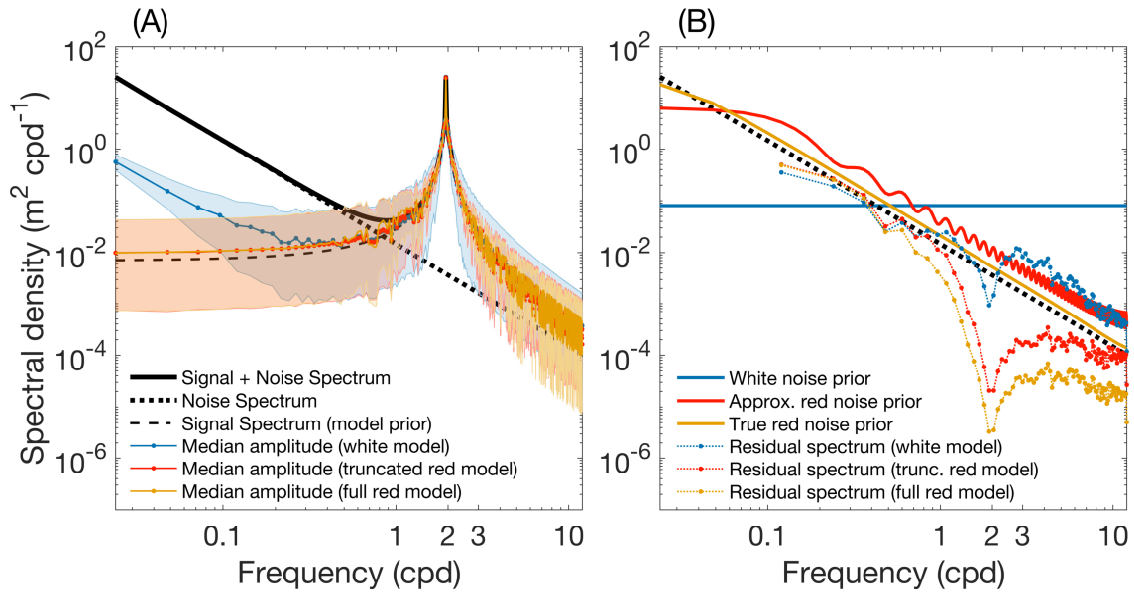
Figure 1.4. (A) Synthetic true spectra (noise and signal) and harmonic amplitudes calculated using red_tide, and (B) assumed and calculated noise spectra. Shading indicates 90% credible intervals of amplitude, with thick lines indicating median amplitude based on the estimated distribution of coefficients. All quantities are scaled to spectral density units for comparison. The true spectra are used to construct the noise and signal components of the time series (1001 hourly data points) from which harmonic amplitudes are calculated. The noise spectra in (B) indicate the Fourier transforms of the residual covariance matrix **R** used in each application of the red_tide procedure. All noise spectra used to construct **R** have the same variance, which approximates the 42% contribution of true noise variance to the true total variance. The residual spectra are calculated using Welch's method with a Hanning window and 9 overlapping segments. The tidal peak is at the M_2 frequency and is convolved with a spectrum of slope -2 in order to simulate real tidal modulation that leads to tidal cusps.
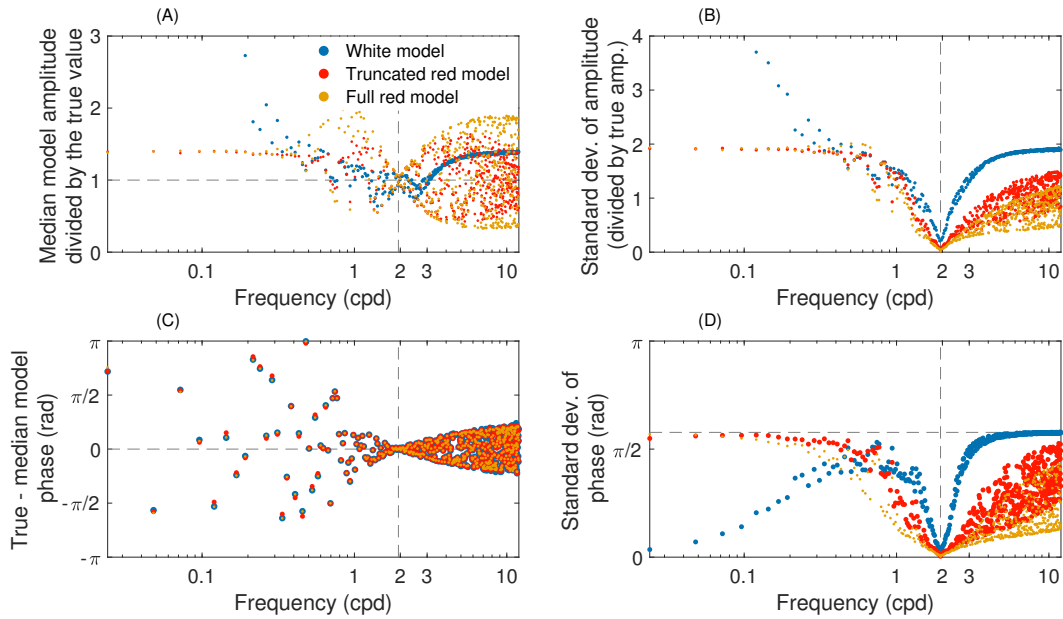
Figure 1.5. For the same time series and analyses in Figure 1.4, (A) model amplitude fraction of truth, (B) standard deviation of the normalized amplitude, (C) phase difference between the true phase and the model estimates, and (D) standard deviation of the estimated phase after being centered on the interval $(-\pi, \pi)$.
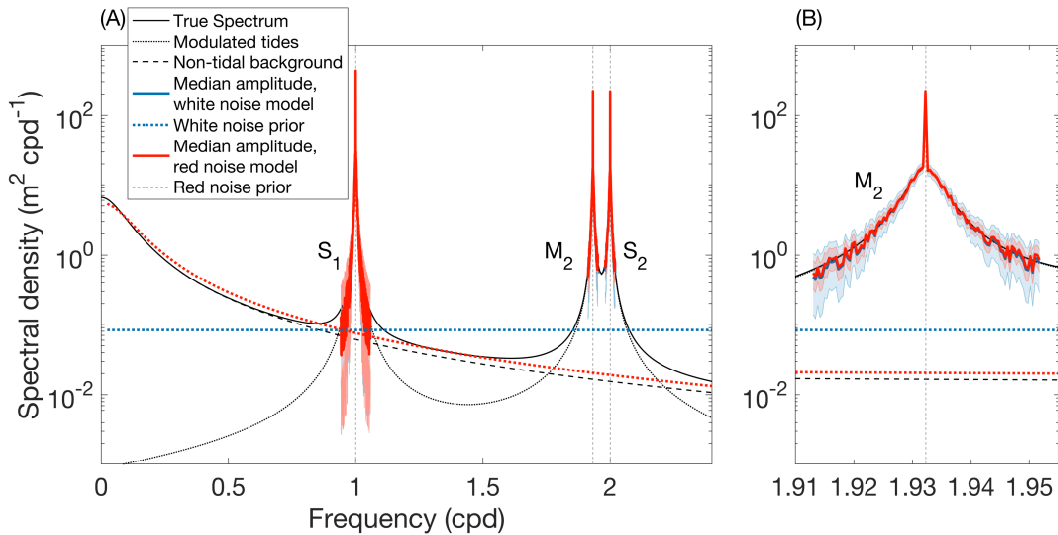
Figure 1.6. Synthetic spectrum, harmonic amplitudes calculated from corresponding time series (7 years sampled hourly) using red_tide (scaled to spectral density units), and spectra used to construct **R**, shown (A) in the range of frequencies typically relevant to tidal harmonic analysis, and (B) at a close view about the M_2 constituent frequency. Shading indicates 90% credible intervals of amplitude. Dashed colored lines indicate the Fourier transforms of the residual covariance matrix **R** used in each estimation. Both of these noise spectra assume the same total signal to noise variance ratio of 1.
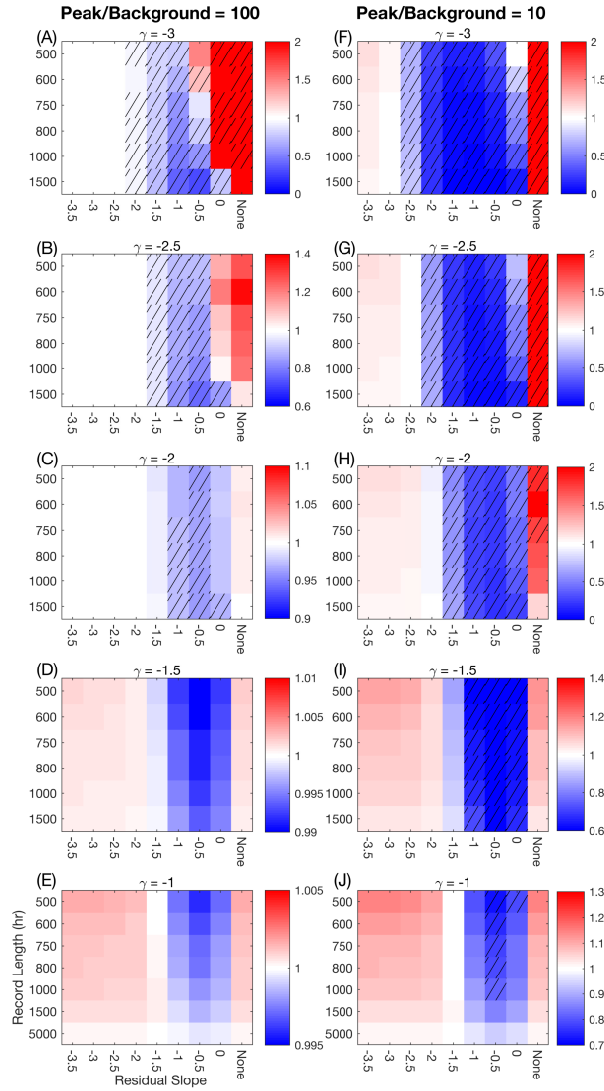
Figure 1.7. Mean ratios of M$_2$ harmonic amplitudes to true amplitudes using red_tide and ordinary least squares. Panels (A-E) correspond to time series with tidal spectral power density 100 times greater than that of the random background process at the M$_2$ frequency. Panels (F-J) correspond to time series with tidal spectral power density 10 times greater than that of the random background process at M$_2$. The random background process to which tides are added has a spectral slope indicated by γ. Time series have 24,000 hourly samples but are only analyzed up to 1500 hours, except for $\gamma = -1$, which was also analyzed at 5000 hours. Each plot corresponds to N Monte Carlo simulations with constant spectra and random phase, where $N = 10,000$ for $\gamma \leq -2$ and $N = 50,000$ for $\gamma > -2$. The spectral slope of the noise term used in red_tide (horizontal axis) and the length of the record (vertical axis) are varied, with noise slope "none" corresponding to the OLS solution. Color bars are centered at unity and scale differently across panels for clarity. Hatched cells indicate combinations of record length and assumed noise spectral slope that resulted in the value 1 (estimated amplitude equal to true amplitude) falling outside the interquartile range for the 10,000 or 50,000 member ensemble.
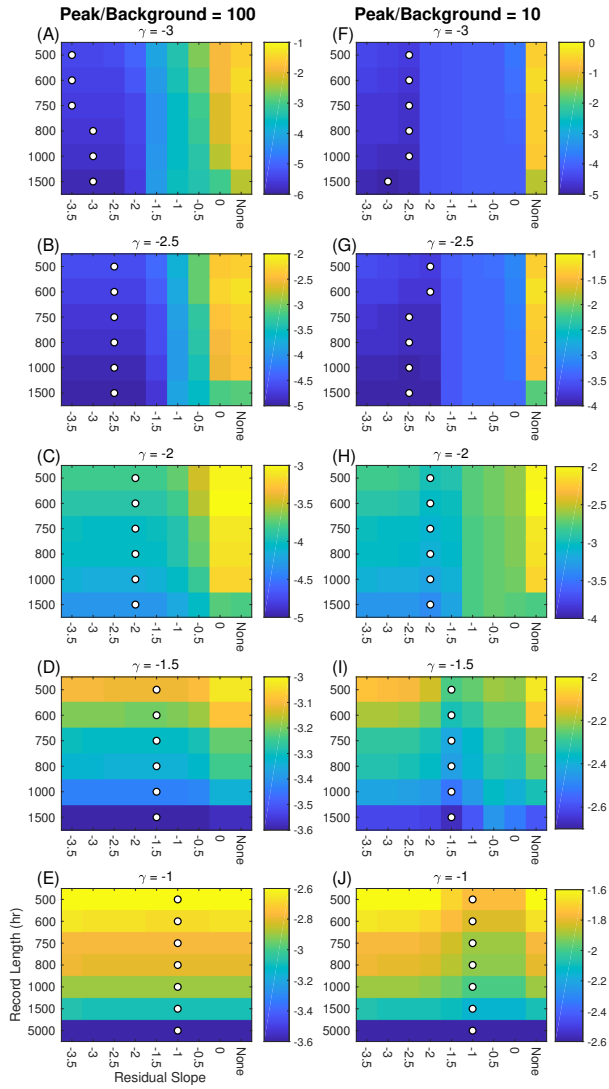
Figure 1.8. Log$_{10}$ of normalized model parameter variance for the M$_2$ constituent calculated from the same artificial data described in Figure 1.7 with the same figure layout. Filled dots indicate the minimum value of its row, and color bars scale differently across panels for clarity.
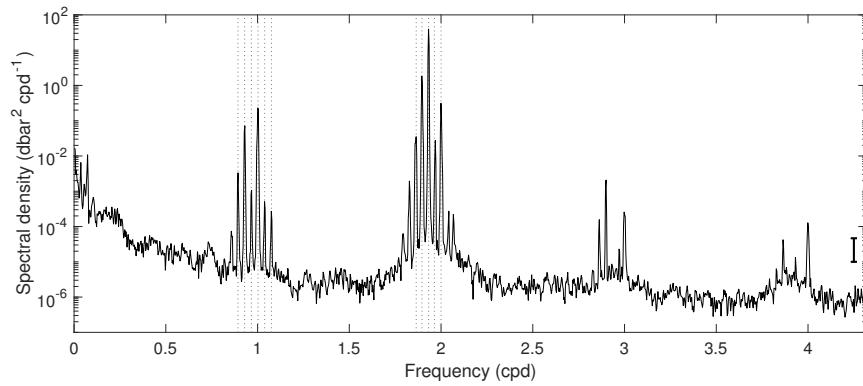
Figure 1.9. Power spectrum of bottom pressure from site 51406 (8.48°S, 125.03°W), part of the DART array. Higher harmonics of the tidal frequencies are present at resolved frequencies higher than roughly 4 cpd (not shown). Vertical dashed gray lines denote prominent tidal constituents, which are the dominant sources of variance in bottom pressure time series. This time series of length 3 years and 6 months (Feb. 12, 2011 to Sept. 6, 2014), sampled every 15 seconds, is averaged over 1 hour intervals, with the power spectrum estimated using Welch's method. The 95% confidence intervals are indicated by the error bar in the lower right, plotted at frequency 4.25 cpd. This corresponds to a ratio valid at all magnitudes of spectral density.
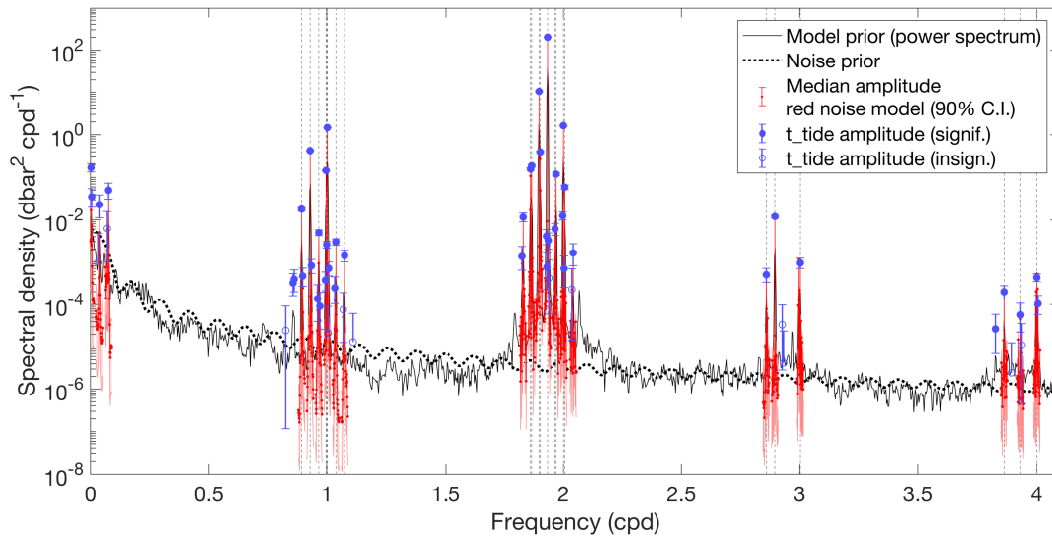
Figure 1.10. Diagonal of model parameter covariance matrix (linear interpolated power spectrum, black solid line), harmonic amplitudes calculated in red_tide using a red noise covariance (spectral slope = -3/2 to empirically match observed spectrum), and results using t_tide. Shading indicates 90% credible intervals. Select diurnal and semidiurnal tidal frequencies are indicated by dashed gray vertical lines. The t_tide amplitudes are from constituents automatically chosen by the package, with amplitudes considered by t_tide to be "significant" filled in, otherwise unfilled, and with 90% confidence bars. All values are in units of spectral density in order to compare quantities defined on different frequency intervals.
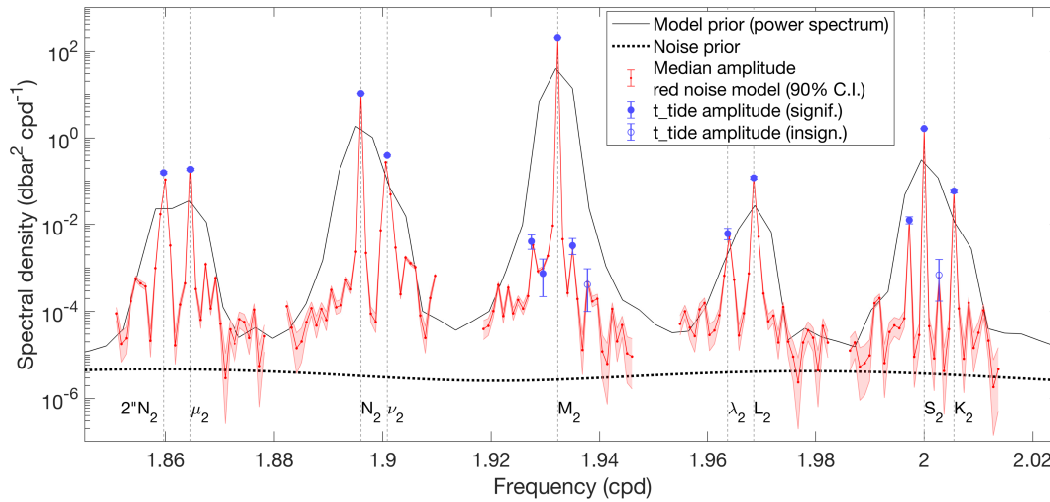
Figure 1.11. Results from Figure 1.10 focused on the semidiurnal band. Prominent tidal constituent frequencies, including those not resolved in the power spectrum, are indicated by vertical dotted lines and labeled according to their conventional Darwin symbols.
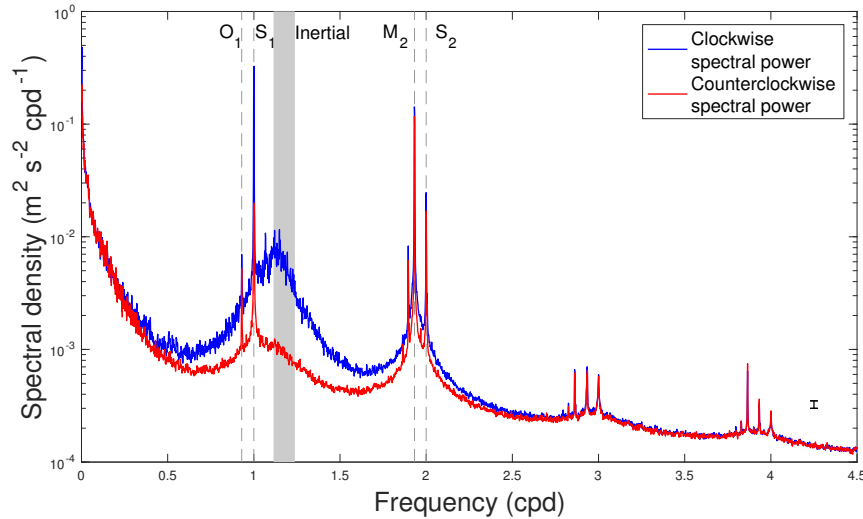
Figure 1.12. Rotary power spectrum of surface current **u** from HFR stations along the central California coast, regionally averaged over 1191 grid points. Time series of length 9 years and 3 months with at least 50% data coverage are sampled hourly. Dashed gray lines denote, in order of increasing frequency, the O_1, S_1, M_2, and S_2 constituents, while the gray band denotes the range of inertial frequencies across the latitudes spanned by the data. Higher harmonics of the tidal frequencies are present at resolved frequencies higher than those shown. The 95% confidence ratio is indicated by the error bar in the lower right, plotted at frequency 4.25 cpd. This corresponds to a ratio valid at all magnitudes of spectral density.

51
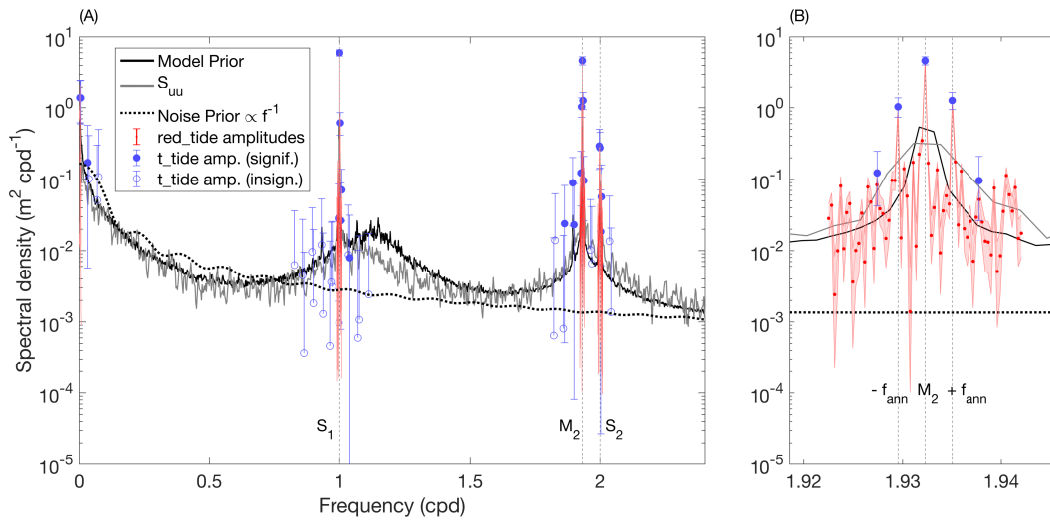
Figure 1.13. Harmonic amplitudes (normalized to units of the power spectrum) of a high-frequency radar time series (zonal component u), with the time series power spectrum S_{uu} and domain-averaged power spectrum (used in **P**) shown for comparison. Model coefficient uncertainty (90% credible interval) is represented by shaded intervals. The covariance used to build **R** is truncated at 300 hours time lag and is built from a spectrum of spectral slope -1 (dotted line). Amplitudes from t_tide, both significant and non-significant, are included for comparison. (B) Close-up of the M_2 band centered at the principal lunar semidiurnal frequency. The vertical dashed lines on either side of the central M_2 line are spaced by the annual frequency

Chapter 2

Characterizing Tides in the California Current System from High-Frequency Radar

Abstract

Over nine years of hourly surface current data from high-frequency radar (HFR) off the US West Coast are analyzed using a Bayesian least-squares fit for tidal components. The high spatial resolution and geographic extent of HFR data allows us to assess the spatial structure of the non-phase-locked component of the tide, while the duration and sampling rate allow the resolution in the frequency domain of discrete tidal lines corresponding to well-known constituents and the broadband elevated continuum near tidal frequencies resulting from amplitude and phase modulation of the tides, known as cusps. The FES2014 tide model is used to remove the barotropic component of tidal surface currents in order to evaluate its contribution to the fraction of phase-locked variance and spatial structure. The mean time scale of modulation is 243 days for the M_2 constituent and 181 days for S_2, with overlap in their range of values. These constituents' modulated amplitudes are significantly correlated in several regions, suggesting forcing by similar processes. Within the frequency band $M_2 \pm 5$ cycles per year, an average of 48% of energy is not at the phase-locked frequency. When we remove the barotropic model, this increases to 64%. In both cases, there is substantial regional variability. This indicates that a large fraction of tidal energy is not easily removable for the purposes of interpreting satellite altimetry data. The spatial autocorrelation of the non-phase-locked variance fraction drops to zero by 150 km, comparable to the width of the swath of the recently launched Surface Water and Ocean Topography (SWOT) altimeter.

2.1 Introduction

Barotropic tidal currents are known to dissipate in shallow seas via bottom boundary layer drag and in the deep ocean via scattering over topography and subsequent conversion to internal tides (Baines, 1982; W. H. Munk & Wunsch, 1998; Egbert & Ray, 2000; R. D. Ray & Cartwright, 2001; Garrett & Kunze, 2007). The deep ocean conversion of barotropic to baroclinic (internal) tides accounts for 25-30% of the total barotropic energy loss and ultimately may provide roughly half of the mixing energy required to maintain the global overturning circulation (Egbert & Ray, 2000). Thus the internal tide plays an important role in the world ocean. Surface signatures of internal tides have long been detected from satellite altimetry (e.g. R. D. Ray & Mitchum, 1996), analysis of which using traditional tidal harmonic methods relies on years of observations to extract highly repeatable phase-locked signals (R. D. Ray & Zaron, 2011). However, the total energy of the internal tide has a substantial component that is not phase-locked to the astronomical forcing terms due to modulation and scattering of propagating internal tides by mesoscale processes (Mitchum & Chiswell, 2000; Chiswell, 2002; Rainville & Pinkel, 2006). Tidal signals have long been known to alias to well-defined low frequencies when sampled by satellites, the orbital periods of which are longer than tidal periods (Parke et al., 1987), and models to account for this phenomenon have been implemented R. Ray (1993). Unlike the large-scale, predictable barotropic tide, this non-phase-locked baroclinic component is difficult to remove from altimeter data and therefore poses a challenge to present and future high-spatial-resolution satellite missions, as it can contaminate mesoscale and submesoscale signals (Chavanne & Klein, 2010; Savage, Arbic, Alford, et al., 2017; Zaron & Ray, 2018). Due to its relevance to altimetry corrections and ocean mixing, the non-phase-locked tide has recently been a widely investigated topic, both observationally and computationally (e.g. R. D. Ray & Zaron, 2011; Ponte & Klein, 2015; Zaron, 2015; Dunphy et al., 2017; Huang et al., 2018; Savage et al., 2020), with the non-phase-locked tide estimated and mapped at a global scale, both from satellite altimetry (Zaron, 2017) and from output from HYCOM, a global ocean general circulation model (Shriver

et al., 2014; Savage, Arbic, Richman, et al., 2017; Nelson et al., 2019). Although model-guided estimation of this tidal component has been demonstrated at a regional scale (Egbert & Erofeeva, 2021), a regional study of the spatial structure of the non-phase-locked tide using observations with high spatial and temporal resolution has not been conducted to our knowledge. Fully characterizing this component observationally can help validate numerical studies and ultimately provide critical information needed to assess tidal variability in newer high-resolution satellite missions, such as the recently-launched Surface Water and Ocean Topography (SWOT) mission, which aims to resolve sea surface height at 15-30 km wavelengths (Morrow et al., 2019).

The goals of this study are to use high-frequency radar (HFR) measurements of surface currents to quantify the non-phase-locked tide. HFR-derived surface currents are a valuable source of spatially high-resolution, decade-long hourly observations (Terrill et al., 2006), and could therefore prove useful for characterizing non-phase-locked tidal variability and spatial scale. The U.S. Integrated Ocean Observing System (IOOS) High Frequency Radar Network (HFRNet) is one of many national networks that comprise the Global HFR Network (Roarty et al., 2019). HFRNet differentiates observation sites by region, with the United States West Coast (USWC) array providing over a decade of publicly available data spanning the Pacific Coast of North America from the U.S.-Mexico border to the mouth of the Columbia River. HFR surface currents have been used to study coastal jets, eddies, near-inertial motions, and tidal currents, from the submesoscale to the mesoscale (e.g. Erofeeva et al., 2003; Kurapov et al., 2003; Kosro, 2005; Chavanne & Klein, 2010; Kim et al., 2010, 2011; Kim & Kosro, 2013); however, these studies did not perform detailed analysis of the non-phase-locked tide.

In order to characterize the non-phase-locked tide in the California Current System (CCS), we examine USWC HFR data using the red_tide harmonic analysis package, which uses a Bayesian approach that accounts for spectrally red background energy and allows for tidal cusps (Kachelein et al., 2022). We then compare our analysis with a barotropic tidal model and further partition the tidal signal into baroclinic and barotropic components in order to investigate the share of non-phase-locked energy that remains after removing the barotropic component,

which is more coherent with astronomical forcing and is typically removed when analyzing altimeter observations (Chavanne & Klein, 2010) or when describing the non-phase-locked tide (e.g. R. D. Ray & Zaron, 2011; J. Wang et al., 2018). We also compare the modulation of the amplitude of M_2 and S_2 currents in order to evaluate the processes that lead to modulation.

This study uses USWC HFRNet observations to quantify the degree to which regional tidal surface currents are non-phase-locked. We build upon previous studies that investigate this problem on a global scale at lower spatial resolution. We begin with an overview of the data and model in this region (Section 2), followed by a brief description of the tidal analysis package and an overview of tidal modulation (Section 3). We then present the results of this analysis, including the spatial structure and amplitude of non-phase-locked tidal currents, the time scales of their modulation, and a comparison of the modulation of the M_2 and S_2 constituents (Section 4), and we end with a discussion of these results (Section 5) and conclusions in the context of future work and the implications of our analysis for satellite altimeter and scatterometer missions (Section 6).

2.1.1 Defining the Non-Phase-Locked Tide

Tides can be subdivided into "phase-locked" and "non-phase-locked" components. The "phase-locked" portion of the tide is usually associated with the barotropic tide, meaning that it depends largely on ocean depth and is highly regular such that it can be readily predicted and removed. The non-phase-locked component of the tide has an amplitude and phase that vary in time due to interactions of internal tides propagating through a varying medium. The interaction of tides with non-tidal processes leads to so-called tidal cusps in the frequency domain (W. H. Munk et al., 1965; Radok et al., 1967; Matte et al., 2013), continuous prominences of spectral power which encompass a range of frequencies centered on tidal forcing frequencies. This feature is common to tidal peaks and is not an artifact of spectral analysis techniques. Though we refer to this variable tide as the non-phase-locked tide, we interpret it in the frequency domain as a non-*frequency*-locked signal at near-tidal, but not exactly tidal, frequencies. This is

because the astronomical forcing is assumed to occur at exactly the known tidal frequencies for the purposes of this study.

The non-phase-locked tidal variance is not a standardized quantity and is often calculated as the difference in spectral power before and after the removal of the phase-locked tide, usually via a least squares fit (Nelson et al., 2019). For example, in their analysis of Topex/Poseidon and Jason satellite altimeter data, R. D. Ray & Zaron (2011) calculated the non-phase-locked tidal energy as the difference between the wavenumber spectra of sea surface height (SSH) along a ground track before and after removal of 12 major tidal constituents via point by point harmonic analysis, within the bands of mode-1 and mode-2 semidiurnal wavenumbers. Similarly, Savage et al. (2020) analyzed mooring data and semi-idealized internal tide model output and calculated the total tidal component by bandpass filtering time series around the M_2 constituent frequency $\omega_{M_2} \pm 0.4$ cpd, which they then separate into phase-locked and non-phase-locked components via least squares fits to the M_2 and S_2 harmonics.

2.2   Data

2.2.1   High-Frequency Radar (HFR)

HFR observations of surface currents are obtained by a land-based network of antennas that measure the radio signal backscattered off of the ocean surface. Surface velocities are inferred using a multiple signal classification (MUSIC) algorithm (Schmidt, 1986; Barrick & Lipa, 1997). The HFR National Network (Terrill et al., 2006; Roarty et al., 2019) provides archived surface ocean velocities mapped to grids at 1 km, 2 km, and 6 km spacing (Fredj et al., 2016). We examine only 6 km gridded data, which are available over nearly the entire continental US West Coast and resolve the $O(10)$ km mode-1 internal semidiurnal tide while providing data from the coast out to 200 km offshore; this is farther than either the 1 km or 2 km gridded products, which also do not provide continuous coverage along the coast.

These gridded fields are produced using an unweighted least-squares fit performed on

radial surface current data. For information about the fitting procedure see Lipa & Barrick (1983); Gurgel (1994). For the 6 km gridded product, the unweighted least squares fit approach is performed on radial data within a 10 km search radius centered on the grid location (Otero, 2021, personal communication), producing zonal and meridional current u and v as well as the zonal and meridional geometric dilution of precision (DOPX and DOPY), which are measures of the uncertainty of the current estimate due to antenna geometry (U.S. Integrated Ocean Observing System, 2016). The horizontal dilution of precision (HDOP) is the sum of DOPX and DOPY and is limited to HDOP ≤ 1.25 for quality control in the publicly available data set we use.

Our study region extends from 32.24°N to 46.43°N and is focused on the California Current System (CCS). For quality control, we require 50% temporal coverage at each location that we analyze. This level of coverage is typically available within 100-200 km from the coast, as shown in Figure 2.1a. Offshore observations are limited by the locations of antennas that comprise the observation system: gridded data quality control requires at least three radial velocity observations from at least two antennas in order to estimate velocity **u** (Otero, 2021, personal communication), as the number of radial velocity observations within a grid point's search radius and their bearing angles influence the dilution of precision (DOP) (Kim et al., 2008). Hourly data are used from 1 January 2012 through 1 April 2021. Figure 2.1b shows the mean flow of the surface current in this domain, which is characterized by a broadly equatorward surface current and offshore advection, consistent with drifter observations of the CCS (Centurioni et al., 2008). Figure 2.1c shows $\sqrt{(\sigma_u^2 + \sigma_v^2)/2}$, the square root of eddy kinetic energy. Regions with the highest values for this quantity typically fall outside our coverage threshold; it is unknown which regions of high total energy are physical or a consequence of poor DOP due to antenna placement, though the fact that high energy areas along the coast are generally limited to regions with low coverage suggests an effect of the observation network, hence our threshold.

2.2.2 FES2014 Barotropic Tidal Model

In order to distinguish the baroclinic tides from coherent barotropic signals, we use the 2014 version of the Finite Element Solution (FES2014) barotropic tidal model (F. H. Lyard et al., 2021) as an estimate of the barotropic component of the total tidal current. This model was designed to provide tidal elevation predictions for use with satellite altimetry data, and it also provides tidal current maps for 34 tidal constituents. FES2014 model constituent frequencies that we fit include the O_1, S_1, K_1, N_2, M_2, S_2, and K_2; however, we only compare HFR and FES2014 at the M_2, as it is the most energetic constituent. The model, which assimilates both tide gauge and satellite sea surface height data, is run on an unstructured grid. This grid improves performance near coasts and shelves. The original model output is gridded onto a 1/16° (~6.9 km) resolution regular grid; for this study, the phase and amplitude of the global model are mapped to the 6 km grid of the HFR data in the CCS region via bilinear interpolation. The current from FES2014 is treated as the barotropic tidal component of the total signal observed by HFR, and is assumed to occur exactly at the tidal forcing frequency. This is justified by the apparent lack of an underlying domain-wide phase coherence at near-tidal frequencies from harmonically decomposed HFR data (not shown) compared to the partially spatially-coherent phase seen at the tidal forcing frequency in HFR. The removal of the modeled barotropic component from harmonically decomposed HFR data is outlined in the appendix.

2.3 Methods

2.3.1 Tidal Harmonic Analysis

Harmonic analysis of tides is frequently performed using ordinary least squares and sinusoidal basis functions (Foreman, 1977; Pawlowicz et al., 2002). Several software packages use modified procedures for the least squares fit, such as iteratively re-weighted least squares (Leffler & Jay, 2009), directly embedding long-period signals (Foreman et al., 2009), or incorporating non-tidal river flow (Matte et al., 2013). Some packages seek to improve and unify various

methods to serve as a general tool (Codiga, 2011).

Surface currents contain signals from several sources, including the tides, wind, and low-frequency variability, and therefore are energetic across a wide range of frequencies. These HFR time series are analyzed using harmonic analysis techniques designed for non-stationary tides in the presence of spectrally colored noise, where the term "noise" is used to describe any non-tidal variability that may complicate the estimation of tidal parameters. Tidal harmonic amplitudes are given by Bayesian maximum posterior estimation assuming Gaussian linear statistics, as described by Kachelein et al. (2022). The appendix provides a brief overview of this method using modified notation following Ide et al. (1997).

2.3.2   Calculating the Non-Phase-Locked Tide

We define the non-phase-locked tide similarly to previous studies but limit our scope to frequencies that are explicitly fit via equation (2.7) (see Appendix for more information). We therefore define the phase-locked component as the harmonic fit at a tidal forcing frequency f_i, the i'th element of $\mathbf{f}$; and the non-phase-locked component as the sum of the other harmonic fits within a narrow frequency band, from f_{i-N} to f_{i+N}, such that the phase-locked fraction of variance can be defined from harmonic coefficients as

$$F_{\mathrm{PL}} = \frac{a_i^2 + b_i^2 + c_i^2 + d_i^2}{\sum_{j=i-N}^{i+N}(a_j^2 + b_j^2 + c_j^2 + d_j^2)}, \tag{2.1}$$

where a and b are respectively the sine and cosine harmonic coefficients that comprise $\hat{\mathbf{x}}_u$, c and d are the same for $\hat{\mathbf{x}}_v$, and i is the index of the tidal frequency in question. The non-phase-locked fraction of variance is simply $F_{\mathrm{NPL}} = 1 - F_{\mathrm{PL}}$. A similar approach for calculating both phase-locked and non-phase-locked tidal components directly from amplitudes given by a least squares fit has been used previously (e.g. Dushaw et al., 1995). However, energy at frequencies that are not explicitly included in the tidal fitting process is missed in this definition. In order to identify the full contribution of the non-phase-locked tide, the fitted harmonics should span

enough frequencies in a band centered on the central tidal frequency such that most of each tidal constituent's total energy is included in the fit. The width of a cusp is related to the spectrum of the process that non-linearly forces it (Matte et al., 2013), so fitting more nearby frequencies (widening the band) will allow more near-tidal variance to be captured. Therefore we fit the central frequency ± 5 cpy for both M_2 and S_2, which captures modulation of these constituents by processes with periods of 0.2 years and longer.

2.3.3 Tidal Modulation

Following Chiswell (2002), a tidally-driven, amplitude- and phase-modulated process $a(t)$ can be written

$$a(t) = A(t)\cos[2\pi f_{\text{tidal}}t + \phi(t)], \tag{2.2}$$

where f_{tidal} is the ordinary frequency of some tidal constituent, $A(t)$ is the time-varying amplitude, and $\phi(t)$ is the time varying phase. The frequency-domain transform of a time series with a modulated tidal component contains information about the structure of the modulation. When phase is modeled as a first-order autoregressive process with some time scale $1/\lambda$, the power spectrum of $a(t)$, assuming that amplitude is constant in time, i.e. $A(t) = A_0$, is modeled by a Lorentzian (or Cauchy) distribution plus a mean (Zaron, 2022):

$$S_a(f) = S_0 + \frac{\alpha}{\pi\lambda\left[1 + \left(\frac{f - f_{\text{tidal}}}{\lambda}\right)^2\right]}, \tag{2.3}$$

where α and λ are parameters that can be estimated via a non-linear least squares fit. If $A(t)$ is modeled as a first-order autoregressive process, $S_a(f)$ is also of this form via the convolution of the spectrum of a sinusoidal tide and the spectrum of $A(t)$ (von Storch & Zwiers, 2003).

We assume that all observed modulated tidal signals are baroclinic, as barotropic tides are negligibly modulated (Colosi & Munk, 2006). Tidal energy becomes non-stationary via both amplitude and phase modulation of tidal constituents. Mitchum & Chiswell (2000), for example,

found that the low-frequency variability observed in both the amplitude and phase of the M_2 constituent is correlated with low-frequency sea level variations, which in turn are correlated with stratification variations, the proposed modulation mechanism. Therefore, information about the processes that modulate tidal constituents should be recoverable from the fitted tidal harmonic coefficients. We focus on amplitude modulation by finding the time-varying amplitude (*envelope*) that modulates the higher-frequency oscillating tide (*carrier wave*).

Figure 2.2 shows a simplified example of tidal modulation in both the time and frequency domains. The top panel (a) shows the apparent phase and amplitude modulation of one tidal constituent by the addition of a different constituent of a similar frequency (here M_2 and S_2). This fortnightly interference pattern simulates the well-known and ubiquitous spring-neap cycle at the lunar synodic fortnightly frequency (Parker, 2007). The power spectrum of the total signal (blue, right) is comprised of two peaks at f_{M_2} and f_{S_2} while that of the recovered non-sinusoidal modulating envelope (red) has a fundamental frequency at $f_{MSf} = f_{S_2} - f_{M_2}$ with harmonics of decreasing power at integer multiples of f_{MSf}, including frequencies at which the low amplitude peaks are not plotted. The middle panels (b) show a low-frequency sinusoidal envelope (red) modulating the amplitude of a high-frequency sinusoidal carrier signal (blue) at the M_2 frequency. The frequency of the modulating envelope, f_{mod}, in (b) is set to the spring-neap frequency f_{MSf} for the sake of comparison and visual clarity (amplitude modulation as observed in HFR data is most evident on an annual cycle). In the frequency domain, the power spectrum of the envelope is a single peak at the modulating frequency f_{mod}, while the modulated time series has a power spectrum of three peaks: a higher-amplitude central peak at the carrier frequency and two lower-amplitude peaks at $f_{M_2} \pm f_{mod}$. The bottom panels (c) show the general case of amplitude modulation by a broadband process, here an AR(1) autoregressive process with coefficient 0.95 (for more information, see Chapters 10.3 and 11.2.3 of von Storch & Zwiers, 2003). The broadband envelope has a power spectrum that is centered around the carrier frequency in the power spectrum of the modulated time series. The gray power spectrum is calculated from the data as the square of the discrete Fourier transform, while the smooth blue spectrum is given

by the convolution of the theoretical spectrum of the AR(1) process (red) with that of the tidal carrier wave. In both cases of amplitude modulation (b-c), the power spectrum of the envelope is centered around the line spectrum of the sinusoidal carrier wave, while in the case of adding constituents (a), there is apparent phase and amplitude modulation that complicate the simple frequency domain convolution of envelope and carrier.

We fit the HFR data to each prominent tidal constituent: O_1, S_1, N_2, M_2, and S_2. Frequencies are fitted in a narrow band centered on the tidal frequency, spanning ± 5 cpy, and spaced by $\Delta f = 1/9$ cpy. The modulated current for any of these constituents can be viewed as a slowly-varying ellipse, whose inclination, semi-minor axis, and semi-major axis are approximately constant across a single tidal period but evolve over time. We estimate the evolution of the semi-major axis length of the fitted vector surface currents at the two major semidiurnal constituents, M_2 and S_2, in order to compare the modulation of these two independently forced constituents in a way that takes into account all four vector components. Within the latitudes of the domain, these are the major propagating constituents that we expect to be modulated by non-tidal processes, and because they are at similar frequencies, we expect them to be modulated similarly.

The semi-major axis is a physically relevant quantity that describes the maximum tidal current amplitude over a tidal period regardless of the orientation and eccentricity of the current ellipse. For a single frequency component, this quantity is constant and can be calculated from the four total harmonic coefficients of u and v at that frequency, but for the evolving ellipse, which contains energy at multiple frequencies, we numerically estimate the semi-major axis length as the maximum value of $\sqrt{u_{\text{band}}^2 + v_{\text{band}}^2}$ over each integer ceiling of the tidal period in hours (i.e. 13 hour intervals for M_2 and 12 hour intervals for S_2 in order to ensure that the maximum in each cycle is sampled). The time series of this value is taken as the "semi-major axis envelope". In constructing the time series for each modulated tidal component (u_{M_2}, v_{M_2}, u_{S_2}, and v_{S_2}), we multiply the harmonic coefficients by a Hanning window 10 cpy wide and centered on the tidal frequency in order to reduce spectral leakage. The Pearson correlation coefficient of the

64

envelopes of these two constituents is then calculated at each location in order to estimate the extent to which they are modulated by the same processes.

To estimate the time scale λ^{-1} of tidal modulation related to non-annual, low-frequency processes, whether of phase or amplitude, we fit harmonic amplitudes directly to (2.3) for M_2, S_2, and O_1 within windows of $\pm$ 5 cpy using the iterative Gauss-Newton algorithm (Bishop et al., 1985). Energetic peaks are manually removed so that only the cusp is fit: tidal peaks, peaks of the annual modulation, and amplitudes at frequencies corresponding to modulation of 3 years or longer typically prevent the least-squares method from converging to curves with reasonable parameters. The mean and uncertainty of the time scale is estimated from 100 cusp widths obtained by fitting the average of harmonic amplitudes from each percentile of cusp-integrated energy, which yields 100 values for λ^{-1}. If the fit to a percentile-averaged cusp does not converge to a reasonable value, for example when the harmonic amplitudes do not average to a cusp shape, that fit is discarded before estimating the mean time scale and its standard deviation.

2.4   Results

2.4.1   Spectral Properties of Tidal Currents

The rotary power spectral density for velocity, $S_{\rm rot}$, is shown in Figure 2.3, calculated at each location and averaged within latitudinal bands of fixed width (3.2° $\approx$ 356 km). Both plots are on a linear frequency axis and logarithmic spectral power axis. Spectra are distinguished by latitude because the inertial frequency varies from approximately 1 cpd to 1.5 cpd over the 1500 km meridional span of the study domain. The numerical values on the vertical axis correspond only to the spectrum of data in the 32.0°N-35.2°N range (blue); for clear differentiation, each successive spectrum is scaled by a factor of ten from its lower neighbor. Away from tidal and near-inertial frequencies, surface currents in this region have a spectral slope of approximately -1, i.e. $S \propto \text{frequency}^{-1}$, indicated by the black curve. This slope holds at frequencies less than about 9 cpd, above which the spectra behave like white noise with spectral power orders of magnitude

lower than the processes of interest.

The principal tidal peaks, clustered around 1 cpd and 2 cpd and denoted by vertical dashed lines, are prominent in all latitude bands. High-frequency, low-amplitude overtides (frequency > 2 cpd) are also detected in surface currents, up to 7 cpd (not shown). This suggests that the observation network has a noise floor low enough to allow for the detection of signals that are orders of magnitude lower in power than the processes of primary interest to this study.

The center of the broad but distinct peak of the near-inertial response increases in frequency with increasing latitude as expected (note that negative frequencies correspond to clockwise polarization). Additionally, the near-inertial peak becomes narrower and more energetic with increasing latitude, as evidenced in the right panel of Figure 2.3. These spectra are directly comparable to the analysis of Kim et al. (2011), who examined HFR data in the same region when the observational network was less than three years old and who resolved the same features (wind-driven and tidal peaks, including overtides, and a spectral slope of -1) across different grid resolutions.

At all latitudes, the semidiurnal tidal band exhibits a prominent cusp. The diurnal constituents, however, are flanked by a smaller cusp than the semidiurnal peaks are: this is most evident in the highest latitude spectrum (44.8°N-48.0°N), where the broad peak of the near-inertial response has minimal overlap with the diurnal band, shown up close in Figure 2.3b. At the latitudes considered here, semidiurnal internal waves can internally propagate but diurnal ones cannot propagate. The difference in cusp prominence for the superinertial semidiurnal versus subinertial diurnal signals is consistent with the internal semidiurnal wave field being modulated by low-frequency changes in stratification that are not expected to modulate the diurnal tide.

2.4.2   Harmonic Decomposition of Tidal Currents

Figure 2.4 shows the domain-wide average of the harmonic amplitude, defined as the sum of the squares of the sine and cosine coefficients, at 6 frequency bands (low-frequency, O_1 lunar

66

diurnal, S_1 solar diurnal, N_2 lunar semidiurnal, M_2 lunar semidiurnal, and S_2 solar semidiurnal). These amplitudes are normalized by the non-constant frequency step Δf in order to obtain units of spectral power that can be compared despite differing values of Δf. Harmonic amplitudes at near-inertial frequencies are not shown in this study. Low-frequency signals, in particular the annual cycle S_a, are energetic, prominent, and predominantly meridional in direction (distinction not shown), possibly due to the seasonally reversing surface flow near the coast of the CCS (Todd et al., 2011) and other low-frequency variability of this current, which is predominantly parallel to the coast and therefore predominantly meridional in direction. The meridional component v similarly dominates at most higher frequencies. The prominent peak at 1 cpd is non-tidal and apparently driven by diurnal winds, as the true S_1 tide is driven primarily by diurnal atmospheric loading and is known to be much smaller than the primary gravitational tides (R. D. Ray & Egbert, 2004).

At semidiurnal frequencies, which are superinertial at these latitudes and therefore are expected to have a substantial baroclinic component, the subtraction of the FES2014 model typically reduces the total variance as observed by HFR. At M_2, the median fraction of remaining variance is 0.46, with 75.2% of locations exhibiting a decrease in variance. For S_2, the median is 0.54, with 76.2% of locations exhibiting a decrease. For N_2, these values are 0.52 and 70.2%. At diurnal frequencies, which are subinertial, subtraction of the FES2014 model does not typically reduce the total variance as observed by HFR. For O_1, the median fraction after model subtraction is 1.16, with 44.4% of locations reduced in variance, while for K_1 these values are 1.38 and 31.1%. The S_1 wind-driven current is much stronger than the modeled S_1 barotropic tide, and thus the median variance fraction after model subtraction is 0.99, with a narrow range of variability (90% of locations fall between 0.85 and 1.03).

2.4.3   Phase Structure

The phases of the M_2 harmonic of both u and v are shown in Figure 2.5. Panels (a,d) show the phase from harmonically decomposed HFR data, panels (b,e) show the phase of currents

67

given by the FES2014 barotropic tide model, and panels (c,f) show the phase of the M_2 harmonic of HFR after removal of the barotropic model. The spatial structure of the phase of decomposed HFR data exhibits domain-wide (~1000 km-scale) coherence like the barotropic model, as well as smaller 50 to 100 km-scale structure which is not present in the model. The meridional current v is particularly coherent across the domain (Figure 2.5d), with an angular mean of 30° and 80% of grid points falling within ±45° of this mean. At all other frequencies in the M_2 band, including the frequencies corresponding to annual modulation, phases are distributed more uniformly (not shown) and therefore lack the partial domain-wide coherence observed at exactly the M_2 frequency.

After removal of the modeled barotropic component (Figure 2.5c,f), the apparent underlying domain-wide coherence largely disappears, leaving behind the smaller scale phase structure that is not visible in the barotropic tidal current and is suggestive of internal propagation. In the Southern California Bight, the phase generally varies across smaller distances than in the topographically simpler regions north of Point Conception (approximately 34.5°N).

2.4.4 Non-phase-locked Energy

Figure 2.6 shows the square root of the eddy kinetic energy of the M_2 tidal current (2.6a,c) and the fraction of energy that is considered phase-locked (2.6b,d). Panels 2.6a and 2.6b display these quantities from harmonically decomposed HFR data only, while panels 2.6c and 2.6d show these quantities after subtracting the barotropic model output. A general reduction in M_2 current energy is observed after barotropic removal, with a median variance reduction of 54%. The sites with the most energetic M_2 currents from HFR are at the mouth of the Columbia River, off Cape Mendocino, at the mouth of San Francisco Bay, and in the Southern California Bight, especially over the Santa Cruz Basin. These sites remain energetic after barotropic removal.

Histogram representations of the quantities in panels 2.6b and 2.6d are shown in Figure 2.7. The median fraction of phase-locked current variance from HFR alone is 53% with an interquartile range between 38% and 66%. After barotropic removal, the median is 33% with an interquartile

range between 22% and 46%, indicating that the barotropic component represents a significant fraction of the phase-locked variance.

To evaluate the dominance of the barotropic component, we define the "total variance" as the sum of the variance of the modeled barotropic current and the variance of the difference between observations and the model:

$$F_{\mathrm{BT}} = \frac{\mathrm{var}(u_{\mathrm{FES}}) + \mathrm{var}(v_{\mathrm{FES}})}{\mathrm{var}(u_{\mathrm{FES}}) + \mathrm{var}(v_{\mathrm{FES}}) + \mathrm{var}(u_{\mathrm{HFR}} - u_{\mathrm{FES}}) + \mathrm{var}(v_{\mathrm{HFR}} - v_{\mathrm{FES}})}. \tag{2.4}$$

This is equivalent to one minus the baro*clinic* fraction of total variance.

Figure 2.8a shows the modeled barotropic variance as a fraction of this total variance, with the barotropic component dominating near the Oregon Coast and at locations where the observations are weakly energetic. Shaded regions indicate where the denominator of equation (2.4) is below the 20th percentile (7.05×10^{-4} m^{2}s^{-2}). Figure 2.8b shows the fraction of phase-locked current variance within the $f_{\mathrm{M_2}} \pm 5$ cpy band that is attributed to the barotropic model, which is considered to be completely phase-locked. The quantity in this panel equals that of Figure 2.6b minus that of Figure 2.6d. Alternatively, this can be viewed as the change in non-phase-locked variance fraction after subtracting the model. Regions with negative values have greater current variance after the removal of the barotropic component. A median of +0.17 is attributed to the model, with the 10% most affected locations attributing ≥ 0.45 fraction of variance to the model. A general increase in non-phase-locked energy fraction is observed after barotropic removal. Because observed and modeled currents are vector quantities, however, their subtraction does not necessarily result in a reduction of tidal amplitude. For example, regions with low observed tidal variability may have opposite-pointing barotropic and baroclinic contributions that sum to a low-amplitude signal, with an increase in non-phase-locked energy fraction after vector subtraction of the barotropic component. A *decrease* in non-phase-locked variance fraction occurs at 25% of locations, predominantly where the angle between the observations and model is greater than 45° and thus where subtraction of the barotropic model does not substantially

reduce the magnitude of the residual (not shown). These regions of non-phase-locked fraction decrease are also disproportionately where the observed phase-locked M_2 variance (i.e. peak only) is less than 3.63×10^{-4} m^{2}s^{-2}, the 25th percentile of that quantity (not shown).

The spatial autocorrelation of the phase-locked variance fraction is shown in Figure 2.9, which shows the different spatial scales at which this quantity varies for each main constituent. The correlation is averaged in 5 km-wide concentric bands except for the second bin, which is 10 km wide in order to obtain sufficient realizations. Spatial lags extend to a maximum of 150 km separation between observation locations. The decrease in correlation as a function of distance differs by constituent: the distance at which the correlation of this quantity drops to 0.5 is 45 km for O_1, 101 km for S_1, 24 km for M_2, and 29 km for S_2, based on a linear interpolation of the discrete correlation. For M_2, the correlation of the fraction of phase-locked variance after barotropic removal is plotted in red, with a distance of 15 km at which correlation drops to 0.5. Error estimates are omitted due to reasonable convergence from averaging over thousands of realizations, from 7270 to 1.77×10^5 depending on the bin. Correlation does not meaningfully differ when averaged over 10 km-wide bins: when interpolated to 1 km intervals, the difference in correlation between the two bin choices has a mean of zero and a standard deviation less than 0.007 for each tidal constituent. The diurnal O_1 and semidiurnal M_2 and S_2 tidal constituents all exhibit autocorrelations that decrease to zero at distances less than 150 km, while the wind-driven S_1 only drops to 0.5 at this distance, indicating that both propagating (superinertial, here semidiurnal) and non-propagating (subinertial, here diurnal) tides in this region have smaller spatial scales of non-phase-locked variability than diurnal S_1 surface currents. The autocorrelation of non-phase-locked variance fraction is reduced for M_2 at every non-zero lag after barotropic model removal; this may be due to the larger scales at which the barotropic phase varies (see Figure 2.5), the removal of which may intensify the effect of the barotropic residual's smaller scale.

2.4.5 Tidal Envelopes

The map in Figure 2.10 shows the correlation between the M_2 and S_2 semi-major axis envelopes, which is mostly positive: for a 90% significance level, 38.2% of locations have a significant positive correlation, 60.0% fall within the non-significance range, and 1.9% have a significant negative correlation (or 8.4% of all negative values). Striped areas indicate regions where the absolute value of the correlation coefficient (correlation at zero lag) is less than the 90% significance level (Press et al., 1988) given the estimated number of degrees of freedom at each location based on the characteristic timescale of the M_2 and S_2 envelopes (for estimating timescales, see Chapter 17, von Storch & Zwiers, 2003). The significance level is spatially inhomogeneous, with a mean of 0.267 and standard deviation of 0.048. Regions with energetic semidiurnal tides, as shown by Figure 2.6a, exhibit significant positive correlation between the M_2 and S_2 envelopes, including the Southern California Bite, San Francisco Bay, Cape Mendocino, and the mouth of the Columbia River. Five locations from these regions with significant cross-correlation at zero lag are selected and labeled (1-5), with the time series of their M_2 and S_2 semi-major axis envelopes plotted in Figure 2.10. The envelopes at each location vary at different time scales. Location (1) near the mouth of the Columbia River is strongly annually modulated: both constituents' envelopes are consistently stronger in summer, with currents roughly three times greater in magnitude than in winter. Location (2) is near Cape Mendocino and the Mendocino Ridge, and the constituents at this location vary at long time scales, with greater magnitude in the first half of the record than in the second half, though the summer magnitude of S_2 in 2018-2020 increases while that of M_2 does not. Location (3) near the entrance to the San Francisco Bay has a less heavily modulated M_2 envelope, while the S_2 envelope is seemingly strongly semiannually modulated. This is caused by the linear addition of the K_2 tidal constituent ($f_{K_2} = f_{S_2} + f_{S_{sa}}$), which is prominent at this location, and does not appear to be due to amplitude modulation of the S_2 constituent: the amplitudes of the harmonic constituents in the semidiurnal band (not shown) are only energetic at f_{K_2} and not $f_{S_2} - f_{S_{sa}}$.

Location (4), between Santa Cruz Island and the mainland, also exhibits a strong K_2 peak, as well as annual modulation of both constituents. Location (5) is farther offshore, south of Santa Cruz and Santa Rosa Islands, and has a less dominant K_2 relative to S_2. The correlation coefficient of 0.77 observed here is greater than those of 99% of locations evaluated. Location (5) is tidally energetic (for M_2, $\sqrt{(\sigma_u + \sigma_v)/2} = 0.05913$ m s^{-1}) with high coverage (94.1%). Overall, the variance of the K_2 constituent is distributed similarly to those of M_2 and S_2 (not shown).

2.4.6 Modulation Time Scale

The time scale λ^{-1} is given by fitting harmonic amplitudes to equation (2.3). Locations were sorted into percentile bins by total energy of the tidal band of interest (including the central peak), averaged by percentile, and fitted at non-peak frequencies to obtain a range of estimates of λ^{-1}. Values for time scales are similar whether the non-linear parameter of the fit is λ or λ^{-1} (values shown here were calculated from the former). The mean time scales obtained by this method are $\lambda_{O_1}^{-1} = 218$ days (standard deviation 37 days, 90% falling between 172 and 275 days), $\lambda_{M_2}^{-1} = 243$ days (standard deviation 31 days, 90% between 195 and 288 days), and $\lambda_{S_2}^{-1} = 181$ days (standard deviation 24 days, 90% between 153 and 218 days). Of the 100 percentile-sorted fits for each of the cusps, two for O_1 and one for M_2 were omitted from averaging due to a lack of convergence to reasonable values. Least squares fits to the median harmonic amplitude are shown in red in Figure 2.4.

2.5 Discussion

In this study, we sought to partition the variability of tidally-driven surface currents between phase-locked and non-phase-locked components. Our analysis of HFR data in the CCS indicates an inhomogeneous distribution of both phase-locked and non-phase-locked variability of the two prominent semidiurnal constituents M_2 and S_2. Most of our analysis is focused on M_2 due to its prominence, while diurnal tides, which are non-propagating, are not investigated as extensively. The FES2014 barotropic tide model is used to remove the barotropic component

of the M_2 tidal current, which accounts for a substantial share of the observed phase-locked variance. The model appears to reasonably remove the domain-wide coherent component of M_2 tidal currents, leaving behind a more uniform distribution of tidal phase suggestive of the surface expression of baroclinic tides, the wavelengths of which are much smaller than the domain. Before removal, a median of 53% (interquartile range between 38% and 66%) of M_2 current variance is phase-locked, compared to a median of 33% (IQR between 22% and 47%) after removal. Overall, 75% of grid locations show a decrease in M_2 phase-locked variance fraction after model subtraction and the distance at which the phase-locked variance fraction is 50% correlated drops from 24 km to 15 km. Our analysis yields a *mean* of 48% *non*-phase-locked variance before removal and 64% after. These are both higher than the 44% near-global total non-phase-locked fraction calculated by Zaron (2017) from SSH, which was considered a lower bound due to the analysis of only mode-1 internal tides as estimated from wavenumber spectra. The results of that study are not directly comparable to ours, as the two use a different definition for non-phase-locked variance fraction. Further differences between our analysis and that of Zaron (2017) include the examination of surface currents as opposed to surface height, a domain that is limited to a single coastal region, and our distinction by constituent but not by internal wave mode. Additionally, the high spatiotemporal resolution of HFR data may limit the value of comparisons with analyses of global altimeter data without carefully accounting for these differences.

The analysis in this study provides a conservative estimate of the non-phase-locked fraction of tidal current variance, as our harmonic analysis is limited to the tidal forcing frequency ± 5 cpy (see Figure 2.4). This is a narrow band in the frequency domain relative to the width of the total semidiurnal cusp evident from power spectra in Figure 2.3. We chose this band because it allows for a clear distinction between energy associated with separate tidal peaks while still fitting to near-tidal energy that corresponds to amplitude and phase modulations at time scales of a few months.

At both diurnal and semidiurnal frequencies, tidal peaks are well-defined and are at least

one order of magnitude more energetic than at near-tidal frequencies. All tidal bands exhibit tidal cusps, though those of the low-amplitude O_1 and N_2 are less pronounced. These cusps are similar in form for all major tidal bands at this frequency resolution and narrow bandwidth. This contrasts with the difference in cusp size over a wider range of frequencies between diurnal and semidiurnal bands in the power spectra. Secondary peaks rise above the cusps at ± 1 cpy around the central tidal peak. Some of these peaks are constituents of the astronomical forcing: the K_1 tide, for example, is a major tidal constituent with $f_{K_1} = f_{S_1} + f_{S_a}$. For M_2, there is a weak solar gravitational effect on the moon's orbit that results in seasonal perturbations impacting M_2, but these perturbations are very small: processes driven by solar radiation with a strong seasonal cycle, including stratification, wind, and sea ice coverage, however, are much more important to the modulation of M_2 (Müller et al., 2014).

Regions with energetic M_2 tidal signals, including known internal tide generation sites like the Mendocino Escarpment (Althaus et al., 2003) and Santa Cruz Basin (Buijsman et al., 2012), exhibit more strongly phase-locked signals. This is expected for locally generated internal tides, as such waves have not propagated far from their generation sites and become more incoherent through interaction with varying stratification and currents. Remotely generated internal tides propagating from outside the observed region are expected to lose coherence with forcing between their generation and observation sites and thus have a similar phase-locked energy fraction throughout the study region. The slowly evolving envelopes of the M_2 and S_2 current semi-major axes exhibit significant correlation coefficients (cross-correlation at zero lag) at several locations in the domain, including near local generation sites (Figure 2.10). The envelopes are assumed to contain information about external modulation, as the forcing mechanisms of these constituents (lunar and solar, respectively) are independent and highly temporally coherent and therefore should not impart modulation, except for long-period modulation like that of the 18.6-year lunar nodal cycle (Doodson & Lamb, 1924; Cherniawsky et al., 2010). The fact that high cross-correlation tends to be observed near generation sites, where modulation related to propagation through a varying medium is expected to be weakest, may be due to the higher

signal-to-noise ratio near generation sites allowing even weak modulation to be detectable. Signals farther from generation sites may have dropped below minimum noise levels required for recovering envelopes or they may require more frequencies to be fitted due to extensive modulation. Additionally, we hypothesize that stratification variability at the generation sites is responsible for the weaker but non-zero tidal modulation observed at these locations, and is the most easily-observed amplitude modulation that both tidal constituents exhibit.

The processes that modulate the M_2 and S_2 currents appear to vary throughout the study domain, suggesting local modulation is captured in Figure 2.10. The strongly annual modulation at the mouth of the Columbia River is likely due to the substantial seasonal change in stratification associated with the seasonally shifting Columbia River plume (B. M. Hickey et al., 1998; B. Hickey et al., 2005; Thomas & Weatherbee, 2006; Burla et al., 2010). The Columbia River plume is known to be tidally pulsing and to generate internal waves, a process separate from the topographical generation of internal tides (Nash & Moum, 2005). The summer magnitudes of both constituents at this location are roughly three times the magnitudes in winter. Near Cape Mendocino, the envelopes of M_2 and S_2 current ellipses are also highly correlated but at longer, nonperiodic time scales, which is also typical of most locations with correlated amplitudes. Annual modulation is also observed throughout the domain, as evidenced by peaks at $f_{M_2} \pm f_{S_a}$ and $f_{S_2} \pm f_{S_a}$ in Figure 2.4, suggesting that both seasonal and interannual processes modulate tidal currents and lead to the non-phase-locked tide. The correlation of the modulation of the independently-forced M_2 and S_2 currents indicates that both respond similarly to the same processes in the ocean.

Lorentzian fits to the cusps of O_1, M_2, and S_2 indicate phase and/or amplitude modulation of these constituents at time scales in the range of 180 to 240 days. Zaron (2022) performed a similar fit for the M_2 cusp and obtained latitude band averaged values for λ^{-1} ranging from 151 days to 234 days. Our analogous value of $\lambda_{M_2}^{-1} = 243$ days is larger than the largest value in that study, though our range of values overlaps that of Zaron (2022). Other differences between the fitting procedures may explain this discrepancy: we omit tidal peaks and fit coastal surface

current data to a Lorentzian probability density function within a band of $f_{M_2} \pm 5$ cpy and $\Delta f =$ 1/9 cpy, while Zaron (2022) uses FES2014 to remove barotropic tides (among other corrections) to fit global altimeter data to a Lorentzian cumulative density function within a band of $f_{M_2} \pm$ 12.8 cpy and $\Delta f = 1/4$ cpy. The relative agreement of time scales given by a similar analysis of different physical quantities suggests both are modulated by the same processes.

2.6   Conclusions

We use least squares harmonic analysis to characterize the tidal components of over 9 years of hourly coastal surface currents as observed by high-frequency radar. The HFR network provides data at a spatial resolution comparable to that of recent and upcoming Earth-observing satellite missions. The length and sampling rate of these records enable us to recover details about the modulation of tidal processes at time scales up to a few years. The removal of the more predictable barotropic tide using a high-resolution finite element model enables analysis of the baroclinic component, which is expected to be much more sensitive to interaction with ocean processes due to its slower propagation speed and dependence on stratification.

The phase structure of tides in the California Current indicates a complicated field of internally propagating waves originating from several generation sites, with an inhomogeneous distribution of tidal energy and non-phase-locked tides. Near these generation sites, tidal currents are more energetic and more phase-locked, though the modulations of the amplitudes of the two dominant semidiurnal tides (M_2 and S_2) are most significantly correlated in these regions. These constituents are independently forced, suggesting that their amplitudes are modulated by the same processes. Stratification likely drives this, with tidal currents in the vicinity of the seasonal Columbia River plume exhibiting strongly annual modulation, while elsewhere modulation varies at a wider range of frequencies. Fitting harmonic amplitudes (a spectral quantity) to a Lorentzian distribution indicates that the time scales of amplitude and/or phase modulation are 243 days for M_2, 181 days for S_2, and 218 days for the non-propagating O_1; the range of the first of these

76

values overlaps with those of the M_2 analysis of Zaron (2022), who examined global satellite altimetry data.

The length scale at which the fraction of phase-locked M_2 and S_2 surface current variance varies is on the order of tens of kilometers, which is resolved by HFR and contemporary and future satellite missions. The removal of the barotropic component via the FES2014 model reduces the correlation length scale of this quantity in M_2 (24 km before, 15 km after removal) and eliminates the domain-wide consistency of that constituent's phase.

With the recent launch of SWOT, incoming data will provide a global view of tidal variability at scales presently available only near well-observed coastlines. This study provides the share of non-phase-locked tidal surface current variance as well as the temporal and spatial scales of this quantity, which could help inform the interpretation and assimilation of SWOT data (e.g. Metref et al., 2020, whose methodology omits internal tides). Additionally, the specific distribution of tidal variability off the US West Coast can help inform regional studies within this domain, for example to target or avoid locations with strongly modulated tidal surface currents. Combining results with stratification records, the predictability of tidal amplitude modulation could be assessed in future work, with implications for satellite missions and a more comprehensive understanding of the transformation of barotropic tides ultimately to turbulence.

2.7 Appendix for Chapter 2

The appendix summarizes the methodology of Kachelein et al. (2022), which is used in this study. It also describes how the baroclinic tidal component is estimated from the HFR data and barotropic tide model output.

2.7.1 Overview of Harmonic Decomposition of HFR Data

A time series $u(t)$ that is discretely and possibly unevenly sampled may be modeled as the sum of sinusoids at M prescribed frequencies $\mathbf{f}$:

$$\mathbf{u} = \mathbf{H}\mathbf{x}_u + \mathbf{r}_u, \tag{2.5}$$

where $\mathbf{u}$ is the N-long column vector representation of $u(t)$, $\mathbf{H}$ is the N-by-$2M$ regressor matrix of sines and cosines at each frequency in $\mathbf{f}$, $\mathbf{x}_u$ are coefficients to the sines and cosines, and $\mathbf{r}_u$ is the residual time series that represents the unmodeled component of the time series. We use f for ordinary frequency rather than angular frequency $\omega = 2\pi f$ throughout. This is not to be confused with the common use of f to refer to the Coriolis parameter, the frequency of inertial oscillations, which are fitted but not investigated in this study.

The ordinary least squares approach (e.g. Draper & Smith, 1998; Foreman & Henry, 1989; Wunsch, 1996; Menke, 2018) seeks to find the $\mathbf{x}_u$ that minimizes the sum of the squares of $\mathbf{r}_u$. With only this constraint, the solution is

$$\hat{\mathbf{x}}_u = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{u}, \tag{2.6}$$

where the hat indicates that this is an estimator for some unknown true quantity. Given prior knowledge of $\mathbf{u}$, additional terms can be included in this solution that account for properties of $\mathbf{u}$ that equation 2.6 does not consider, including the covariance of $\mathbf{x}_u$ and $\mathbf{r}_u$. Assuming independent Gaussian statistics of $\mathbf{x}_u$ and $\mathbf{r}_u$ and working from Bayes' theorem (see Lee, 2012; Kachelein et

al., 2022), $\mathbf{x}_u$ can be estimated by

$$\hat{\mathbf{x}}_u = (\mathbf{H}^{\mathrm{T}}\mathbf{R}_{uu}^{-1}\mathbf{H} + \mathbf{P}_{uu}^{-1})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}_{uu}^{-1}\mathbf{u}, \tag{2.7}$$

where $\mathbf{R}$ denotes the expected autocovariance matrix of the residual of the subscripted quantity, here $\mathbf{R}_{uu} = \langle \mathbf{r}_u\mathbf{r}_u^{\mathrm{T}}\rangle$ and $\mathbf{P}$ denotes the expected autocovariance matrix of the model parameters of the subscripted quantity, here $\mathbf{P}_{uu} = \langle \mathbf{x}_u\mathbf{x}_u^{\mathrm{T}}\rangle$. These are estimated from the power spectrum of the data, $S_{uu}(f)$. Because the coefficients $\mathbf{x}_u$ are the amplitudes of sines and cosines, they also represent an estimate of the spectral power of $\mathbf{u}$ and therefore their autocovariance matrix $\mathbf{P}_{uu}$ may be estimated from the power spectrum at frequencies in $\mathbf{f}$. Similarly, the residual autocovariance $\mathbf{R}_{uu}$ may be estimated from the non-tidal component of the power spectrum via its Fourier transform.

2.7.2 Removal of Barotropic Estimate

The harmonically decomposed HFR data are given as sine and cosine amplitudes, while the FES2014 barotropic tide model output is given by amplitude A and Greenwich phase lag G, which is related to tidal phase by $\phi = (V_\circ + u) - G$ (F. Lyard, 2022, personal comm.), where $(V_\circ + u)$ is the equilibrium argument for some reference time (Parker, 2007), in our case 2012-01-01 00:00:00 UTC. Values for the equilibrium argument can be computed or referenced in tidal charts, e.g. Zetler (1982). The sine amplitude a and cosine amplitude b can be calculated from total amplitude and phase by:

$$a = -A\sin\phi, \qquad b = A\cos\phi \tag{2.8}$$

Calculating the difference between two vector quantities that are sinusoidal in time is simpler when written as separate sine and cosine amplitudes than when written as total amplitude times a phase-shifted cosine:

$$u_3(t) = u_1(t) - u_2(t) = a_3\sin(\omega t) + b_3\cos(\omega t) = A_3\cos(\omega t + \phi_3), \tag{2.9}$$

where $a_3 = a_1 - a_2$ and $b_3 = b_1 - b_2$.

Chapter Two, in part, is currently being prepared for submission for publication of the material in the Journal of Geophysical Research: Oceans. Kachelein, L., S. T. Gille, M. R. Mazloff, and B. D. Cornuelle. Characterizing Non-phase-locked Tidal Currents in the California Current System using High-frequency Radar (In Preparation). The dissertation author was the primary investigator and author of this material.

Figures (temporary section)
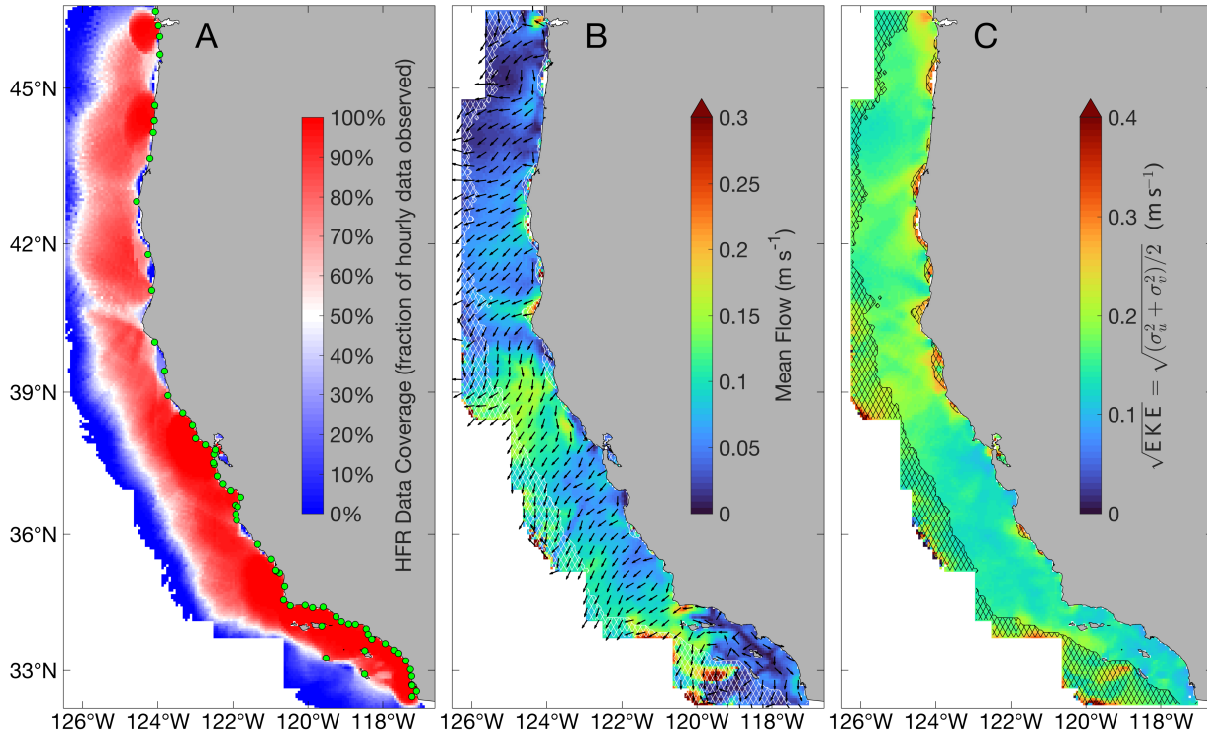
Figure 2.1. (a) High-frequency radar (HFR) network coverage in the study domain (percent hourly observations available). Green dots indicate locations of HFR stations. (b) Time-mean flow of the total current **u**, including at locations not analyzed for tides. Colors indicate vector magnitude |**u**|, uniform-length arrows indicate direction **u**/|**u**|. (c) Square root of the eddy kinetic energy of **u**. Hatching in panels b and c indicates regions with less than 50% coverage, which were not analyzed for tides.

Figure 2.2. Constructed examples of (a) apparent modulation from linear addition of two tidal constituents and amplitude modulation by (b) sinusoidal and (c) broadband 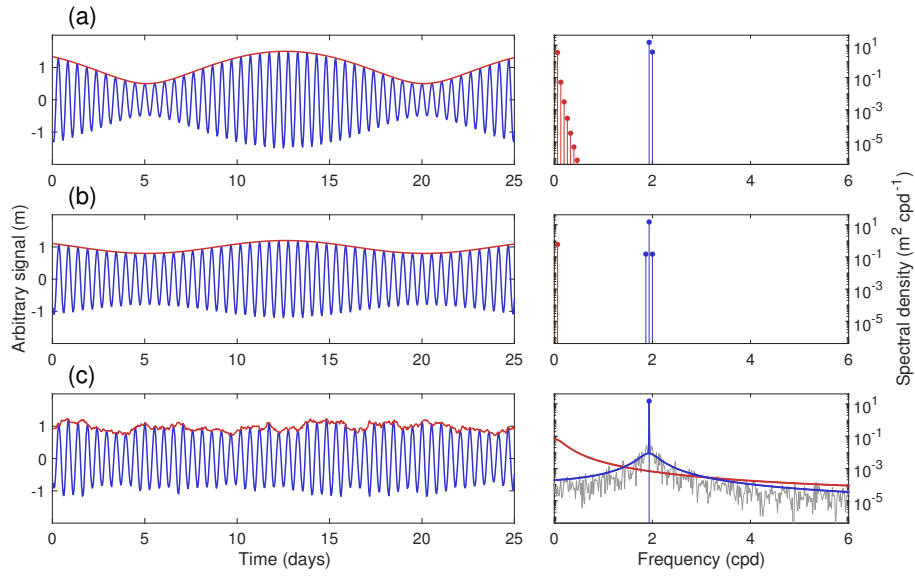processes. The square of the Fourier transform of the broadband-modulated tidal signal is shown in gray in the bottom-right panel.
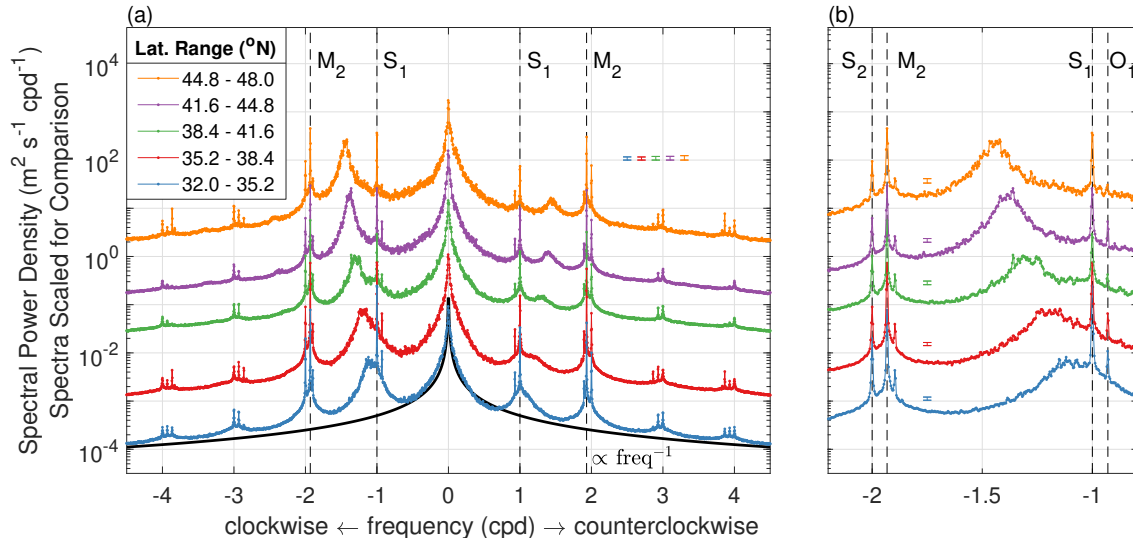
Figure 2.3. (a) Rotary power spectral density of surface current **u** from HFR averaged within five latitudinal bands, plotted on a linear frequency axis and logarithmic spectral power density axis. Spectra are vertically offset to avoid overlap. A black curve corresponding to a spectral slope of -1 is plotted for reference. Panel (b) displays the diurnal to semidiurnal band in the negative (clockwise) frequencies. Small colored vertical error bars indicate the 95% confidence ratio, which corresponds to a ratio valid at all magnitudes of the correspondingly colored spectral density.
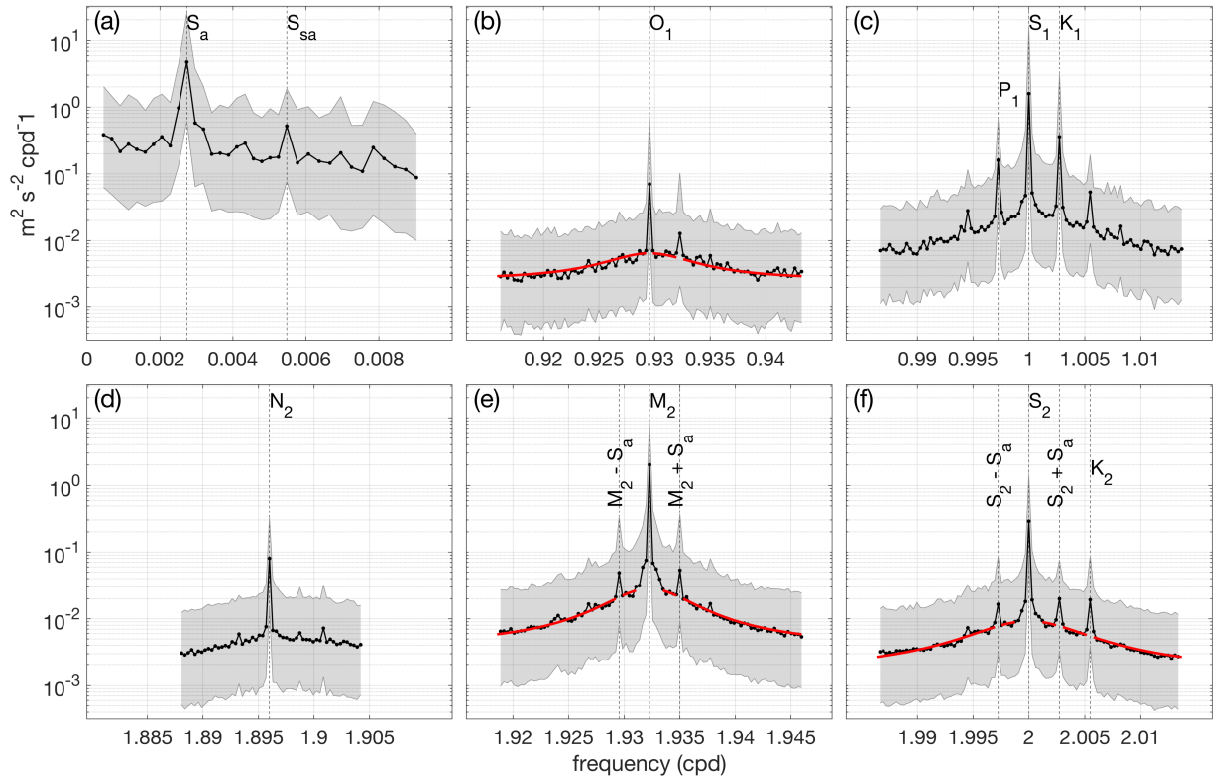
Figure 2.4. Domain-median total HFR current variance $\frac{1}{2}(\sigma_u^2 + \sigma_v^2)$ as given by harmonic decomposition, divided by (variable) frequency step $\Delta f(f)$ to give power spectral density. Shading indicates 5th and 95th percentiles. Prominent tidal constituents are labeled and all panels span the same vertical interval. Panel (a), which shows the variance recovered from harmonic decomposition at frequencies $\leq 1/110$ cpd, is scaled horizontally to accommodate fewer fitted frequencies, while panels (b-f) are all equally horizontally scaled. The red curves in panels (b), (e), and (f) are the non-linear least squares fitted curves to tidal cusps plotted at the fitted frequencies.
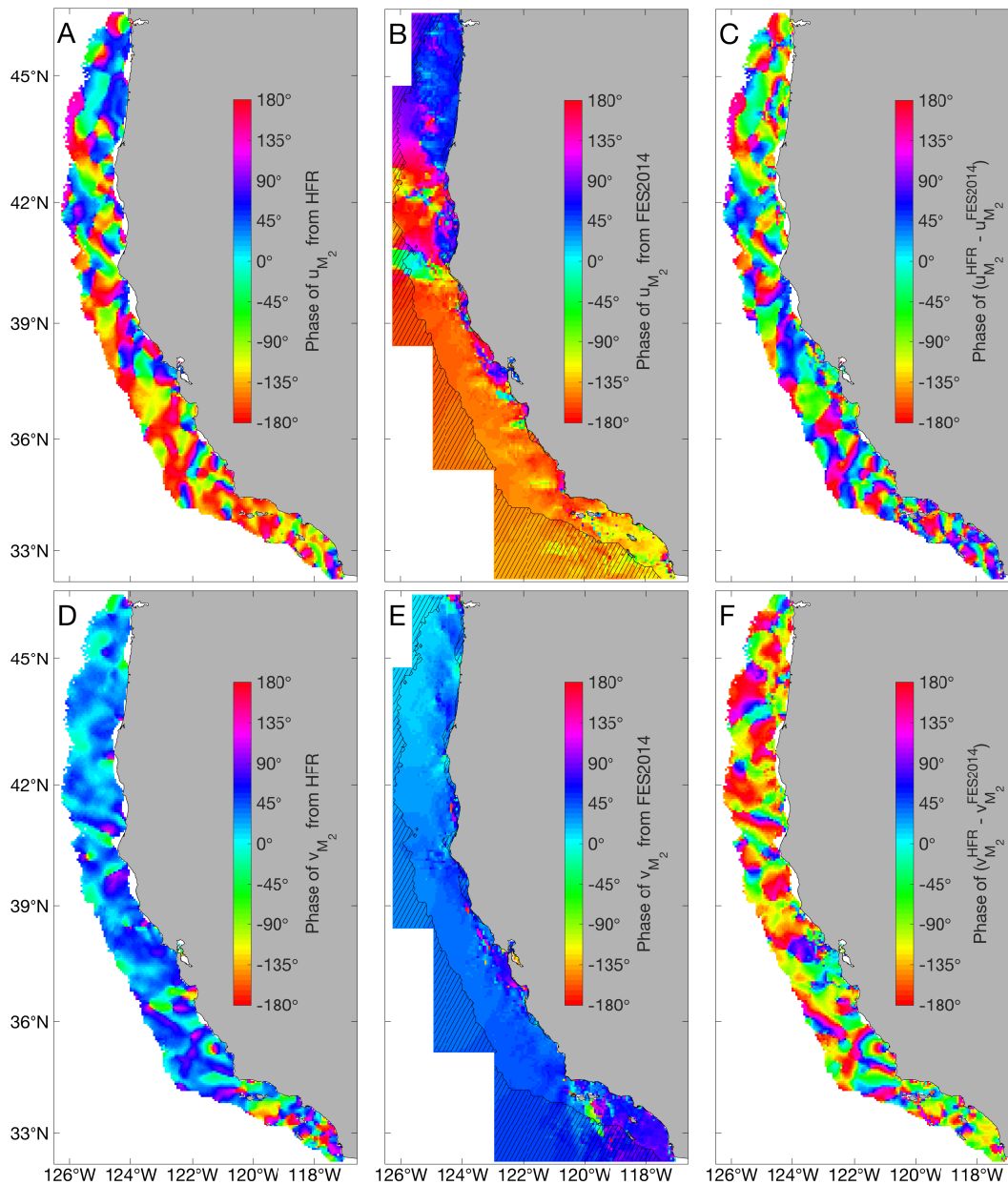
Figure 2.5. Tidal phase (degrees) of harmonically decomposed surface currents at the M_2 frequency. (a-c) show phase of the *u*-component, (d-f) show phase of the *v*-component. (a,d) show results from decomposed HFR data, (b,e) show results from the FES2014 tide model, and (c,f) show results from the difference between the data and model. Masking on panels (b,e) indicate locations where the model is defined but HFR data were not analyzed.
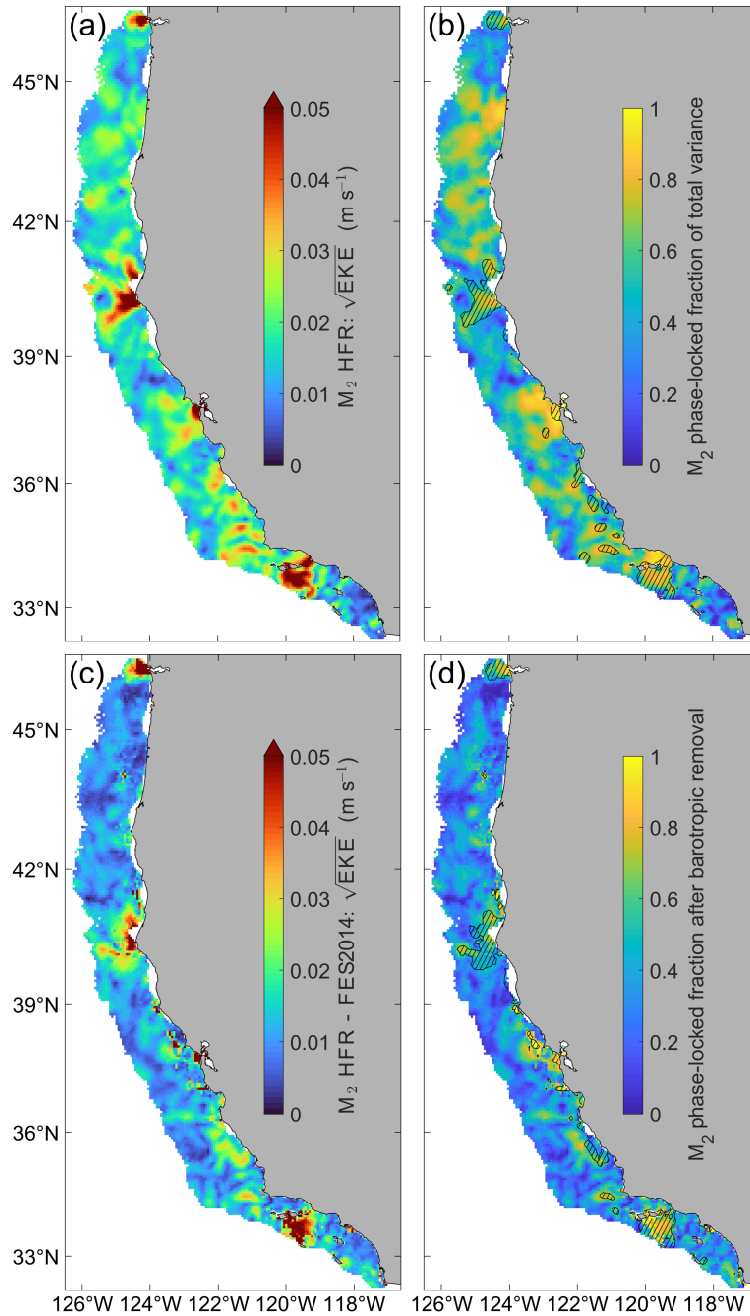
Figure 2.6. (a) Square root of eddy kinetic energy of the M_2 component of harmonically decomposed HFR data. (b) Fraction of M_2-band variance considered to be phase-locked without removal of the modeled barotropic component. (c-d) are analogous to (a-b) but with the modeled barotropic component removed. The "total variance" of the M_2 component is calculated within the frequency band $f_{M_2} \pm 5$ cpy. Striped regions in (b) indicate the top decile of variance, where the total variance within the M_2 band is $\geq 2.8 \cdot 10^{-3}$ $m^2 s^{-2}$, while striped regions in (d) indicate the top decile after barotropic removal, where the total variance within the M_2 band is $\geq 2.4 \cdot 10^{-3}$ $m^2 s^{-2}$.
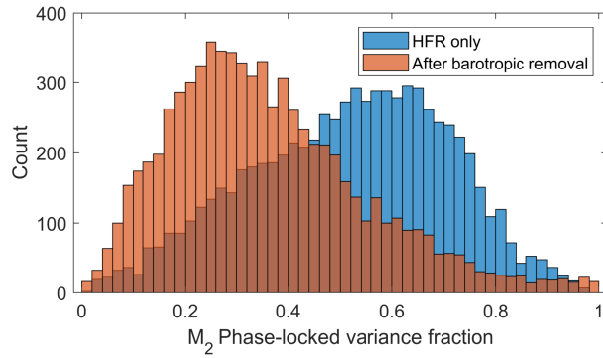
Figure 2.7. Histogram representation of the quantities in Figure 2.6 (b) and (d).
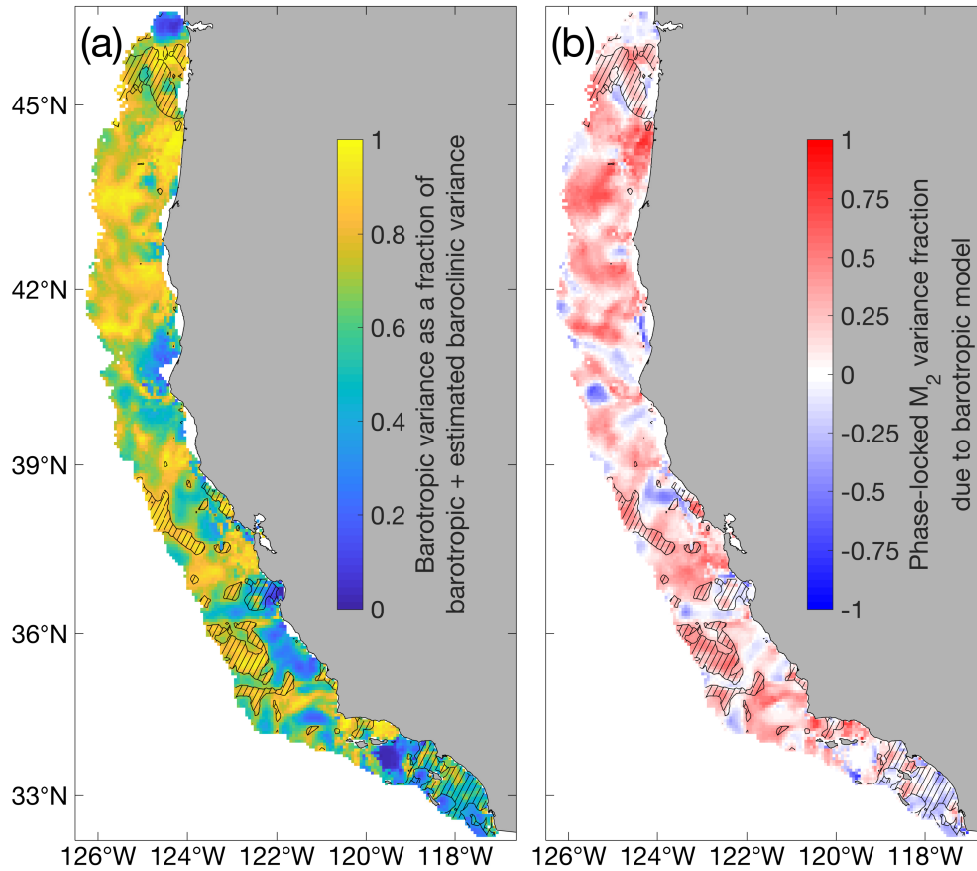
Figure 2.8. (a) Modeled barotropic M_2 variance as a fraction of total variance (modeled barotropic plus estimated baroclinic) and (b) the fraction of phase-locked M_2 variance attributable to the barotropic modeled component (the difference of Figure 2.6b and 2.6d). Negative values in panel (b) indicate greater current variance after the removal of the barotropic component. Regions with the 20% lowest total variance (the denominator in panel a) are hatched, i.e. where modeled barotropic variance plus estimated baroclinic variance is $\leq 7.05 \times 10^{-4}$ m^{2}s^{-2}.
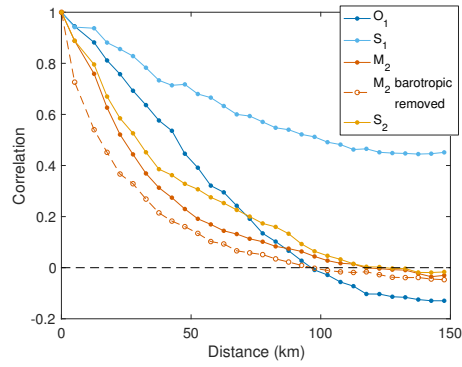
Figure 2.9. Isotropic correlation of phase-locked fraction of variance as a function of distance. Correlation drops to 0.5 at 45 km for O_1, 101 km for S_1, 24 km for M_2 (15 km for M_2 after barotropic removal), and 29 km for S_2. Error estimates are not shown due to strong convergence across $O(10^3)$-$O(10^5)$ realizations.
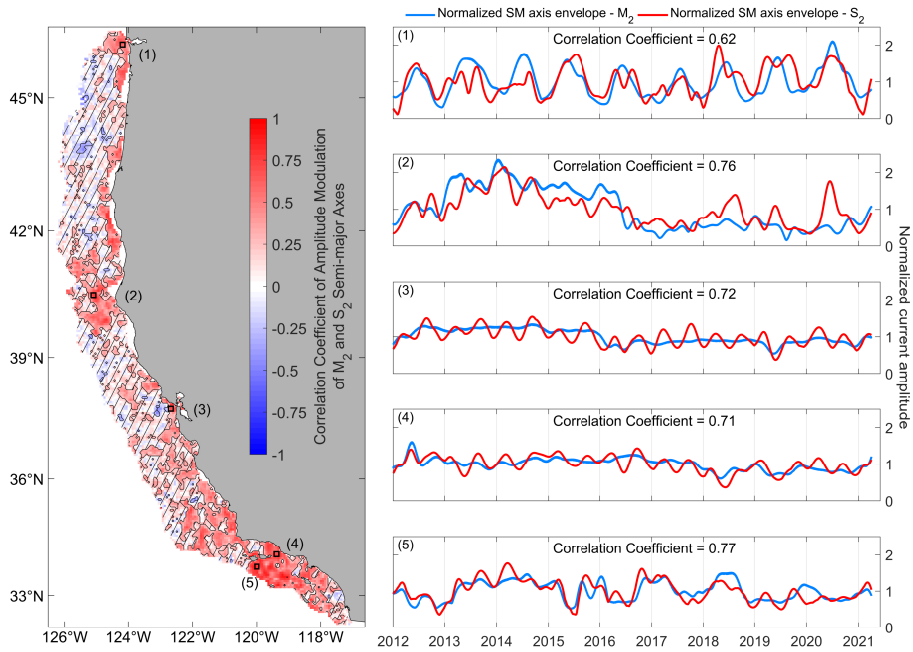
Figure 2.10. The correlation coefficients of the recovered modulation of the M_2 and S_2 vector current semi-major axes, calculated from the upper envelopes of $\sqrt{u_{M_2}^2 + v_{M_2}^2}$ and $\sqrt{u_{S_2}^2 + v_{S_2}^2}$. The modulating envelopes (alternatively the time-evolution of the semi-major axes) contain variance at fitted frequencies ≤ 5 cpy (periods ≥ 73 days), with coefficients in the frequency domain multiplied by a Hanning window in order to reduce spectral leakage. Striped regions indicate that the absolute value of the correlation coefficient is less than the 90% significance level calculated from the estimated number of degrees of freedom from the characteristic timescale of the M_2 and S_2 envelopes at each location. (1-5) are the normalized modulating envelopes for the semi-major axes of $\vec{u}_{M_2}$ and $\vec{u}_{S_2}$ at each location labeled on the map.

Chapter 3

The Diurnal Cycle in the California Current System: Currents from High-Frequency Radar and Winds from ERA5

Abstract

We evaluate 9 years of hourly HFR surface current data and contemporary 10-meter wind fields from ERA5 using harmonic analysis in order to characterize the diurnal current in the California Current System, evaluate its coherence with diurnal winds, and examine the frequency-dependent response of currents to wind forcing via the deflection angle, a scalar quantity that can be used to evaluate the surface ocean response to wind forcing and compare this to theoretical models. The time of maximum wind speed in the diurnal cycle was found to increase with distance from the coast, whereas the time of maximum current speed is more uniform in the winter but less uniform in summer, when current ellipses exhibit low eccentricity. For energetic winds and currents, the optimal time of day to sample via sun-synchronous orbit in order to maximize the amplitude of the observations is between approximately 0400 and 0500 or 1600 and 1700 local mean time. Currents and winds were found to be highly coherent at the diurnal frequency, especially between the dominant clockwise (anticyclonic) components. The phase difference between these quantities reasonably approximates the current deflection angle, which is approximately 45° to the right of high-amplitude winds as predicted by classical Ekman theory but which increases toward 90° at lower wind speeds offshore. At most other frequencies, the deflection angle was greater than 45° except for frequencies below approximately 0.1 cycles per day, which are typically between 0° and 45°. This deviation from both classical and more complicated spectral models suggests additional physical considerations for modeling air-sea interaction.

89

3.1 Introduction

The circulation of the global ocean relies upon a continual input of energy (W. Wang & Huang, 2004), with an estimate on the order of 2 TW required to mix the abyssal ocean (W. H. Munk & Wunsch, 1998), with winds and tides as the only sources of mechanical energy available to drive interior mixing. Both tides (Baines, 1982; W. H. Munk & Wunsch, 1998; Egbert & Ray, 2000; R. D. Ray & Cartwright, 2001; Garrett & Kunze, 2007) and winds are the subjects of investigation, with the wind-driven input divided further into wind stress contribution to the geostrophic component of surface flow, estimated at 0.85 TW to 1.0 TW (Scott & Xu, 2009); contribution to near-inertial motions, estimated at 0.5 TW to 0.7 TW (Watanabe & Hibiya, 2002; Alford, 2003); and contribution at subinertial frequencies, estimated at 2.4 TW (W. Wang & Huang, 2004), though recent estimates of wind work on ageostrophic currents have been revised significantly upward (Yu et al., 2018). Additionally, surface current feedback and thermal feedback can result in the ocean forcing the winds (Villas Bôas et al., 2019), further complicating any simple descriptions of air-sea interactions. The diurnal frequency, a highly energetic component of the wind (Deser & Smith, 1998; Dai & Deser, 1999), is subinertial poleward of 30°N/S, near-inertial in the vicinity of this critical latitude, and is the frequency of solar radiative forcing on the oceans. Hence its contribution to the global wind work on the oceans is expected to be prominent and complicated, as is the ocean's diurnal response, which includes interactions between the diurnally heated upper layer and wind-forced velocity component, resulting in diurnal jets (Price et al., 1986; Masich et al., 2021).

The diurnal air-sea interaction is therefore relevant to understanding the broader global ocean energy budget, and due to its ubiquity at the surface, it is important to any observation system that aims to understand the Earth's climate. This includes satellites at repeat orbits that cannot resolve the diurnal frequency. Sun-synchronous missions, including most scatterometers (Rodríguez et al., 2019), fall into this category; not only do such orbits not provide enough information to resolve details of diurnal processes, but they also effectively sample the diurnal cycle

90

at its Nyquist frequency (Villas Bôas et al., 2019). In order to avoid sampling (generally elliptical) vector surface currents at a low-amplitude phase of their diurnal cycle, thus underestimating total current variance, prior knowledge of the predominant amplitude and phase of the diurnal cycle may be important for the planning of future sun-synchronous, vector-observing satellite missions, such as the proposed Ocean DYnamics and Surface Exchange with the Atmosphere (ODYSEA) mission, formerly the Winds and Currents Mission (WaCM) (Rodríguez et al., 2019). The purpose of this study is to evaluate diurnal winds and currents in the California Current System (CCS) and provide an overview of diurnal processes in this region, both to inform potential future sun-synchronous satellite missions and to improve understanding of these processes in relation to previous observations and theoretical models.

To achieve this, we analyze gridded surface current data from land-based high-frequency radar stations, which use scatterometry, and a gridded reanalysis product for 10-meter winds. Our analysis consists of a Bayesian harmonic least squares fit originally formulated for tides in Chapter 1, published in Kachelein et al. (2022), and implemented in Chapter 2; this method can be applied to any frequency. From 9 years of data, we calculate mean current/wind ellipses, squared coherence of rotary components, phase difference, and the deflection angle of current from wind as a function of frequency, with emphasis on the diurnal cycle. We begin with an overview of the data in Section 3.2 and analysis methods in Section 3.3, which also includes an overview of a previous Ekman-type model describing the frequency-dependent response of the upper ocean boundary layer to wind forcing. Section 3.4 presents the results of our analysis, followed by a discussion in Section 3.5 and concluding remarks in Section 3.6. Mathematical formulas are included in the Appendix.

3.2 Data

The meridional and zonal components of vector winds and vector surface currents are analyzed in this study. For notational clarity, ocean currents are written lowercase as $\vec{u} = u\hat{\mathbf{i}} + v\hat{\mathbf{j}}$,

and winds are written uppercase as $\vec{U} = U\hat{\mathbf{i}} + V\hat{\mathbf{j}}$. In the appendix where coherence and other quantities are explained for generic vector quantities, only lowercase is used.

3.2.1 High-Frequency Radar

The high-frequency radar time series analyzed are the same as those examined in Chapter 2, with hourly time series from 1 January 2012 through 1 April 2021, gridded at 6 km spacing in a study region extending from 32.25°N to 46.43°N (see Terrill et al., 2006; Roarty et al., 2019, for more information about the Global High Frequency Radar Network). HFR velocity measurements are generally assumed to be observed in the upper 1 meter (Kim et al., 2009; Lana et al., 2016). A similar analysis is performed using the fitting procedure outlined in Chapter 1, but in order to examine seasonal differences and coherence, we examine 6-month-long segments at coarser frequency resolution as outlined in Section 3.3.1. When analyzing segments, only those taken from years (1 April through 31 March) with at least 50% data coverage were used from time series with at least 50% total coverage; the number of summers and winters at each location that satisfy this criterion are displayed in Figure 3.1 alongside HFR station locations.

3.2.2 Wind Product

We use hourly 10-meter wind speeds on a ~31 km (~0.28°) grid from the ERA5 reanalysis implemented by the European Centre for Medium Range Weather Forecasts (ECMWF) (Hersbach et al., 2018, 2020). It is the fifth generation of atmospheric reanalysis of the ECMWF and is based on the Integrated Forecasting System (IFS) Cy41r2. Billions of atmospheric and ocean observations are assimilated, resulting in an improvement over previous reanalyses. Currently, ERA5 covers the time period from 1940 to the present; we use data from 2012 to 2021, contemporary with the HFR records used in this study. Wind grid cells that contain HFR grid cells are used for coherence analysis, while all wind grid cells over the ocean and coastline are used to evaluate spectral properties of the wind in order to maximize the degrees of freedom, resulting in smoother covariance estimates for the harmonic analysis described in Section 3.3.1.

The 10-meter wind product exhibits seasonality, with a strong spectral peak at the annual frequency (not shown) and a seasonal difference in the distribution of wind direction. Figure 3.2 shows wind roses (polar probability density functions) of ERA5 winds from 2011 to 2021 sorted by month after removal of the 11-year mean wind vector for locations approximately co-located with HFR observations: $\overline{U(t)} = 2.24$ m s^{-1}, $\overline{V(t)} = $ -2.97 m s^{-1}, angle -53.0° from eastward, corresponding to wind blowing southeastward, tilted parallel to the coastline of the study region. Bins correspond to the direction toward which the wind blows and each wind rose is scaled to the maximum histogram bar radius in order to appear visually consistent. April through September winds are characterized by a narrower range of directions, with most wind vectors pointing southeastward along the coast, while winds from October through March exhibit greater angular variability. We designate the former as "summer" and the latter as "winter" for the purposes of this study.

A known issue with this data set is a mismatch in near-surface wind speed between assimilation cycles, occurring from 0900 to 1000 and 2100 to 2200 UTC[1]. This results in small but noticeable jumps at 1000 and 2200 UTC in both U and V, typically smaller than the amplitude of the diurnal cycle. Assuming these jumps are consistently in one direction, this would be equivalent to adding a Dirac comb with 12-hour spaced finite peaks to the wind time series, with an effect in the frequency domain of adding a Dirac comb to the wind power spectrum with peaks spaced at 2 cpd, which would not affect the power at the diurnal frequency. However, for time series jumps that are different for morning versus evening, spectral power at the diurnal frequency is affected. The spectral power of a discrete diurnal Dirac comb with amplitude 1 is 0.69% that of a sinusoid with amplitude 1; because the magnitude of observed jumps appears to be less than the amplitude of the diurnal cycle of the wind, we assume that the impact of any spectral artifacts present in ERA5 10-meter winds is negligible.

---

[1]For more information, see point 8 at
https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation#heading-Knownissues.

3.2.3 Spectral Properties

The rotary power spectra (e.g. Gonella, 1972) of both currents and winds used in this study are shown together in Figure 3.3. As both quantities are velocities, their spectra share the same units of spectral power ($m^2\,s^{-2}\,cpd^{-1}$). Wind spectra are averaged over all locations in the domain located over the ocean or bordering an ocean grid point, while the current spectra displayed here are averaged over observations taken between 41.6°N and 44.8°N in order to capture the latitude-dependent near-inertial response centered at about 1.4 cpd. All spectra displayed are calculated using Welch's method on overlapping segments multiplied by a Hanning window. For this figure, wind time series are divided into 365 days-long segments and current time series are divided into 281.5 days-long segments. These are not the same lengths of segments that are analyzed later for estimating coherence. Visual indications of error ratios are omitted due to the large number of realizations leading to reasonably smooth spectra.

Wind spectra display prominent peaks at the diurnal frequency and at every one of its resolvable harmonics. The narrow width of the peaks and lack of a discernible increase in background spectral power at nearby frequencies indicate minimal modulation at time scales resolvable with these segment lengths. The higher-order harmonics of the diurnal cycle may be influenced by the twice-daily mismatches associated with the ERA5 assimilation cycle discussed in Section 3.2.2; the assumption that these mismatches are negligible for our analysis is supported by the fact that the amplitudes of the peaks of these higher-order harmonics are two to three orders of magnitude smaller than those of the diurnal, semidiurnal, and low-frequency processes examined in this study. The broadband background power of the wind is roughly proportional to ω^{-2} for frequencies above 0.1 cpd and below 4 cpd, and the clockwise polarized component dominates at most frequencies.

Surface currents also display prominent peaks at the diurnal and semidiurnal frequencies, as well as at tidal frequencies, such as O_1 and M_2. Unlike winds, currents are only significantly clockwise-polarized between approximately 0.4 cpd and 3 cpd, with the strongest polarization

observed in the near-inertial frequency band centered around 1.4 cpd due to the near-inertial response, which is resonant in this frequency band at these latitudes. Energy at higher harmonics of the diurnal frequency is present in current spectra, though less prominently than in wind spectra. The broadband background power of the current is roughly proportional to ω^{-1} at all resolved frequencies.

3.3   Methods

3.3.1   Harmonic Analysis

In order to facilitate analysis of the coherence of wind and current, we fit to a different set of frequencies than in Chapter 2, which examined records spanning over 9 years. Here, we fit both wind and currents to the same set of frequencies within the same 6-month segments, summer (defined as 1 April through 30 September) and winter (defined as 1 October through 31 March), which necessitates more widely-spaced frequencies. These two seasons are defined such that summer corresponds to prevailing winds that are alongshore toward the south-east, while winter corresponds to winds that display a wider range of angles, though still tending to exhibit the same alongshore direction as in summer or its reverse, as in Figure 3.2. For currents, the same procedures as in Chapter 2 are used (five latitude-dependent spectra to build $\mathbf{P}_u$ and $\mathbf{P}_v$ and a spectral slope of -1 and a residual variance fraction of 0.6 to build $\mathbf{R}_u$ and $\mathbf{R}_v$), with matrices resized for shorter segments and fewer frequencies. Seasons are not considered when constructing covariance matrices for currents. For winds, the domain-averaged spectra S_{UU}^{s}, S_{VV}^{s}, S_{UU}^{w}, and S_{VV}^{w} of 10-meter winds over the ocean are used to construct separate $\mathbf{P}_U$ and $\mathbf{P}_V$ for summer and winter, where spectra are the average of 7263 squared discrete Fourier transforms (807 locations, 9 seasons each of summer and winter). These are different from the spectra displayed in Figure 3.3, which combine seasons. No windowing was performed in estimating the 4 seasonal wind spectra: data near the beginning and end of each segment were therefore not de-emphasized and convergence to reasonably smooth spectra was achieved from the large

number of segments. The residual covariance matrices **R** for both components and seasons were estimated using these same spectra at unfitted frequencies, a residual variance fraction of 0.1, a cutoff of 50 hours, and a triangular tapering window to reduce spectral artifacts. The $M = 157$ frequencies to which both wind and current were fit include 104 low frequencies (less than 100/182 cpd), 11 frequencies centered on O_1, 11 frequencies centered on the diurnal cycle (S_1), 11 frequencies centered on M_2, 11 frequencies centered on the semidiurnal cycle (S_2), and each of the higher-order harmonics of S_1 up to 11 cpd. The frequencies in the O_1, S_1, M_2, and S_2 bands are spaced by the fundamental frequency 2 cpy, as are those in the low-frequency band except for the 8 lowest fitted frequencies, which are spaced by 1 cpy (lowest frequency 1 cpy) in order to reduce the effect of enforcing 6-month periodicity on a record that is known to incompletely resolve a strong annual cycle.

3.3.2   Coherence and Rotary Analysis

We evaluate the relationship between 10-meter winds and surface currents by estimating squared coherence γ^2 (hereafter simply "coherence"). This quantity is the frequency-domain analogue of the squared correlation coefficient (von Storch & Zwiers, 2003), and is used here to investigate the effect of wind on current and the predictability of current from the ERA5 reanalysis product. In addition to coherence, we estimate the phase angle ϕ (hereafter "phase"), which describes the fraction of a cycle by which one process leads the other. We estimate the coherence and phase between rotary components (clockwise and counterclockwise rotating) of both wind and current (e.g. Gonella, 1972), as both quantities are dominated by rotating processes at the frequencies of interest and a rotationally-invariant framework is appropriate for the angled coastline of the domain. The phase of counter-rotating components only describes their relative angle at time 0, with all possible differences in phase visited over one cycle, hence we present this value but do not examine it in detail and instead focus on co-rotating coherence and phase.

The estimation of coherence is outlined in the Appendix, as is the rotary framework implemented here. The commonly-implemented method of using averaged Fourier transforms of

96

two quantities to compute cross-spectra, and hence coherence, is modified to use the sine and cosine coefficients given by harmonic analysis, which allows for fitting over gaps in HFR data and for fitting arbitrary frequencies that are not periodic over the record length, reducing spectral leakage. Seasonal segments are few in number (at most 9 seasons for each of summer and winter) and cannot be overlapped, thus they are not multiplied by any windowing function.

3.3.3 Frequency-dependent Response of Upper Oceanic Boundary Layer to Wind

Elipot & Gille (2009) evaluated several idealized Ekman-type models of wind forcing of the oceanic boundary layer, comparing them to surface drifter observations in the Southern Ocean from 30°S to 60°S. In this study, we compare observations to their "best" model, denoted 1b in that study. The theoretical frequency-dependent transfer function from Elipot & Gille (2009), which uses ν for ordinary frequency, to which we compare our result is

$$\mathbf{H}_{1b}(\nu, z) = \frac{e^{-i\pi/4}}{\rho\sqrt{(2\pi\nu + f_{\mathrm{Cor}})K_0}} \frac{\sinh\left[(1+i)(h-z)/\sqrt{2K_0/(2\pi\nu + f_{\mathrm{Cor}})}\right]}{\cosh\left[(1+i)h/\sqrt{2K_0/(2\pi\nu + f_{\mathrm{Cor}})}\right]}, \quad (3.1)$$

which we evaluate at the surface ($z = 0$) and for Coriolis frequencies f_{Cor} corresponding to the latitudes of this study. We use $h = 35$ m and $\rho = 1027$ kg m^{-3}, the first of which falls within the range of values that that study found to best explain drifter behavior at Southern Hemisphere latitudes equivalent to those we examine in the Northern Hemisphere. We test a range of values for K_0 such that the misfit between the median observed angle and model angle at each frequency is minimized. We only examine the phase of this quantity (see equation (3.12)), as we do not calculate wind stress $\vec{\tau}$ from $\vec{U}$, which are parallel and thus have the same phase relationship with $\vec{u}$.

3.4 Results

3.4.1 Characteristics of Wind and Currents

Diurnal winds are intensified near the coastline and are most prominent in the Southern California Bight as shown in Figure 3.4a, which shows the semi-major axis or peak wind speed in the diurnal cycle for the colored background, and the diurnal wind ellipses at alternating grid points averaged over summer (black) and winter (white). Winter diurnal ellipses are typically smaller than and similarly elliptical to the co-located summer ellipse. A non-linear least squares fit of peak diurnal wind speed $U_{S_1}^{\text{peak}}$ was performed using the Gauss-Newton algorithm (Bishop et al., 1985), where its functional dependence on distance to the nearest coastline, D_c, is assumed to have amplitude U_c at the coast and decay exponentially with length scale L to some possibly non-zero value in the open ocean U_o:

$$U_{S_1}^{\text{peak}}(D_c) = U_c \exp\left(-\frac{D_c}{L}\right) + U_o.$$

(3.2)

In summer, the parameters U_c, U_o, and L were respectively 1.17 m s^{-1}, 0.29 m s^{-1}, and 71 km ($R^2 = 0.46$), while in winter these parameters are 0.89 m s^{-1}, 0.06 m s^{-1}, and 75 km ($R^2 = 0.41$).

Figure 3.4b shows that the amplitude of the diurnal current is distributed similarly to that of wind, with a wider range of elliptical tilts, moderate eccentricity, and intensification in the Southern California Bight, which is magnified in panel (3.4c) for clarity without changing the scale of the ellipses relative to those in panel (3.4b).

The fraction of total variance accounted for by the diurnal frequency for both winds and currents is displayed in Figure 3.5, which displays the seasonal average across the 9 analyzed summers and winters. The spatial distribution of this quantity is similar to that of the amplitude for both vector processes, with the Southern California Bight exhibiting the greatest relative and absolute energy at the diurnal frequency, and coastal regions generally more energetic than offshore regions. The summer diurnal variance fraction (3.5a,c) is greater than the winter diurnal

variance fraction (3.5b,d), while the pattern of nearshore wind (3.5c,d) appears to follow coastal geometry, indicative of enhanced sea breeze/land breeze generation.

The local time at which diurnal wind and current vectors are aligned with the semi-major axis is shown in Figure 3.6. This alignment occurs twice daily, hence values are shown between 00:00 and 12:00. The average time of maximum diurnal wind does not differ substantially between summer (3.6a) and winter (3.6b): within 71 km of the coast (the exponential decay scale of maximum current in summer) the mean hour in summer is 4:16 am/pm (σ = 2.3 hr after centering the distribution) while in winter within 75 km from shore it is 5:09 am/pm (σ = 1.9 hr). A more coastally limited range of 62 km (two ERA5 grid cells) for both seasons yields 4:03 am/pm (σ = 2.0 hr) and 4:50 am/pm (σ = 1.7 hr) for summer and winter, respectively. The hour of maximum diurnal wind increases with distance from the coast, indicating propagation offshore of a signal that, according to Figure 3.4, grows weaker in this direction. Along the Central California and Central Oregon Coasts, where the phase structure appears to conform to the coast, the peak wind time appears to increase approximately 1-2 hours per 31 km grid cell, equating to approximately 15-30 km h^{-1} ≈ 4-8 m s^{-1}, in rough agreement with Gille et al. (2005). At a wind buoy 220 km offshore of Southern California, Lerczak et al. (2001) found that maximum onshore wind speed occurred 9 hours after local noon, in rough agreement with ERA5 winds.

The average time of maximum diurnal current exhibits seasonal differences: in summer (3.6c), this time is distributed broadly across the 12 hour interval, with a circular mean at 1:15 am/pm (σ = 2.6 hr after centering) while in winter (3.6d) this quantity is more uniform over the domain, with a circular mean of 4:08 am/pm (σ = 1.6 hr). However, diurnal current direction during the summer (not shown) is narrowly distributed: at 0115 PST, the circular mean angle of summer diurnal current vectors is 169° from eastward with σ = 46° after centering.

3.4.2 Coherence between Wind and Currents

The four coherence estimates between the two rotary components each of summer diurnal wind and current are displayed in Figure 3.7. The greatest coherence is observed between

clockwise wind and clockwise current (panel 3.7a), which is the dominant polarization for both processes. South of approximately 39°N, the coherence is very high, while north of this latitude there are regions in which coherence drops below the 95% confidence threshold, which may result from a more limited resolution of current angle by the HFR network; these regions also generally have time series with fewer summers that pass the 50% coverage requirement, resulting in fewer degrees of freedom and hence a higher 95% confidence threshold. Coherence between counterclockwise wind and clockwise current (panel c) is also prominent, which may be explained by the more elliptical winds comprised of a superposition of both polarizations with stable phases. The pixelation along the coast is an artifact of the nearest-neighbor convention of pairing current time series with wind time series and is visible in most of the maps in this chapter that examine quantities that are calculated from both of these quantities. The counterclockwise current is generally less coherent with the wind, except for a few locations (panels 3.7b,d), as the diurnal current ellipses generally exhibit low eccentricity, the result of a weak counterclockwise component. The winter coherence values (Figure 3.8) are similar to those of summer, though with a few differences: maps of $\gamma^2_{\text{winter}}(\text{cw}, \text{ccw})$ and $\gamma^2_{\text{winter}}(\text{ccw}, \text{ccw})$ (the coherence of both polarizations of wind with counterclockwise current, panels 3.8b,d) exhibit greater coherence in general due to the more eccentric winter current ellipses, while $\gamma^2_{\text{winter}}(\text{ccw}, \text{cw})$ and $\gamma^2_{\text{winter}}(\text{ccw}, \text{ccw})$ (the coherence of counterclockwise wind with both polarizations of current, panels 3.8c,d) are nearly zero north of 39°N and more than about 50 km offshore, regions where the diurnal winter winds are much weaker than those in summer.

The phases corresponding to these diurnal coherence estimates are displayed in Figure 3.9 (summer) and Figure 3.10 (winter). In summer, the phase $\phi(\text{cw}, \text{cw})$ (written later as $\phi_{\circlearrowright\circlearrowright}$ for notational convenience) of the highly coherent clockwise-clockwise paring is characterized by the current leading wind by 0 to 9 hours, with longer lead times farther offshore and shorter ones near the coast, especially in regions of strong diurnal currents like the Southern California Bight (Figure 3.9a). The phase associated with the counterclockwise current (panels 3.9b,d) exhibit a wider phase distribution, with structure in phase occurring at smaller spatial scales;

100

because these correspond to mostly insignificant coherence at the 95% confidence level, this structure is not investigated further. The phase of counterclockwise wind and clockwise current (panel 3.9c) exhibits comparable uniformity to panel (3.9a) but at a largely negative phase.

In winter (Figure 3.10), all panels display phases that tend to vary across the domain over large scales. The dominant clockwise-clockwise pairing (panel 3.10a) is similar to its summer analogue, but with higher phase values offshore, indicating current that leads the wind by greater than 6 hours; in the Southern California Bight, however, the phase is largely unchanged from summer. The greater uniformity in phase in panels (3.10b,d) is likely due to a more elliptical current possessing a counterclockwise component that is observable above noise levels or simply more stable in phase than during summer. The small change in panel (3.10c) relative to summer indicates a small shift in the angle between counter-rotating components at time 0.

3.4.3   Difference in Angle between Wind and Current Vectors

The seasonally-averaged angle $\Delta\theta(t_{\max})$ between the diurnal current and wind vectors at the hour of maximum wind speed is shown in Figure 3.11, separated into summer (3.11a) and winter (3.11b). These were obtained by stepping through 24 hours in 15-minute intervals and calculating wind speed and direction from diurnal harmonic coefficients, recording the time $t_{\max}$ at which wind is aligned with its semi-major axis, and finding the angle between $\vec{U}(t_{\max})$ and $\vec{u}(t_{\max})$. Near the coast in several regions, $\Delta\theta(t_{\max})$ is closer to 0° and increases with offshore distance. For all locations and both seasons, the circular mean deflection angle is +76.6° (σ = 38.7° after centering the distribution). Less than 31 km (one wind grid cell width) from the coast, this angle is +52.6° (σ = 39.3°), between 31 and 62 km it is +69.4° (σ = 31.0°), and farther than 62 km it is +90.3° (σ = 35.7°).

This pattern of offshore increase in $\Delta\theta(t_{\max})$ suggests a physical mechanism related to offshore distance, such as diurnal wind intensity, which is coastally intensified. If $\Delta\theta(t_{\max})$ were a function of the coastally intensified diurnal wind intensity, for example, it might result in a distribution of $\Delta\theta(t_{\max})$ that varies with distance from the coast. Figure 3.12 shows the

relationship between the deflection angle and the maximum speed of wind and current, combining summer and winter data. The bivariate histograms of $\Delta\theta(t_{max})$ versus peak diurnal wind speed and peak diurnal current speed are displayed in panels (3.12a) and (3.12b) respectively. The peak in both histograms around $\Delta\theta(t_{max}) = +90°$ corresponds to low amplitude winds and currents farther from shore, while at higher wind and current speeds, $\Delta\theta(t_{max})$ trends broadly toward $+45°$, the deflection angle predicted in the Northern Hemisphere for classical Ekman theory (Ekman, 1905; W. Wang & Huang, 2004), plotted here as a vertical red line. Note that we define the angle of deflection as $\Delta\theta = \theta_{wind} - \theta_{current}$ such that deflection to the right of the current is positive in order to be consistent with the sign of ϕ as we have defined it.

For peak diurnal wind speed > 1 m s^{-1}, the circular mean deflection angle is $+47.0°$ ($\sigma = 34.9°$ after centering) and for peak diurnal wind speed > 2 m s^{-1}, $+41.5°$ ($\sigma = 12.4°$). For peak diurnal current speed > 0.05 m s^{-1}, the circular mean deflection angle is $+54.8°$ ($\sigma = 34.3°$) and for peak diurnal current speed > 0.1 m s^{-1}, $+38.2°$ ($\sigma = 35.6°$).

Figure 3.13 shows the (top) summer and (bottom) winter deflection angles of surface currents from wind direction across fitted frequencies. The distribution of $\Delta\theta(\omega, t_{max})$ is calculated for each fitted frequency using a procedure analogous to that used for the diurnal frequency: wind amplitude is calculated from harmonic coefficients at 100 evenly-spaced times across the period $T = 2\pi/\omega$ to find the time of peak wind t_{max}, from which the angle $\Delta\theta(\omega, t_{max})$ between $\vec{U}(\omega, t_{max})$ and $\vec{u}(\omega, t_{max})$ is found. The annual frequency is not displayed, as its inclusion served only to eliminate the partially resolved annual cycle that would otherwise be treated as a trend in the 6-month analysis segments.

Figure 3.13 omits all locations with insignificant coherence as well as frequencies for which greater than 90% of locations have insignificant coherence, defined as when the average of γ^2_{uu} and γ^2_{vv} for the respective season is less than the 95% confidence threshold (see equation (3.15)); parallel Cartesian coherence is chosen for this as it is agnostic toward rotational relationships and therefore should not over-represent already dominant clockwise-clockwise interactions. An amplitude threshold was tested as an additional criterion for inclusion but was

not found to have a large effect except at the diurnal frequency, which was shown previously to converge toward a +45° deflection angle at high wind and current speeds. The inner bar corresponding to the interquartile range of the diurnal mean deflection angle is shaded blue in order to distinguish it from points at nearby frequencies. Light gray shading indicates the range of inertial frequencies in the study region, while dark gray indicates the phase prediction from the "best" oceanic boundary layer model by Elipot & Gille (2009) within these latitudes and evaluated at the surface, using K_0(summer) = 0.007 m^2 s^{-1} and K_0(winter) = 0.0042 m^2 s^{-1}, which were found to result in the smallest squared difference (for frequencies between 0.033 and 1.2 cpd) between the median (black) points and the "best" model phase evaluated at the mean latitude of the domain (39.3°N). Winter median angles were found to fit the theoretical curve more closely than summer median angles. The other model parameter h = 35 m was taken from Elipot & Gille (2009) as the best estimate for latitudes between 30°S and 38°S, while ρ = 1027 kg m^{-3}. The bounding curves for model phase are vertically offset by ±2° for visual clarity, as they both nearly equal ±45° away from the inertial band.

3.5   Discussion

The coastally-intensified diurnal sea breeze/land breeze from ERA5 agrees with theoretical descriptions of sea breeze, which predict surface wind speeds to decrease away from the coast (e.g. Walsh, 1974), and appears to be stable in phase over decades: Round (1993) found an average time of maximum wind occurrence of 1400 PST for the six-month period ending 30 September 1992 at Fort Ord, California, which falls in the grid cell for which we found 1418 PST (1411 local mean time) for maximum diurnal wind speed averaging over April through September from 2012 through 2020 (note that we use local mean time throughout this study and only use PST here for comparison). The reduction of nearshore diurnal wind amplitude during the winter is expected from the reduced temperature difference between the land and sea surface, which drives the sea breeze process.

The hour of maximum diurnal wind shifts moderately from summer to winter: within the respective seasonal offshore decay scales, winter peak diurnal winds are shifted 53 minutes later on average, which is within one standard deviation of maximum wind time for both seasons. The hour of maximum diurnal current (Figure 3.6c,d), however, exhibits greater seasonal differences, with a more complicated distribution in summer than in winter. This may be explained by the increase in ellipse eccentricity in the winter, with 68% of locations showing an increase over summer values. Additionally, the elliptical tilt roughly aligns across locations in winter, while the nearly circular diurnal ellipses in summer south of 36°N are slightly elongated in different directions such that the elliptical tilt and hence hour of maximum current varies substantially. Where diurnal currents are nearly circular, the importance of the overpass timing of a Sun-synchronous satellite observing ocean currents is reduced, as the difference between maximum and minimum current speed is small. North of 36°N, however, summer current ellipses are more eccentric and are aligned in a variety of directions, while in winter they tend to have alongshore alignment. This complicates strategies for optimizing Sun-synchronous satellite overpass times, as the seasonal rotation and distortion of current ellipses would appear as a signal at the annual frequency. The winter peak current hour does not appear to change substantially with distance from shore aside from a few regions (the southern Oregon Coast, near Cape Mendocino, south of San Francisco Bay, and near San Diego); this contrasts with the offshore-propagating winter diurnal wind, which presumably drives much of the diurnal current variability.

Diurnal currents off the US West Coast are highly coherent with 10-meter winds across 9 summers and winters, with both data sets characterized by strongly clockwise (anticyclonic) polarization. These results are consistent with Pidgeon & Winant (2005), who found clockwise-dominated surface currents coherent with wind stress in the coastal Santa Barbara Channel–Santa Maria Basin region of Southern California. The observed coherence of clockwise currents with counterclockwise winds is likely the result of the latter's stability and more eccentric elliptical diurnal cycle, which results in the counterclockwise component having a stable phase related to that of the dominant clockwise component. The mean point-by-point (nearest-neighbor) ratio of

diurnal peak current to wind speed throughout the domain in summer is 0.063 ($\sigma = 0.024$) and in winter is 0.085 ($\sigma = 0.035$). This overlaps with but is mostly higher than the 3%-5% range seen in previous studies for steady winds (Kim et al., 2009), which is a friction velocity theoretically described by the square root of the product of a drag coefficient and the ratio of the densities of air and water (eq. 6.7 in Weber, 1983). This suggests that a frictional velocity alone may be inadequate for describing the diurnal response. At frequencies less than 0.5 cpd, the peak current to peak wind ratio is typically less than 2% except for frequencies less than ~0.02 cpd, at which the ratio ranges from 2% to 9% (not shown).

The coherence at the diurnal frequency does not appear to substantially decrease away from the coast except north of 39°N; the drop-off in that region suggests that there is some variable offshore distance at which the two vector quantities lose coherence, as both become weaker. However, along most of the coast, such a distance is likely greater than 200 km, the farthest offshore HFR can resolve with at least 50% coverage. Alternatively, the observed coherence drop-off north of 39°N may be related to the observation network, as locations farther offshore are more difficult to observe.

The partitioning of time series into 9 or fewer discontinuous segments for coherence analysis appears to be sufficient to detect meaningful coherence: maps analogous to Figures 3.7-3.8 for the tidal frequencies O_1 and M_2 (not shown), which are powerful in currents but not winds, exhibit low coherence as expected. At the S_2 (semidiurnal) frequency there is partial coherence with complicated spatial structure, which is expected due to the contributions of both wind-driven and tidal currents.

The phase $\phi_{\cup\cup}$ between clockwise winds and clockwise currents (panel (a) in Figures 3.9 and 3.10) reasonably predicts the angle $\Delta\theta(t_{max})$ between the two vectors at the hour of peak wind speed as shown in Figure 3.11, after converting from hours to degrees. This is expected, as the dominant polarization of both wind and current is clockwise, and in the limit of completely clockwise motion of both processes, $\Delta\theta(t_{max})$ and $\phi_{\cup\cup}$ will be equal. The relationship between $\Delta\theta(t_{max})$ and $\phi_{\cup\cup}$ is similar regardless of season: in summer, the circular mean difference

between $\Delta\theta(t_{\max})$ and $\phi_{\cup\cup}$ is 4.4° (σ = 23.8° after centering the distribution), while in winter it is -7.9° (σ = 22.7°).

The inhomogeneous distribution of diurnal deflection angles suggests more complicated physics than the simple assumptions of the steady-state Ekman model (Ekman, 1905) and related models (e.g. Elipot & Gille, 2009). Regions where diurnal currents and/or winds are highly energetic, especially the Southern California Bight, exhibit deflection angles approximating the +45° expected from these models; however, for weaker winds and currents, +90° predominates with a substantial range of angles, with 75% of locations having a deflection angle between +45° and +135°. It is not clear if this positive deviation in the deflection angle is due to a difference between offshore and nearshore depth-varying vertical eddy viscosity coefficients (Bressan & Constantin, 2019), a previously observed diurnal cycle in the deflection angle (Krauss, 1993), or the effect of wave-induced Coriolis–Stokes forcing, which rotates the surface Ekman current to be greater than +45° from the mean wind stress direction (Polton et al., 2005). In support of the Coriolis–Stokes forcing as a factor in this deflection angle is the relatively strong southward Stokes drift off the US West Coast (McWilliams & Restrepo, 1999; Onink et al., 2019), with an offshore increase in depth-integrated wave-driven transport relative to nearshore values in this region (Polton et al., 2005). An ocean response model with a wind stress amplitude-dependent deflection angle could also explain the pattern in Figure 3.12, but this is beyond the scope of this study. Additionally, radiative forcing of the upper ocean at the diurnal frequency may lead to a response that substantially deviates from the dominantly wind-driven response expected at other frequencies and corroborated by the larger deviation from +45° observed at 1 cpd compared to lower frequencies in Figure 3.13.

The deviation from Ekman theory of the response of diurnal currents to diurnal winds is not unexpected for this high-frequency component, as Ekman theory is based on the assumption of steady winds. Comparison of our results to the best-performing frequency domain model in Elipot & Gille (2009) shows a frequency-dependent mismatch in deflection angle, in which the observed low-frequency response is characterized by a small deflection angle between 0° and

+45° that increases with frequency, matching +45° at approximately 0.1 cpd and continuing to increase as the inertial frequency is approached. This contrasts with the global observations of Rio & Hernandez (2003), who used drifting buoys, altimetry, and an earlier ECMWF wind field and found a frequency-dependent response that more closely matches the expectations of the theoretical curve in Figure 3.13 for observations at latitudes poleward of 30°. However, in the CCS region (which was more thoroughly sampled than other regions in that study), Rio & Hernandez (2003) found a larger average deflection angle of +57.6° (negative by the conventions of that study) between ageostrophic velocities and wind stress, and a "vectorial coherency" less than 0.3; during summer, their grid cells corresponding to the CCS were in the 60°-65° range, while in winter they were in the 45°-55° range. At low frequencies, we typically find coherence equal to or greater than 0.3, and all locations contributing to Figure 3.13 exhibit coherence above the 95% confidence threshold, which is 0.312 for 9 seasons.

Rio & Hernandez (2003) explain this tendency toward latitude- and seasonally-dependent deflection angles by the constant depth of drifter measurements (15 m) corresponding to different levels within the Ekman layer, the thickness of which is related to stratification. HFR should not have this issue, as it observes the top 1 m, yet summer deflection angles are still observed to deviate more positively from +45° than winter deflection angles for frequencies greater than 0.1 cpd. Though the drifter depth explanation likely accounts for much of this variability in deflection angle globally, the > 45° values we observe and the similar seasonal modulation at the surface merit additional study.

In general, surface deflection angles have been previously observed to range from 20° to greater than 45° (Bressan & Constantin, 2019), with low-latitude regions experiencing larger average deflection angles and high-latitude regions experiencing values closer to 45°, albeit with increased angles observed during the summer, up to 90° (Rio & Hernandez, 2003). Our observations are in broad agreement with past studies, indicating the potential utility of HFR for evaluating models of upper ocean response to wind forcing and estimating the vertical eddy viscosity coefficient, which may be complex (Price & Sundermeyer, 1999; Lenn & Chereskin,

2009).

3.5.1 Implications for Satellite Orbits

The more uniformly timed maximum diurnal current in winter and the more elliptical diurnal current in summer suggest that if the California Current System is assumed to be globally representative, then sun-synchronous orbital planning should take the winter phase of diurnal currents into greater consideration than the summer phase, despite the lower amplitude of the winter signal. Our results suggest that an overpass between 0400 and 0500 (or 1600-1700) local mean time would best observe diurnal currents at or near their peak amplitude, at least in this region. This time window also samples the wind at its peak magnitude in the coastal region, where its amplitude is greatest, albeit within only 30-90 km of the coast; this is convenient for the proposed ODYSEA mission, which would observe both winds and currents simultaneously (Rodríguez et al., 2019). Though this choice of overpass time may be primarily useful in coastal regions like those examined here, the attenuation of the sea breeze over the open ocean may justify prioritizing sampling for coastal diurnal currents. Additionally, this domain is entirely poleward of 30°N, where the inertial frequency is greater than the diurnal frequency. Equatorward of 30°, differences in the ocean's response as well as the characteristics of diurnal winds will likely need to be taken into account for orbital design (linear theory of sea breeze predicts different results for superinertial versus subinertial regions, e.g. Rotunno (1983)).

3.6   Summary and Conclusions

In this study, we described the magnitude, timing, and seasonality of diurnal winds and currents within about 200 km of the coastlines of California and Oregon. We found that both quantities are intensified near the coast, with typically greater amplitudes in summer compared to winter. Diurnal winds and currents were found to be highly coherent, especially the dominant clockwise components of each. We found that the angle between wind and current vectors differs from the 45° angle expected from steady-state Ekman theory, which is not unexpected

for a non-steady wind. The angle between wind and current vectors deviates from this value at most frequencies, which we discussed in the context of previous studies that investigated air-sea interaction beyond Ekman theory.

Previous studies (e.g. Paduan & Rosenfeld, 1996; Kosro, 2005; Kohut et al., 2006; Kim et al., 2009; Cosoli et al., 2020) have examined currents from HFR and their response to winds over shorter time periods and smaller regions than those of the present study. In our larger domain, we discussed the implications of our findings for designing sun-synchronous orbits of satellite scatterometers that would observe surface ocean currents and winds simultaneously and found that an overpass at a local mean time between 0400 and 0500 or 1600-1700 would typically sample both winds and currents at or near their maximum amplitude along their elliptical diurnal cycles. The nearly circular summer current merits less attention toward orbital timing for that quantity in this region, though winter currents and year-round winds follow sufficiently elliptical paths such that orbital timing may need to consider the diurnal cycle. Winds and currents are highly coherent at the diurnal frequency, especially for co-rotating clockwise components, partially coherent at the semidiurnal frequency, and incoherent at tidal frequencies as expected.

Though we do not investigate the annual modulation of the wind or currents beyond seasonal differences, this is an ongoing topic of investigation (e.g. Giglio et al., 2022) due to its relevance to climate records from satellite wind observations and implications for the understanding of air-sea interaction. The seasonal differences of quantities estimated in this study, such as the peak amplitudes and hours of diurnal currents and winds, the structure of coherence, and the deflection angle of currents, all highlight the importance of the annual cycle for air-sea interaction, meriting further investigation.

The deflection angle of current vectors from simultaneous wind vectors was somewhat larger than expected, though not without precedent. The complicated physics of the diurnal cycle, which includes both wind-driven and radiative processes, may be responsible for its large deflection angle relative to other frequencies, at which wind forcing is expected to dominate. The spatial resolution afforded by this analysis could help inform models of wind-driven surface

109

currents that are more complicated than those examined here and may refine estimates for vertical eddy viscosity.

Acknowledgements

3.7   Appendix for Chapter 3

3.7.1   Rotary Quantities

The zonal and meridional velocity components at times $\mathbf{t}$ can be expressed as the sum of weighted sinusoidal terms at M frequencies plus some residual term:

$$\mathbf{u} = \sum_{m=1}^{M} \left(a_m \cos(\omega_m \mathbf{t}) + b_m \sin(\omega_m \mathbf{t}) \right) + \mathbf{r}_u \tag{3.3}$$

$$\mathbf{v} = \sum_{m=1}^{M} \left(c_m \cos(\omega_m \mathbf{t}) + d_m \sin(\omega_m \mathbf{t}) \right) + \mathbf{r}_v, \tag{3.4}$$

where boldface indicates column vectors. The coefficients of the sinusoids may be estimated using a least squares fit and may be written as complex vectors corresponding to each velocity component:

$$\tilde{\mathbf{u}} = [a_1 + \mathrm{i}b_1, a_2 + \mathrm{i}b_2, ..., a_M + \mathrm{i}b_M]^{\mathrm{T}} \tag{3.5}$$

$$\tilde{\mathbf{v}} = [c_1 + \mathrm{i}d_1, c_2 + \mathrm{i}d_2, ..., c_M + \mathrm{i}d_M]^{\mathrm{T}}. \tag{3.6}$$

Vector processes that follow elliptical paths, such as wind or currents at a single frequency, are conveniently described by rotary components with complex clockwise (↻) and counterclockwise (↺) amplitudes. This formulation is therefore rotationally invariant:

$$\mathbf{u} + \mathrm{i}\mathbf{v} = \sum_{m=1}^{M} \left(R_{\circlearrowleft,m} e^{\mathrm{i}\omega_m \mathbf{t}} + R_{\circlearrowright,m} e^{-\mathrm{i}\omega_m \mathbf{t}} \right) + \mathbf{r}_{\mathrm{rot}} \tag{3.7}$$

where $\mathbf{r}_{\mathrm{rot}} = \mathbf{r}_u + \mathrm{i}\mathbf{r}_v$ is the complex residual. Following Gonella (1972), we can define the clockwise and counterclockwise amplitudes in terms of the sine and cosine coefficients of u and v:

$$R_{\circlearrowleft,m} = \frac{1}{2} (a_m - d_m + \mathrm{i}[c_m + b_m]) \tag{3.8}$$

$$R_{\circlearrowright,m} = \frac{1}{2} (a_m + d_m + \mathrm{i}[c_m - b_m]), \tag{3.9}$$

3.7.2 Coherence

A least squares analogue to the cross-spectrum may be defined for two time series $\alpha(t)$ and $\beta(t)$ with coefficients $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\boldsymbol{\beta}}$:

$$\mathbf{S}_{\alpha\beta} = \left\langle \tilde{\boldsymbol{\alpha}} \tilde{\boldsymbol{\beta}}^* \right\rangle, \tag{3.10}$$

where $*$ indicates complex conjugation and $\langle \rangle$ indicates an average over some number of realizations of coefficients corresponding to segments of the times series. Because these segments

are defined by seasons, they are noncontinuous and are not tapered; this is known as the "chunk" spectral estimator with equivalent degrees of freedom equal to twice the number of segments (von Storch & Zwiers, 2003). Spectral normalization is omitted due to the use of this quantity only in ratios in this study. A least squares analogue to the squared coherence of one Cartesian component of some quantity (e.g. wind) with the Cartesian component of another quantity (e.g. current) may then be defined as (e.g. Bendat & Piersol (2010)):

$$\gamma^2_{\alpha\beta} = \frac{\Re\left[\mathbf{S}_{\alpha\beta}\right]^2 + \Im\left[\mathbf{S}_{\alpha\beta}\right]^2}{\mathbf{S}_{\alpha\alpha}\mathbf{S}_{\beta\beta}} \tag{3.11}$$

where $\boldsymbol{\gamma}^2$ is M elements long, α is either the u or v component of the first quantity of interest, and β is either the u or v component of the second quantity of interest. The phase difference between these two components is

$$\boldsymbol{\phi}_{\alpha\beta} = \arctan\left(-\frac{\Im\left[\mathbf{S}_{\alpha\beta}\right]}{\Re\left[\mathbf{S}_{\alpha\beta}\right]}\right). \tag{3.12}$$

The squared coherence and phase of one rotary component of one vector process with that of another vector process is analogous to the Cartesian definition in equations (3.11)-(3.12) (e.g. Mooers (1973)):

$$\gamma^2_{\alpha\beta} = \frac{\Re\left[\langle\mathbf{R}_\alpha\mathbf{R}^*_\beta\rangle\right]^2 + \Im\left[\langle\mathbf{R}_\alpha\mathbf{R}^*_\beta\rangle\right]^2}{\langle\mathbf{R}_\alpha\mathbf{R}^*_\alpha\rangle\langle\mathbf{R}_\beta\mathbf{R}^*_\beta\rangle} \tag{3.13}$$

$$\boldsymbol{\phi}_{\alpha\beta} = \arctan\left(-\frac{\Im\left[\langle\mathbf{R}_\alpha\mathbf{R}^*_\beta\rangle\right]}{\Re\left[\langle\mathbf{R}_\alpha\mathbf{R}^*_\beta\rangle\right]}\right) \tag{3.14}$$

where $\mathbf{R}$ are M-element vectors of rotary coefficients at each frequency, α represents a rotary component (clockwise or counterclockwise) of the first vector and β is a rotary component of the second vector.

3.7.3 Coherence Critical Values

In order to determine the significance of coherence estimates estimated from the least squares coefficients of N segments, we use the critical value for the $\alpha \times 100\%$ significance level:

$$\gamma^2_{\text{crit}}\bigg|_\alpha = 1 - \alpha^{\frac{1}{N-1}}, \tag{3.15}$$

which corresponds to the value above which $\alpha \times 100\%$ of coherence values of two incoherent processes are expected to fall.

Equation (3.15) has been used at least as early as the 1950s (Goodman, 1957), but it is not clear in many sources why this formula is used. Following from assumptions about the distribution of coherence, von Storch & Zwiers (2003) define the critical value for the $\alpha \times 100\%$ significance level as:

$$\gamma^2_{\text{crit}} = \frac{2F_{\tilde{p}}}{r - 2 + 2F_{\tilde{p}}}, \tag{3.16}$$

where r is the effective degrees of freedom, $\tilde{p} = 1 - \alpha$ (we use $\tilde{p} = 0.95$), and $F_{\tilde{p}}$ is the $\tilde{p}$-critical value of the F-distribution with degrees of freedom 2 and $r - 2$. In this study, the simple chunk method is used and $r = 2N$, thus $F(F_{\tilde{p}}; 2, 2N - 2) = 1 - \tilde{p}$, where F is the cumulative distribution function of the F-distribution [NIST, 2012]:

$$F(x; d_1, d_2) = 1 - \frac{\int_0^{d_1 x/(d_1 x + d_2)} t^{\frac{d_1}{2}-1}(1-t)^{\frac{d_2}{2}-1}dt}{\int_0^1 t^{\frac{d_1}{2}-1}(1-t)^{\frac{d_2}{2}-1}dt} \tag{3.17}$$

$$F(F_{\tilde{p}}; 2, 2N - 2) = 1 - \frac{\int_0^{F_{\tilde{p}}/(F_{\tilde{p}}+N-1)}(1-t)^{N-2}dt}{\int_0^1 (1-t)^{N-2}dt} \tag{3.18}$$

$$= \left(\frac{N-1}{F_{\tilde{p}}+N-1}\right)^{N-1}, \tag{3.19}$$

where the last step is simplified using Mathematica software. Setting this equal to $1 - \tilde{p}$ and

solving for $F_{\tilde{p}}$:

$$F_{\tilde{p}} = (N - 1)\left((1 - \tilde{p})^{\frac{1}{1-N}} - 1\right).\tag{3.20}$$

This may then be substituted into equation (3.16) and simplified, giving equation (3.15):

$$\gamma_{\text{crit}}^2 = 1 - (1 - \tilde{p})^{\frac{1}{N-1}}\tag{3.21}$$

Note that we use the term "95% confidence threshold" to refer to the coherence critical value for the $\alpha \times 100\%$ significance level for $\alpha = 0.05$ (alternatively the $[1 - \alpha] \times 100\%$ confidence level), that is, the value above which we expect $\alpha \times 100\%$ of coherence values for incoherent processes to fall.

Chapter Three, in part, is currently being prepared for submission for publication of the material. Kachelein, L., S. T. Gille, M. R. Mazloff, and B. D. Cornuelle. The Diurnal Cycle in the California Current System: Currents from High-Frequency Radar and Winds from ERA5 (In Preparation, Preliminary Title). The dissertation author was the primary investigator and author of this material.
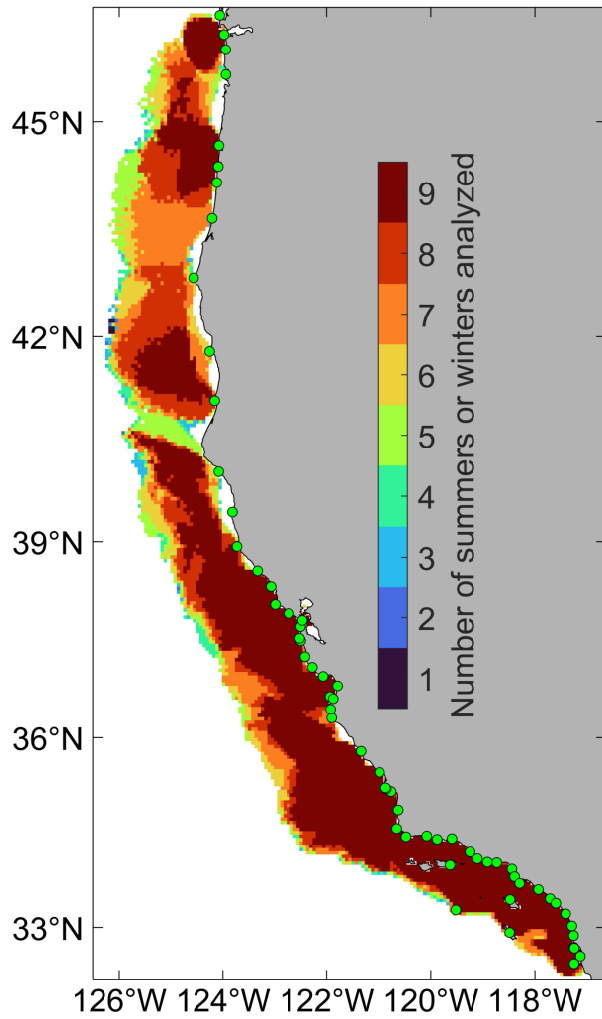
Figures

Figure 3.1. Number of summer (April-September) and winter (October-March) segments of high-frequency radar time series analyzed. Green circles indicate the locations of HFR network stations.
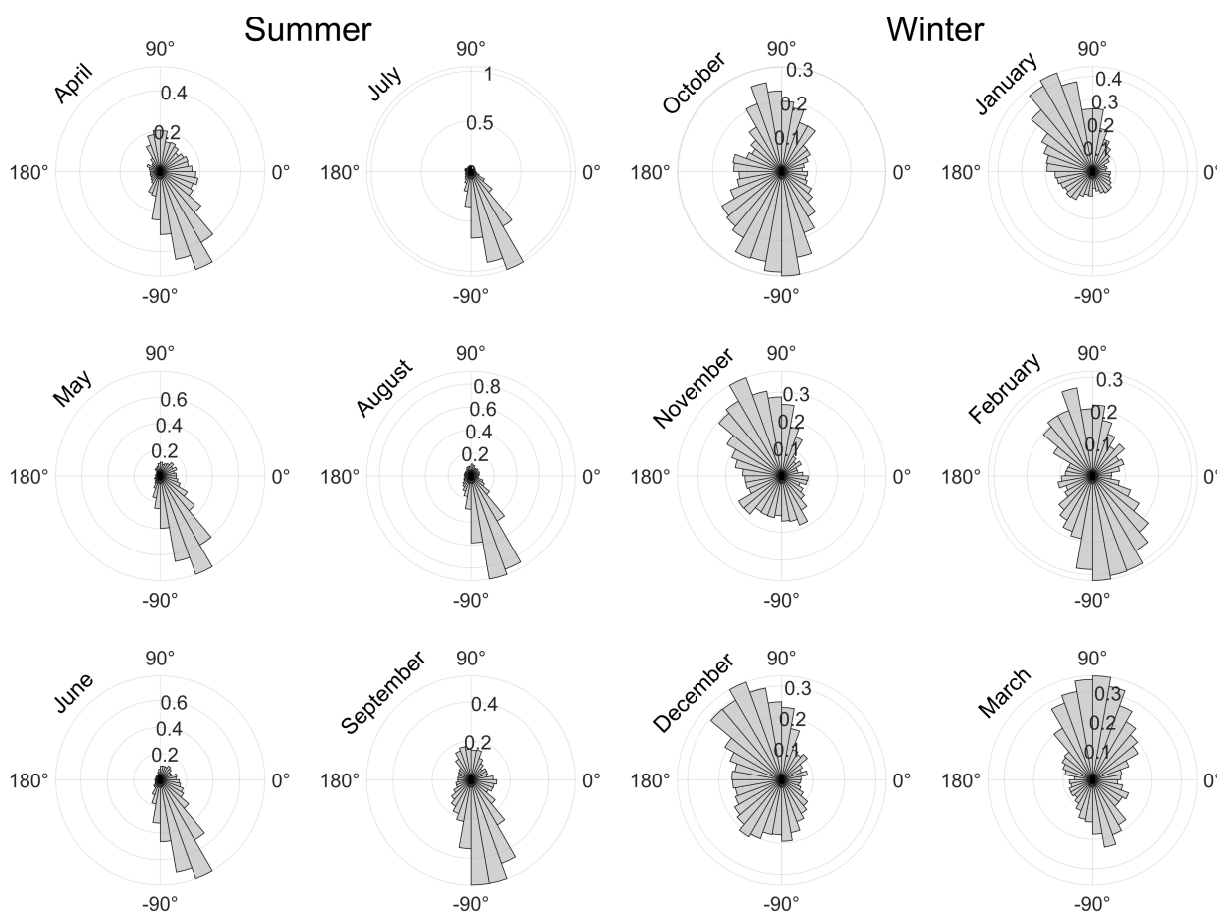
Figure 3.2. Wind roses displaying probability density functions of ERA5 reanalysis downwind angle within the HFR domain (see Figure 3.1) after removal of the 11-year mean wind vector, sorted by month, with 10° bin widths. Summer months, as defined for this study, range from April through September, while winter months range from October through March.
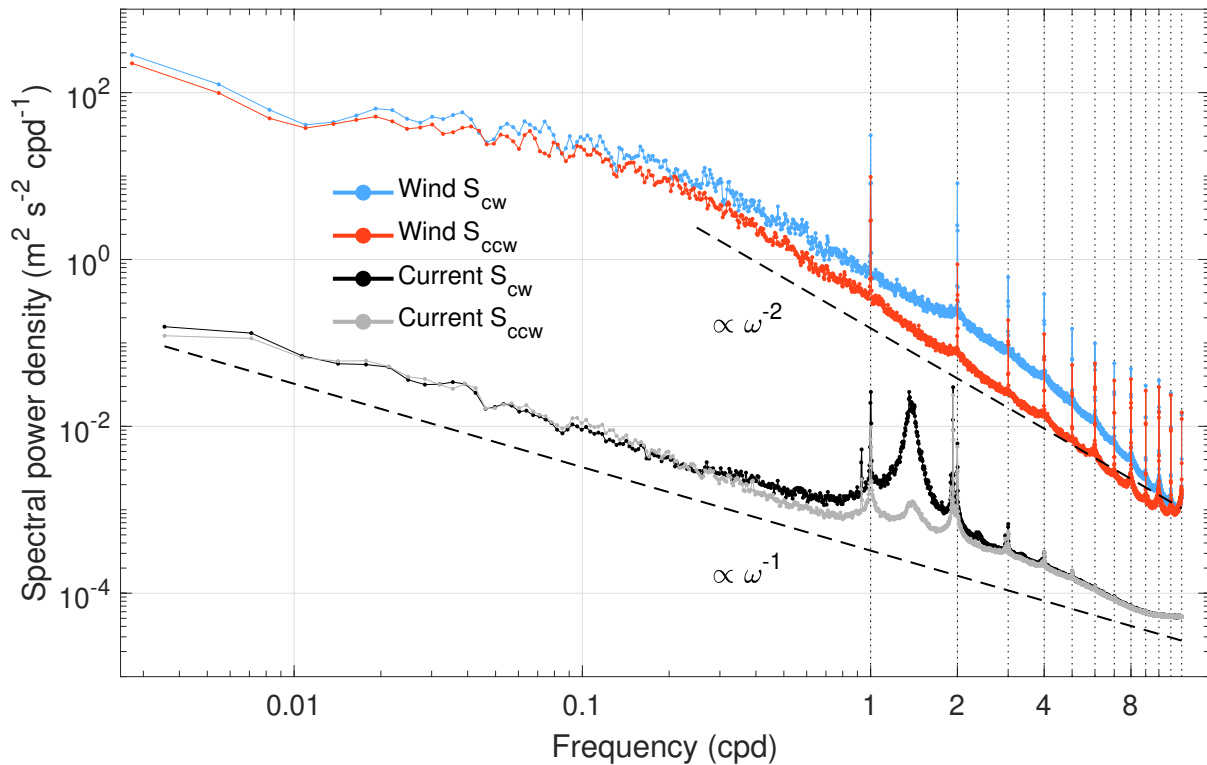
Figure 3.3. Rotary power spectral density of ERA5 10-meter winds over the ocean in the study domain, including beyond the range of HFR (colored curves) and surface currents observed by HFR between 41.6°N and 44.8°N (black and gray curves). The diurnal frequency and each of its harmonic frequencies up to 12 cpd are emphasized with black dotted lines, and the approximate spectra slopes of each data set near the diurnal frequency are indicated with black dashed lines.
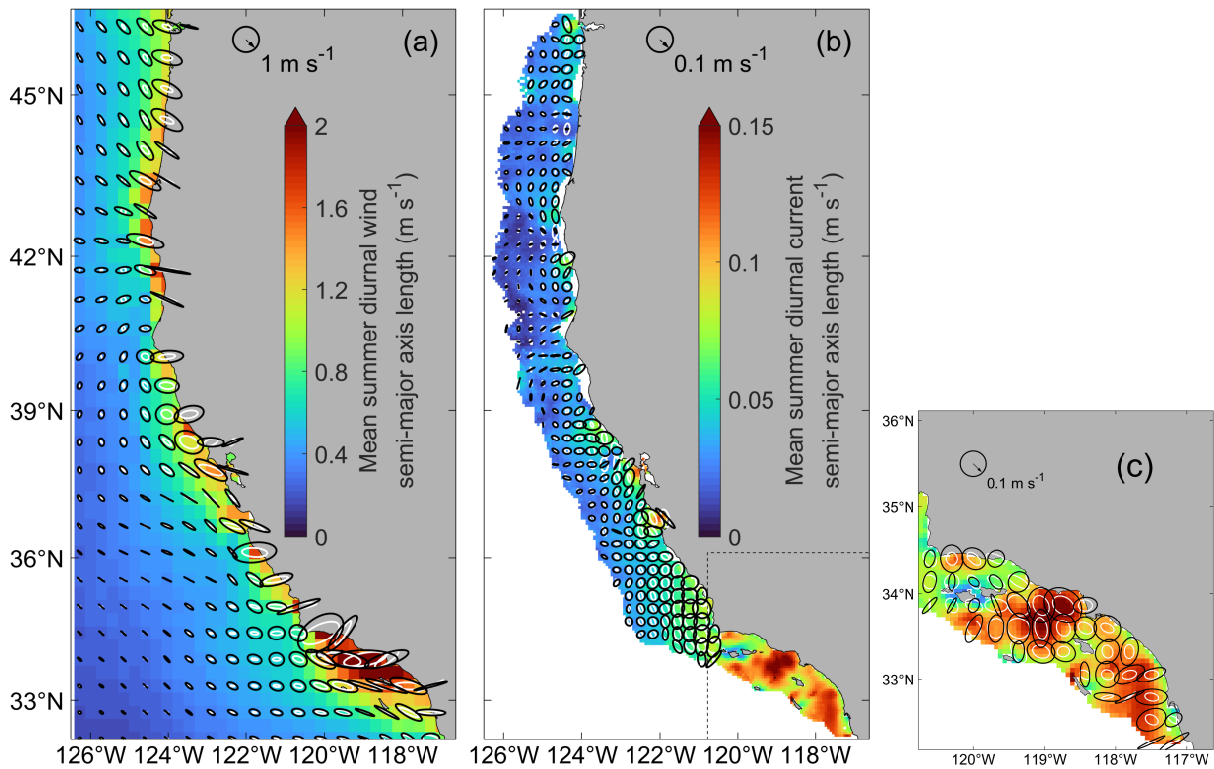
Figure 3.4. Diurnal wind and current ellipses differentiated by season. For all panels, color bars represent summer semi-major axis length, black ellipses correspond to summer, and white ellipses correspond to winter for comparison. (a) Diurnal winds from ERA5, (b) diurnal currents from HFR, and (c) HFR results in the Southern California Bight (box in panel b) magnified in order to prevent ellipse crowding while preserving scale. Displayed values of semi-major axis length and ellipses are calculated from the seasonal average of the four diurnal coefficients corresponding to cosine and sine of u and v.
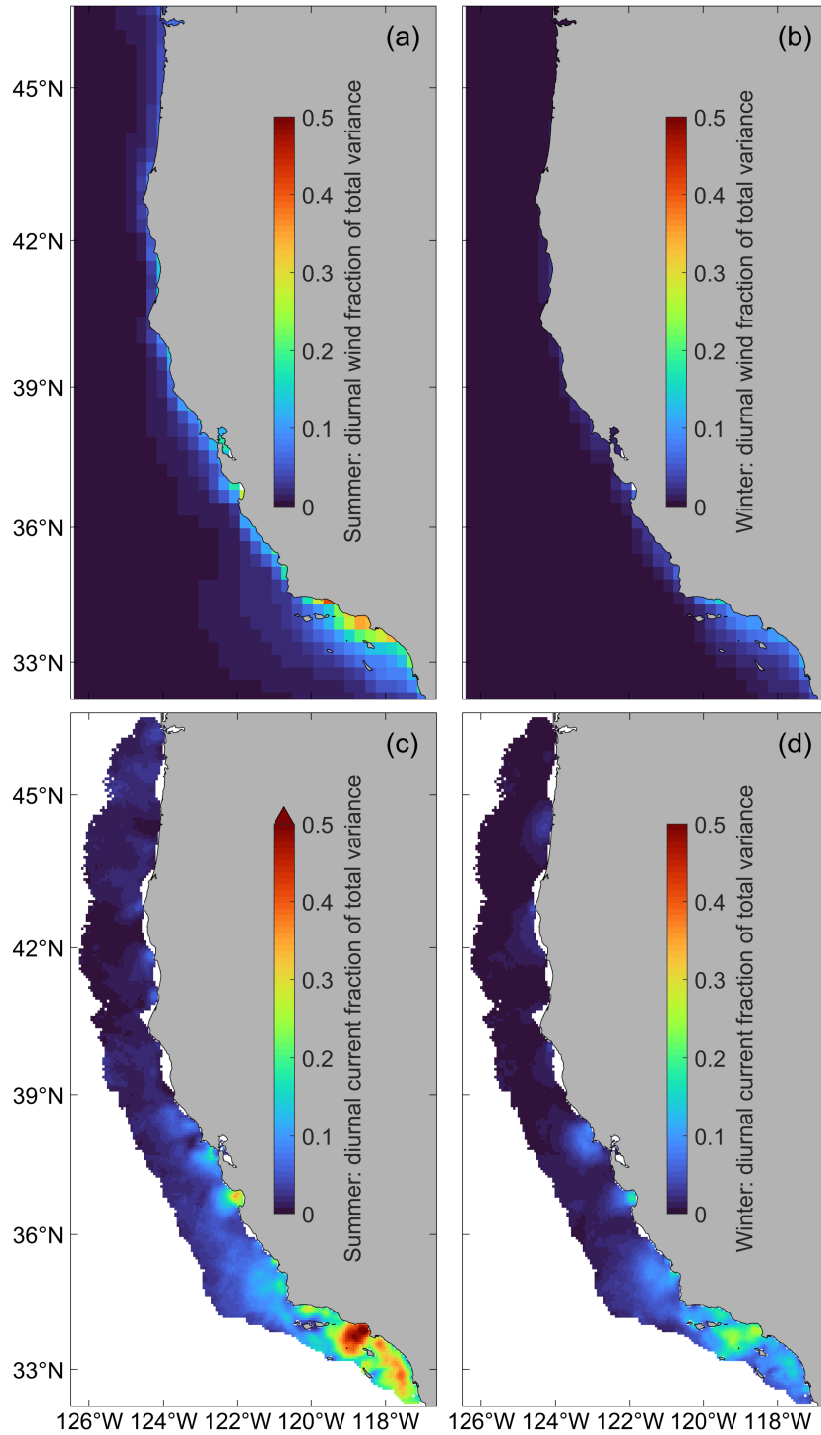
Figure 3.5. Seasonally-averaged fraction of variance at the diurnal frequency calculated via harmonic analysis for ERA5 winds (a,b) and HFR surface currents (c,d). The saturated color bar in panel (c) accounts for 6 grid cells for which diurnal variance comprised greater than half of total current variance, with a maximum value of 0.53.
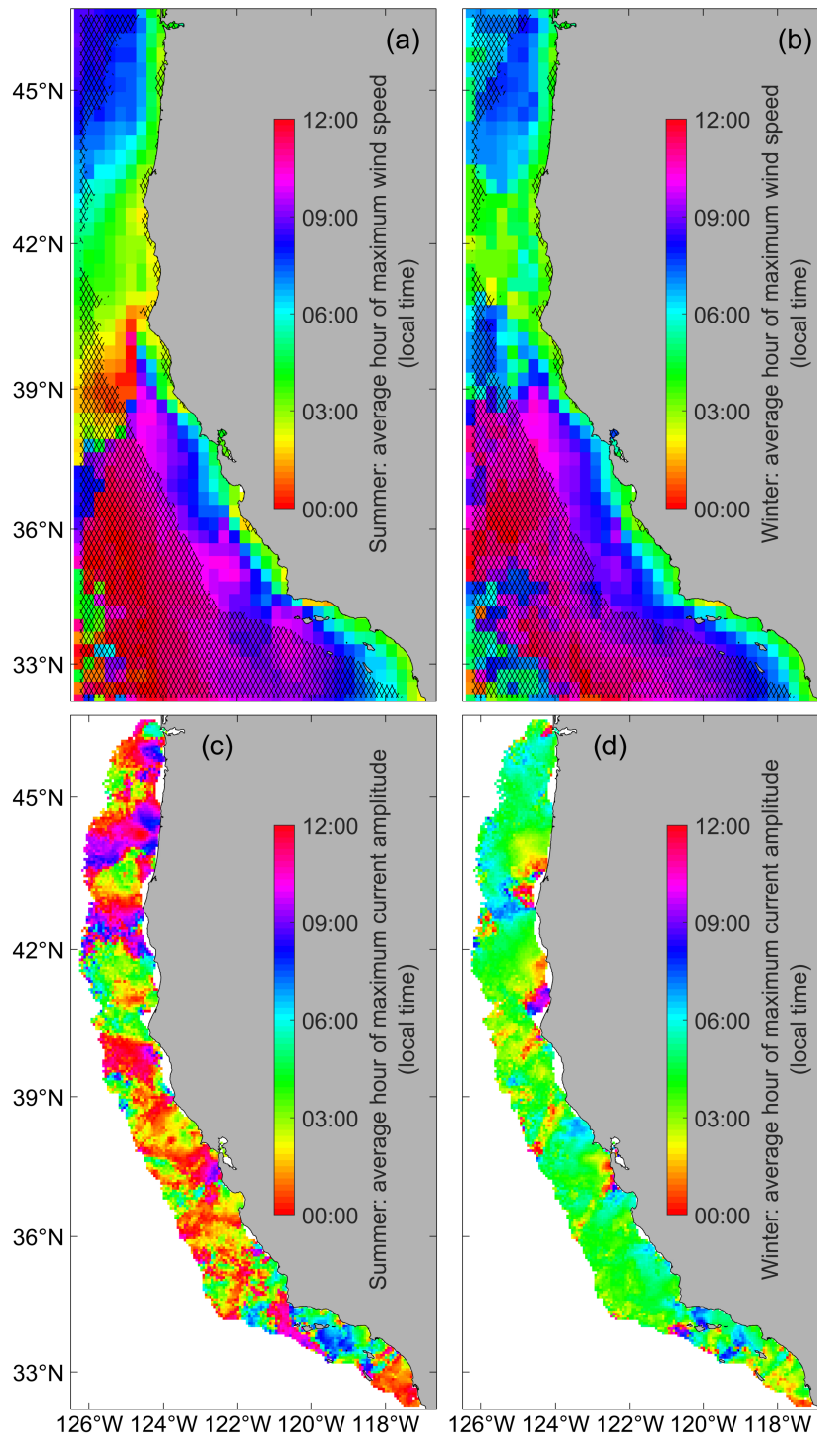
Figure 3.6. Average local time, twice daily, at which diurnal wind (a,b) and diurnal current (c,d) are at maximum amplitude in summer (a,c) and winter (b,d). Hatching in (a,b) indicates where HFR observations are absent.
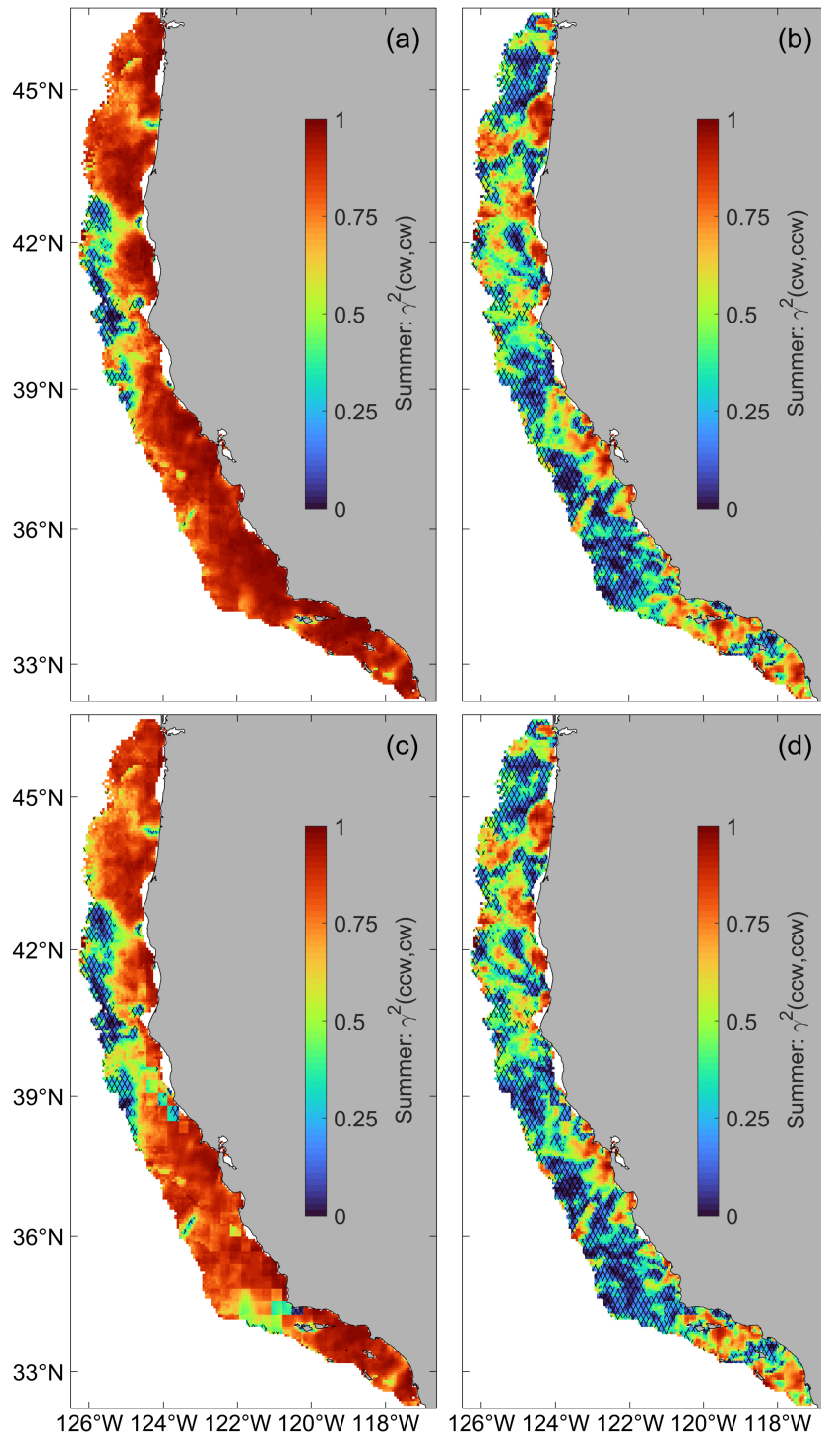
120

Figure 3.7. Rotary coherence γ^2 in summer between clockwise (cw) and counterclockwise (ccw) components of diurnal wind and diurnal surface current. The first argument in parentheses corresponds to wind polarization, the second argument corresponds to current polarization. Hatching indicates values of γ^2 less than the 95% confidence threshold, the value below which 95% of coherence estimates of incoherent processes are expected to fall according to equation (3.15).
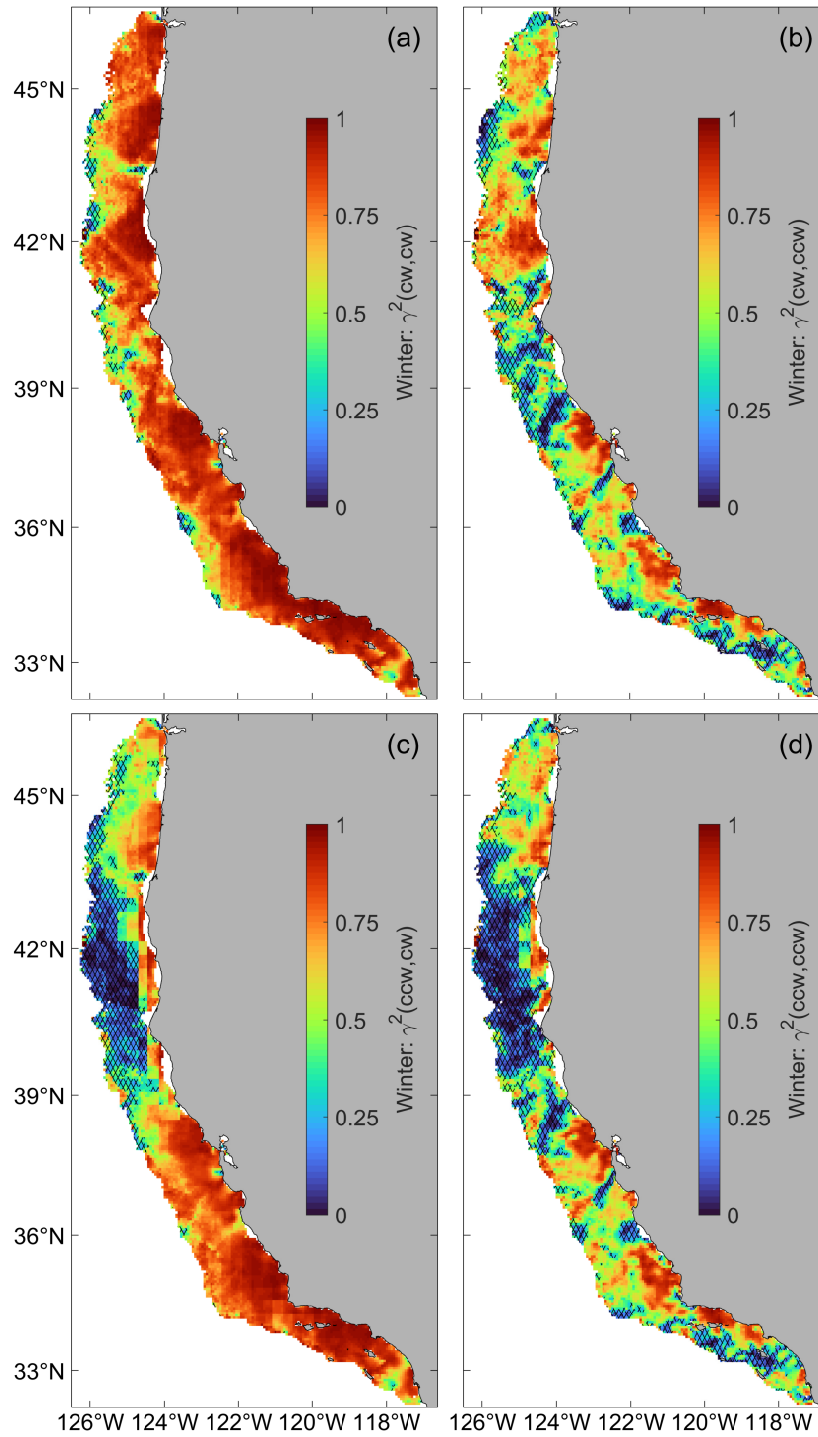
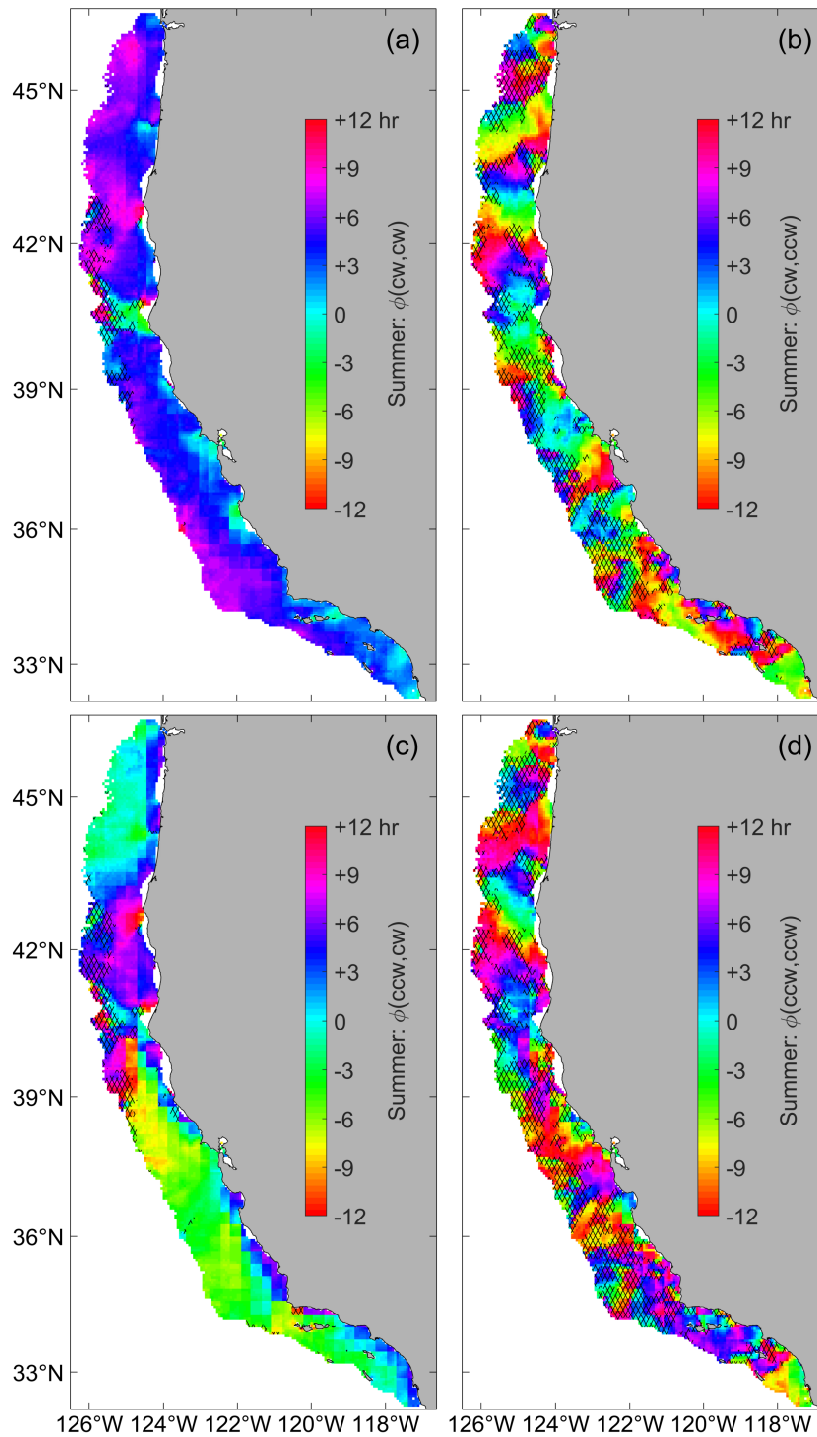**Figure 3.8.** As in Figure 3.7 but for winter months.

Figure 3.9. Phase ϕ of rotary coherence in summer between clockwise (cw) and counterclockwise (ccw) components of diurnal wind and diurnal surface current. Positive values indicate that the current is leading the wind. Hatching is the same as in Figure 3.7.
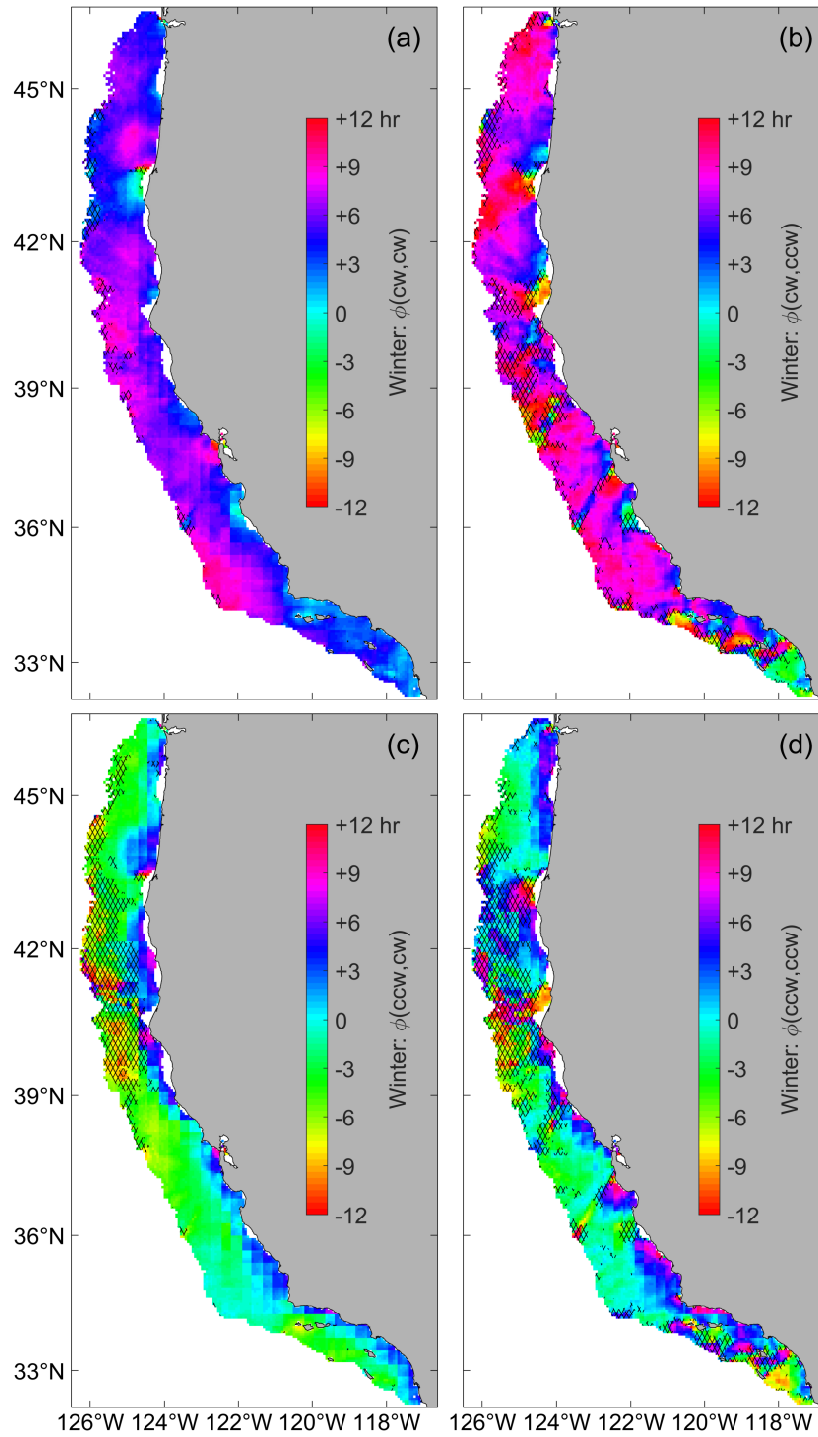
Figure 3.10. As in Figure 3.9 but for winter months.

124
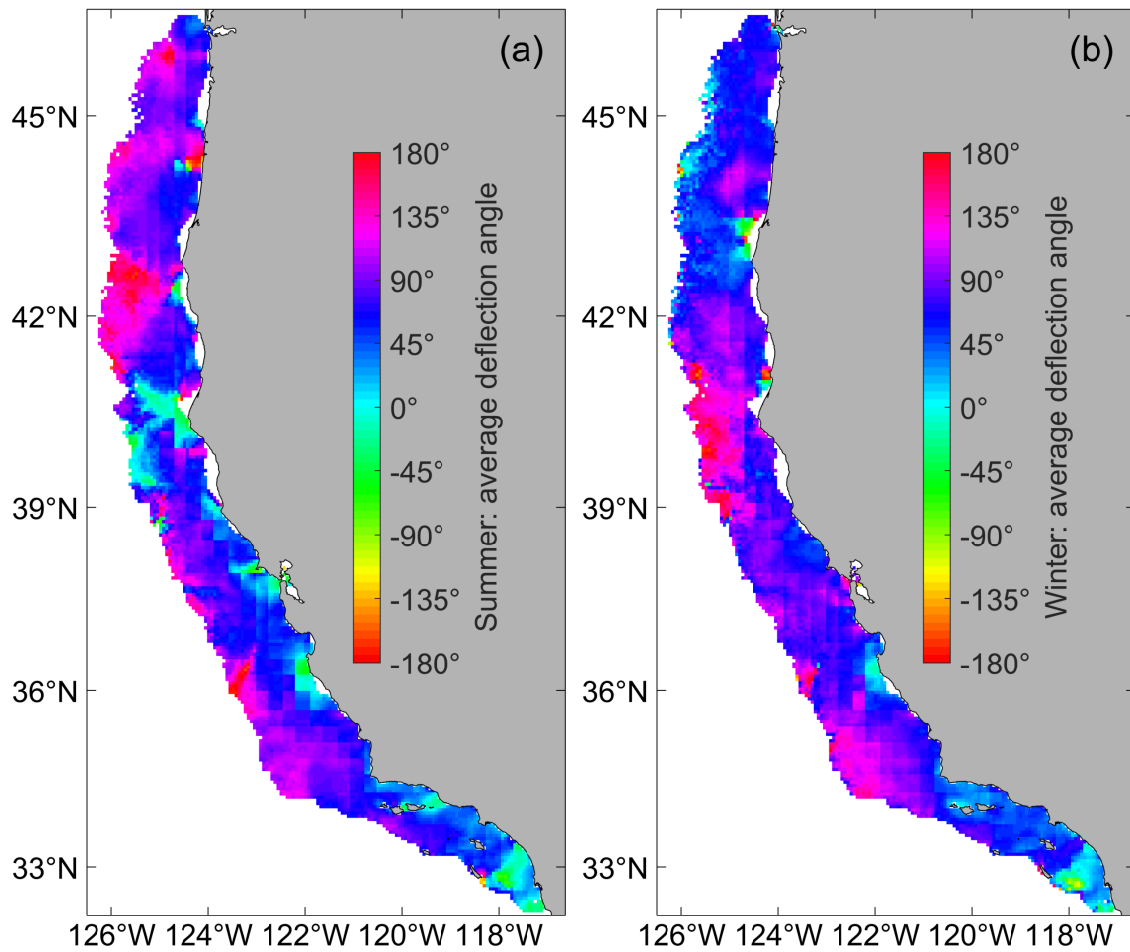
Figure 3.11. Summer (a) and winter (b) deflection angle of current at the time of peak wind speed. A positive deflection angle means the current is to the right from the perspective of the wind direction.

Figure 3.12. Bivariate histograms of combined summer and winter (a) diurnal wind semi-major axis and (b) diurnal current semi-major axis (peak diurnal value) versus deflection angle of current at the time of peak wind speed. Red lines indicate the +45° deflection to the right as predicted at the surface by classical Ekman theory.

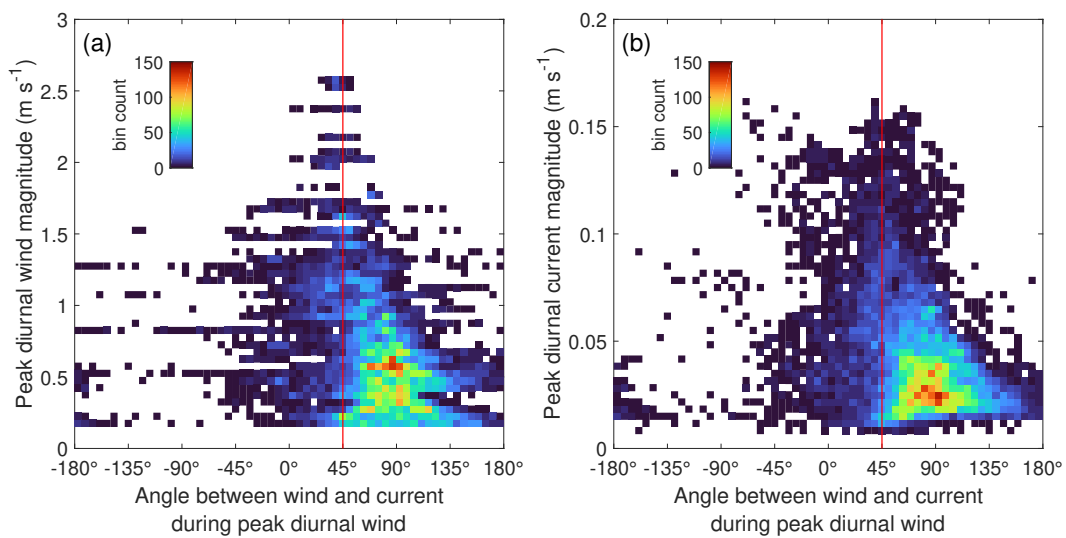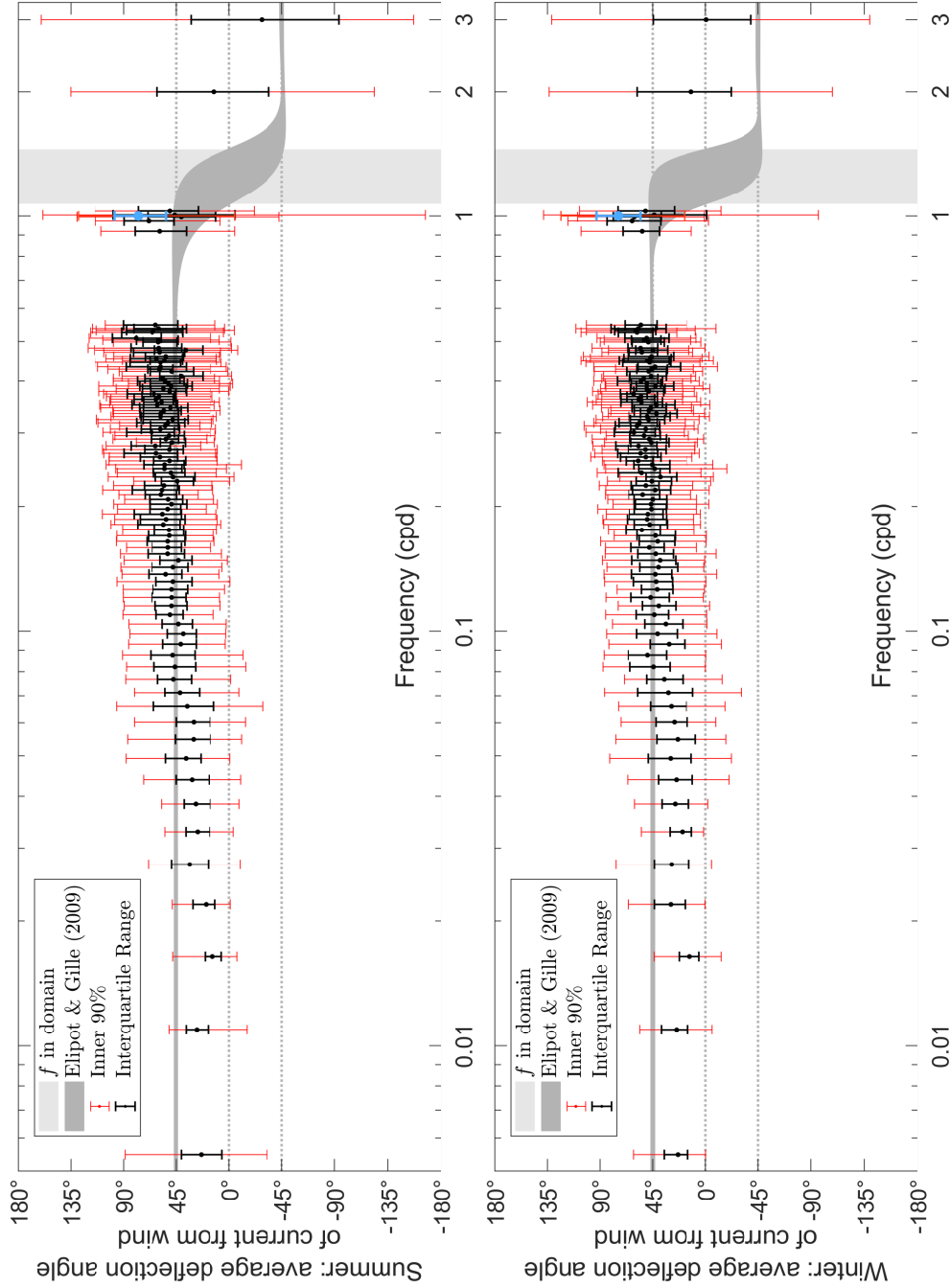**Figure 3.13.** Distributions of (top) summer and (bottom) winter angles of deflection of current from peak wind as a function of frequency. Black markers indicate the median angle, black inner bars indicate interquartile range, and red outer bars indicate the inner 90%. Positive angles indicate leading current. Locations do not contribute to the distributions if they have insignificant coherence (average of γ_{uu}^2 and γ_{vv}^2), and frequencies with < 10% of significantly coherent locations are not plotted. The inertial frequency range within the latitudes of the study is indicated in light gray, while dark gray indicates the phase prediction from the "best" oceanic boundary layer model in Elipot & Gille (2009) within these latitudes, using K_0(summer) = 0.007 m^2 s^{-1}, K_0(winter) = 0.0042 m^2 s^{-1}, h = 35 m, and ρ = 1027 kg m^{-3}, evaluated at the surface. The dark gray curves are vertically offset by $\pm 2°$ for visual clarity.

Summary and Concluding Statements

Diurnal and semidiurnal processes are important components of the ocean's energy budget. Tidal forces input energy into the ocean at known frequencies in a predictable way, but through the generation of internal tides and their interactions with other processes in the ocean, tidal signals are modulated and distorted in their evolution toward turbulence, becoming difficult to sample and interpret by observations with low sampling rates. By contrast, winds input energy into the ocean via wind stress at a range of frequencies, but due to the diurnal cycle in the wind, there is also a substantial diurnal component of surface currents which poses the same difficulty for sampling and interpretation. With the recent, planned, and proposed launches of high-resolution Earth observation satellites, it is more important than ever that these signals are understood in order to aid in the interpretation and analysis of satellite data, from both altimetry and scatterometry, and as a consequence is a topic of extensive study.

This dissertation has sought to contribute to this understanding by developing and proposing methods that are then applied to high-frequency radar observations, an extensive surface velocity data set that captures submesoscale processes. Additionally, we seek to add to the corpus of the California Current System, an exceptionally well-observed region of the global ocean. We outlined the methods in Chapter 1, applied them to characterize tides in Chapter 2, and applied them alongside rotary coherence analysis to examine diurnal wind-driven currents in Chapter 3. Beyond these results, there is still more that can be done with these methods and data sets.

The methods from Chapter 1 can be improved upon for user-friendliness. Though red_tide was written with usability in mind, it requires familiarity with the mathematics of Chapter 1 and

has several inputs that are non-trivial to specify, like the power spectrum, though the default white noise assumption and harmonic constituent frequencies as given by NOAA are enough to provide a cursory evaluation using red_tide. Additionally, translation from MATLAB to other programming languages, including non-proprietary ones, is desirable, and to this end, Chris Turnadge of the Commonwealth Scientific and Industrial Research Organisation of Australia has worked to translate red_tide into Python (github.com/christurnadge/redtidepy). For Chapter 2, a remaining question is how the correlated amplitude modulations of M_2 and S_2 signals arise. It is apparently related to stratification, upon which the propagation of internal waves depends, but the degree to which internal tidal amplitude modulation as observed by HFR can be predicted from stratification records is unknown. Additionally, the relevance of our surface current analysis to sea surface height needs to be fully evaluated for applications to altimetry. For Chapter 3, we only briefly examined the response of the ocean to wind as a function of frequency. A more thorough evaluation of this topic might help inform models of air-sea interaction, as could a time-domain response function approach. Another topic is the transfer function for winds and currents, the examination of which was omitted in the interest of time and focusing on characterizing the diurnal cycle.

Additionally, the energetic near-inertial oscillations seen in power spectra and fit explicitly in the analysis of Chapter 2 were not examined beyond confirmation that the energy within the near-inertial band follows the latitude-frequency relationship expected from theory. Because near-inertial oscillations were found to be most energetic at the offshore edge of our domain of HFR measurements, we chose not to focus on them in any of the chapters, but future work examining this process is a logical addition to the studies of Chapters 2 and 3, which examined different energetic processes in the same diurnal-to-semidiurnal frequency band. Finally, regional studies of HFR data at 1 km and 2 km spatial resolution could reveal processes that are not captured by the more extensive but coarser 6 km data set we examined in Chapters 2 and 3.

The motivation of this thesis was to inform satellite missions while investigating more fundamental oceanographic topics like the behavior of internal tides as they are modulated by

non-tidal processes, and the response of ocean surface currents to wind forcing. The HFR network, especially on the US West Coast, provides an extensive data set for investigating these topics, which are likely only a small subset of the questions HFR can help answer. As more Earth observation satellites are launched and as the HFR network continues to mature, coverage of the ocean's surface will improve and the energetic air-sea boundary will continue to be an increasingly important topic of investigation.

Bibliography

Agnew, D. C. (1992). The time-domain behavior of power-law noises. *Geophysical Research Letters*, *19*(4), 333-336. doi: https://doi.org/10.1029/91GL02832

Alford, M. H. (2003). Redistribution of energy available for ocean mixing by long-range propagation of internal waves. *Nature*, *423*, 159-162. doi: https://doi.org/10.1038/nature01628

Althaus, A. M., Kunze, E., & Sanford, T. B. (2003). Internal Tide Radiation from Mendocino Escarpment. *Journal of Physical Oceanography*, *33*(7), 1510-1527. doi: https://doi.org/10.1175/1520-0485(2003)033⟨1510:ITRFME⟩2.0.CO;2

Baines, P. (1982). On internal tide generation models. *Deep Sea Research Part A. Oceanographic Research Papers*, *29*(3), 307-338. doi: https://doi.org/10.1016/0198-0149(82)90098-X

Barrick, D. E., & Lipa, B. J. (1997). Evolution of Bearing Determination in HF Current Mapping Radars. *Oceanography*, *10*(2), 72–75.

Bendat, J., & Piersol, A. (2010). Random Data. In (p. 417-472). John Wiley & Sons, Ltd. (ISBN:9781118032428) doi: https://doi.org/10.1002/9781118032428

Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., . . . Wyld, H. W. (1985). Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, *50*(6), 903-923. doi: https://doi.org/10.1190/1.1441970

Bressan, A., & Constantin, A. (2019). The Deflection Angle of Surface Ocean Currents From the Wind Direction. *Journal of Geophysical Research: Oceans*, *124*(11), 7412-7420. doi: https://doi.org/10.1029/2019JC015454

Buijsman, M. C., Uchiyama, Y., McWilliams, J. C., & Hill-Lindsay, C. R. (2012). Modeling Semidiurnal Internal Tide Variability in the Southern California Bight. *Journal of Physical Oceanography*, *42*(1), 62 - 77. doi: https://doi.org/10.1175/2011JPO4597.1

Burla, M., Baptista, A. M., Zhang, Y., & Frolov, S. (2010). Seasonal and interannual variability of the Columbia River plume: A perspective enabled by multiyear simulation databases. *Journal of Geophysical Research: Oceans*, *115*(C2). doi: https://doi.org/10.1029/2008JC004964

Centurioni, L. R., Ohlmann, J. C., & Niiler, P. P. (2008). Permanent Meanders in the California Current System. *Journal of Physical Oceanography*, *38*(8), 1690 - 1710. doi: https://doi.org/10.1175/2008JPO3746.1

Chavanne, C. P., & Klein, P. (2010). Can oceanic submesoscale processes be observed with satellite altimetry? *Geophysical Research Letters*, *37*(22). doi: https://doi.org/10.1029/2010GL045057

Cherniawsky, J. Y., Foreman, M. G., Kuh Kang, S., Scharroo, R., & Eert, A. J. (2010). 18.6-year lunar nodal tides from altimeter data. *Continental Shelf Research*, *30*(6), 575-587. (Tides in Marginal Seas - A special issue in memory of Prof Alexei Nekrasov) doi: https://doi.org/10.1016/j.csr.2009.10.002

Chiswell, S. M. (2002, 09). Energy Levels, Phase, and Amplitude Modulation of the Baroclinic Tide off Hawaii. *Journal of Physical Oceanography*, *32*(9), 2640-2651. doi: https://doi.org/10.1175/1520-0485(2002)032⟨2640:ELPAAM⟩2.0.CO;2

Codiga, D. L. (2011, 1). *Unified Tidal Analysis and Prediction Using the UTide Matlab Functions* (Tech. Rep.). Graduate School of Oceanography, University of Rhode Island, Narragansett, RI. Retrieved from ftp://www.po.gso.uri.edu/pub/downloads/codiga/pubs/2011Codiga-UTide-Report.pdf (59pp.)

Coles, W., Hobbs, G., Champion, D. J., Manchester, R. N., & Verbiest, J. P. W. (2011, 11). Pulsar timing analysis in the presence of correlated noise. *Monthly Notices of the Royal Astronomical Society*, *418*(1), 561-570. doi: https://doi.org/10.1111/j.1365-2966.2011.19505.x

Colosi, J. A., & Munk, W. (2006). Tales of the Venerable Honolulu Tide Gauge. *Journal of Physical Oceanography*, *36*(6), 967 - 996. doi: https://doi.org/10.1175/JPO2876.1

Cosoli, S., Pattiaratchi, C., & Hetzel, Y. (2020, Feb). High-Frequency Radar Observations of Surface Circulation Features along the South-Western Australian Coast. *Journal of Marine Science and Engineering*, *8*(2). doi: https://doi.org/10.3390/jmse8020097

Dai, A., & Deser, C. (1999). Diurnal and semidiurnal variations in global surface wind and divergence fields. *Journal of Geophysical Research: Atmospheres*, *104*(D24), 31109-31125. doi: https://doi.org/10.1029/1999JD900927

Deser, C., & Smith, C. A. (1998). Diurnal and Semidiurnal Variations of the Surface Wind Field over the Tropical Pacific Ocean. *Journal of Climate*, *11*(7), 1730 - 1748. doi: https://doi.org/10.1175/1520-0442(1998)011⟨1730:DASVOT⟩2.0.CO;2

Dickey, T. D., & Bidigare, R. R. (2005, Jun.). Interdisciplinary oceanographic observations: the wave of the future. *Scientia Marina*, *69*(S1), 23–42. doi: https://doi.org/10.3989/scimar.2005.69s123

Doodson, A. T., & Lamb, H. (1921). The harmonic development of the tide-generating potential. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *100*(704), 305-329. doi: https://doi.org/10.1098/rspa.1921.0088

Doodson, A. T., & Lamb, H. (1924). Perturbations of harmonic tidal constants. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *106*(739), 513-526. doi: https://doi.org/10.1098/rspa.1924.0085

Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis* (3rd ed.). John Wiley & Sons, Ltd. doi: https://doi.org/10.1002/9781118625590

Dunphy, M., Ponte, A., Klein, P., & Le Gentil, S. (2017, 01). Low-Mode Internal Tide Propagation in a Turbulent Eddy Field. *Journal of Physical Oceanography*, *47*. doi: https://doi.org/10.1175/JPO-D-16-0099.1

Dushaw, B. D., Howe, B. M., Cornuelle, B. D., Worcester, P. F., & Luther, D. S. (1995). Barotropic and Baroclinic Tides in the Central North Pacific Ocean Determined from Long-Range Reciprocal Acoustic Transmissions. *Journal of Physical Oceanography*, *25*(4), 631 - 647. doi: https://doi.org/10.1175/1520-0485(1995)025⟨0631:BABTIT⟩2.0.CO;2

Egbert, G. D., & Erofeeva, S. Y. (2021). An Approach to Empirical Mapping of Incoherent Internal Tides With Altimetry Data. *Geophysical Research Letters*, *48*(24), e2021GL095863. doi: https://doi.org/10.1029/2021GL095863

Egbert, G. D., & Ray, R. D. (2000). Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data. *Nature*, *405*, 775-778. doi: https://doi.org/10.1038/35015531

Egbert, G. D., & Ray, R. D. (2001). Estimates of M2 tidal energy dissipation from TOPEX/Poseidon altimeter data. *Journal of Geophysical Research: Oceans*, *106*(C10), 22475-22502. doi: https://doi.org/10.1029/2000JC000699

Eich, M. L., Merrifield, M. A., & Alford, M. H. (2004). Structure and variability of semidiurnal internal tides in Mamala Bay, Hawaii. *Journal of Geophysical Research: Oceans*, *109*(C5). doi: https://doi.org/10.1029/2003JC002049

Ekman, V. W. (1905). On the influence of the earth's rotation on ocean-currents. *Ark. Mat. Astr. Fys.*, *2*(11).

Elipot, S., & Gille, S. T. (2009). Ekman layers in the Southern Ocean: spectral models and observations, vertical viscosity and boundary layer depth. *Ocean Science*, *5*(2), 115–139. doi: https://doi.org/10.5194/os-5-115-2009

Erofeeva, S. Y., Egbert, G. D., & Kosro, P. M. (2003). Tidal currents on the central Oregon shelf: Models, data, and assimilation. *Journal of Geophysical Research: Oceans*, *108*(C5).

doi: https://doi.org/10.1029/2002JC001615

Ferrari, R., & Wunsch, C. (2009). Ocean Circulation Kinetic Energy: Reservoirs, Sources, and Sinks. *Annual Review of Fluid Mechanics*, *41*(1), 253-282. doi: https://doi.org/10.1146/annurev.fluid.40.111406.102139

Fleming, L. E., Kirkpatrick, B., Backer, L. C., Walsh, C. J., Nierenberg, K., Clark, J., . . . Baden, D. G. (2011). Review of Florida red tide and human health effects. *Harmful Algae*, *10*(2), 224-233. doi: https://doi.org/10.1016/j.hal.2010.08.006

Foreman, M. G. G. (1977). Manual for tidal heights analysis and prediction. Pacific Marine Science Rep. 77-10 [Computer software manual]. Patricia Bay, BC, Canada.

Foreman, M. G. G., Cherniawsky, J. Y., & Ballantyne, V. A. (2009). Versatile Harmonic Tidal Analysis: Improvements and Applications. *Journal of Atmospheric and Oceanic Technology*, *26*(4), 806-817. doi: https://doi.org/10.1175/2008JTECHO615.1

Foreman, M. G. G., & Henry, R. F. (1989, September). The harmonic analysis of tidal model time series. *Advances in Water Resources*, *12*, 109-120. doi: https://doi.org/10.1016/0309-1708(89)90017-1

Fredj, E., Roarty, H., Kohut, J., Smith, M., & Glenn, S. (2016). Gap Filling of the Coastal Ocean Surface Currents from HFR Data: Application to the Mid-Atlantic Bight HFR Network. *Journal of Atmospheric and Oceanic Technology*, *33*(6), 1097-1111. doi: https://doi.org/10.1175/JTECH-D-15-0056.1

Garrett, C., & Kunze, E. (2007). Internal Tide Generation in the Deep Ocean. *Annual Review of Fluid Mechanics*, *39*(1), 57-87. doi: https://doi.org/10.1146/annurev.fluid.39.050905.110227

Giglio, D., Gille, S. T., Cornuelle, B. D., Subramanian, A. C., Turk, F. J., Hristova-Veleva, S., & Northcott, D. (2022). Annual Modulation of Diurnal Winds in the Tropical Oceans. *Remote Sensing*, *14*(3). doi: https://doi.org/10.3390/rs14030459

Gille, S. T., Llewellyn Smith, S. G., & Lee, S. M. (2003). Measuring the sea breeze from QuikSCAT Scatterometry. *Geophysical Research Letters*, *30*(3). doi: https://doi.org/10.1029/2002GL016230

Gille, S. T., Llewellyn Smith, S. G., & Statom, N. M. (2005). Global observations of the land breeze. *Geophysical Research Letters*, *32*(5). doi: https://doi.org/10.1029/2004GL022139

Gonella, J. (1972). A rotary-component method for analysing meteorological and oceanographic vector time series. *Deep Sea Research and Oceanographic Abstracts*, *19*(12), 833 - 846. doi: https://doi.org/10.1016/0011-7471(72)90002-2

Goodman, N. R. (1957). *On the Joint Estimation of the Spectra, Cospectrum and Quadrature Spectrum of a Two-dimensional Stationary Gaussian Process* (Unpublished doctoral dissertation). Princeton University, Princeton, NJ.

Gurgel, K.-W. (1994). Shipborne measurement of surface current fields by HF radar. In *Proceedings of oceans'94* (Vol. 3, p. III/23-III/27 vol.3). doi: https://doi.org/10.1109/OCEANS.1994.364167

Haigh, I. D., Eliot, M., & Pattiaratchi, C. (2011). Global influences of the 18.61 year nodal cycle and 8.85 year cycle of lunar perigee on high tidal levels. *Journal of Geophysical Research: Oceans*, *116*(C6). doi: https://doi.org/10.1029/2010JC006645

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., . . . Thépaut, J.-N. (2018). ERA5 hourly data on single levels from 1959 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. ((Accessed on 12-Apr-2022)) doi: https://doi.org/10.24381/cds.adbb2d47

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., . . . Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. doi: https://doi.org/10.1002/qj.3803

Hewitt, E., & Hewitt, R. E. (1979, June 01). The Gibbs-Wilbraham phenomenon: An episode in fourier analysis. *Archive for History of Exact Sciences*, *21*(2), 129–160. doi: https://doi.org/10.1007/BF00330404

Hickey, B., Geier, S., Kachel, N., & MacFadyen, A. (2005). A bi-directional river plume: The Columbia in summer. *Continental Shelf Research*, *25*(14), 1631-1656. doi: https://doi.org/10.1016/j.csr.2005.04.010

Hickey, B. M., Pietrafesa, L. J., Jay, D. A., & Boicourt, W. C. (1998). The Columbia River Plume Study: Subtidal variability in the velocity and salinity fields. *Journal of Geophysical Research: Oceans*, *103*(C5), 10339-10368. doi: https://doi.org/10.1029/97JC03290

Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, *12*(1), 55-67. doi: https://doi.org/10.1080/00401706.1970.10488634

Huang, X., Wang, Z., Zhang, Z., Yang, Y., Zhou, C., Yang, Q., . . . Tian, J. (2018). Role of Mesoscale Eddies in Modulating the Semidiurnal Internal Tide: Observation Results in the Northern South China Sea. *Journal of Physical Oceanography*, *48*(8), 1749-1770. doi: https://doi.org/10.1175/JPO-D-17-0209.1

Ide, K., Courtier, P., Ghil, M., & Lorenc, A. C. (1997). Unified Notation for Data Assimilation : Operational, Sequential and Variational. *Journal of the Meteorological Society of Japan. Ser.*

II, *75*(1B), 181-189. doi: https://doi.org/10.2151/jmsj1965.75.1B_181

Kachelein, L., Cornuelle, B. D., Gille, S. T., & Mazloff, M. R. (2022). Harmonic Analysis of Non-Phase-Locked Tides with Red Noise Using the red_tide Package. *Journal of Atmospheric and Oceanic Technology*, *39*(7), 1031 - 1051. doi: https://doi.org/10.1175/JTECH-D-21-0034.1

Kim, S. Y., Cornuelle, B. D., & Terrill, E. J. (2009). Anisotropic Response of Surface Currents to the Wind in a Coastal Region. *Journal of Physical Oceanography*, *39*(6), 1512 - 1533. doi: https://doi.org/10.1175/2009JPO4013.1

Kim, S. Y., Cornuelle, B. D., & Terrill, E. J. (2010). Decomposing observations of high-frequency radar-derived surface currents by their forcing mechanisms: Decomposition techniques and spatial structures of decomposed surface currents. *Journal of Geophysical Research: Oceans*, *115*(C12). doi: https://doi.org/10.1029/2010JC006222

Kim, S. Y., & Kosro, P. M. (2013). Observations of near-inertial surface currents off Oregon: Decorrelation time and length scales. *Journal of Geophysical Research: Oceans*, *118*(7), 3723-3736. doi: https://doi.org/10.1002/jgrc.20235

Kim, S. Y., Terrill, E. J., & Cornuelle, B. D. (2008). Mapping surface currents from HF radar radial velocity measurements using optimal interpolation. *Journal of Geophysical Research: Oceans*, *113*(C10). doi: https://doi.org/10.1029/2007JC004244

Kim, S. Y., Terrill, E. J., Cornuelle, B. D., Jones, B., Washburn, L., Moline, M. A., . . . Kosro, P. M. (2011). Mapping the U.S. West Coast surface circulation: A multiyear analysis of high-frequency radar observations. *Journal of Geophysical Research: Oceans*, *116*(C3). doi: https://doi.org/10.1029/2010JC006669

Kohut, J. T., Glenn, S. M., & Paduan, J. D. (2006). Inner shelf response to Tropical Storm Floyd. *Journal of Geophysical Research: Oceans*, *111*(C9). doi: https://doi.org/10.1029/2003JC002173

Kosro, P. M. (2005). On the spatial structure of coastal circulation off Newport, Oregon, during spring and summer 2001 in a region of varying shelf width. *Journal of Geophysical Research: Oceans*, *110*(C10). doi: https://doi.org/10.1029/2004JC002769

Krauss, W. (1993). Ekman drift in homogeneous water. *Journal of Geophysical Research: Oceans*, *98*(C11), 20187-20209. doi: https://doi.org/10.1029/93JC01898

Kurapov, A. L., Egbert, G. D., Allen, J. S., Miller, R. N., Erofeeva, S. Y., & Kosro, P. M. (2003). The M2 Internal Tide off Oregon: Inferences from Data Assimilation. *Journal of Physical Oceanography*, *33*(8), 1733 - 1757. doi: https://doi.org/10.1175/1520-0485(2003)033⟨1733:TMITOO⟩2.0.CO;2

Lana, A., Marmain, J., Fernández, V., Tintoré, J., & Orfila, A. (2016, Apr 01). Wind influence on surface current variability in the Ibiza Channel from HF Radar. *Ocean Dynamics*, *66*(4), 483–497. doi: https://doi.org/10.1007/s10236-016-0929-z

Lee, P. M. (1997). *Bayesian Statistics: an Introduction* (2nd ed ed.). London: Arnold. (ISBN: 0-471-19481-6)

Lee, P. M. (2012). *Bayesian Statistics: An Introduction* (4th ed.). Wiley. (ISBN: 978-1-1183-3257-3)

Leffler, K. E., & Jay, D. A. (2009). Enhancing tidal harmonic analysis: Robust (hybrid L1/L2) solutions. *Continental Shelf Research*, *29*(1), 78 - 88. (Physics of Estuaries and Coastal Seas: Papers from the PECS 2006 Conference) doi: https://doi.org/10.1016/j.csr.2008.04.011

Lenn, Y.-D., & Chereskin, T. K. (2009). Observations of Ekman Currents in the Southern Ocean. *Journal of Physical Oceanography*, *39*(3), 768 - 779. doi: https://doi.org/10.1175/2008JPO3943.1

Le Provost, C. (2001). Chapter 6 Ocean Tides. In L.-L. Fu & A. Cazenave (Eds.), *Satellite Altimetry and Earth Sciences* (Vol. 69, p. 267 - 303). Academic Press. doi: https://doi.org/10.1016/S0074-6142(01)80151-0

Lerczak, J. A., Hendershott, M. C., & Winant, C. D. (2001). Observations and modeling of coastal internal waves driven by a diurnal sea breeze. *Journal of Geophysical Research: Oceans*, *106*(C9), 19715-19729. doi: https://doi.org/10.1029/2001JC000811

Lipa, B., & Barrick, D. (1983). Least-squares methods for the extraction of surface currents from CODAR crossed-loop data: Application at ARSLOE. *IEEE Journal of Oceanic Engineering*, *8*(4), 226-253. doi: https://doi.org/10.1109/JOE.1983.1145578

Lourens, A., & van Geer, F. C. (2016). Uncertainty propagation of arbitrary probability density functions applied to upscaling of transmissivities. *Stochastic Environmental Research and Risk Assessment*, *30*(1), 237-249. doi: https://doi.org/10.1007/s00477-015-1075-8

Lyard, F. (2022, April). personal communication.

Lyard, F. H., Allain, D. J., Cancet, M., Carrère, L., & Picot, N. (2021). FES2014 global ocean tide atlas: design and performance. *Ocean Science*, *17*(3), 615–649. doi: https://doi.org/10.5194/os-17-615-2021

Masich, J., Kessler, W. S., Cronin, M. F., & Grissom, K. R. (2021). Diurnal cycles of near-surface currents across the tropical pacific. *Journal of Geophysical Research: Oceans*, *126*(4), e2020JC016982. (e2020JC016982 2020JC016982) doi: https://doi.org/10.1029/2020JC016982

Matte, P., Jay, D. A., & Zaron, E. D. (2013). Adaptation of Classical Tidal Harmonic Analysis to Nonstationary Tides, with Application to River Tides. *Journal of Atmospheric and Oceanic Technology*, *30*(3), 569-589. doi: https://doi.org/10.1175/JTECH-D-12-00016.1

McWilliams, J. C., & Restrepo, J. M. (1999). The Wave-Driven Ocean Circulation. *Journal of Physical Oceanography*, *29*(10), 2523 - 2540. doi: https://doi.org/10.1175/1520-0485(1999) 029⟨2523:TWDOC⟩2.0.CO;2

Menke, W. (2018). *Geophysical Data Analysis: Discrete Inverse Theory* (Fourth edition ed.). London, United Kingdom: Elsevier Ltd.

Metref, S., Cosme, E., Le Guillou, F., Le Sommer, J., Brankart, J.-M., & Verron, J. (2020). Wide-Swath Altimetric Satellite Data Assimilation With Correlated-Error Reduction. *Frontiers in Marine Science*, *6*. doi: https://doi.org/10.3389/fmars.2019.00822

Mitchum, G. T., & Chiswell, S. M. (2000). Coherence of internal tide modulations along the Hawaiian Ridge. *Journal of Geophysical Research: Oceans*, *105*(C12), 28653-28661. doi: https://doi.org/10.1029/2000JC900140

Mooers, C. N. (1973). A technique for the cross spectrum analysis of pairs of complex-valued time series, with emphasis on properties of polarized components and rotational invariants. *Deep Sea Research and Oceanographic Abstracts*, *20*(12), 1129-1141. doi: https://doi.org/10.1016/0011-7471(73)90027-2

Morrow, R., Fu, L.-L., Ardhuin, F., Benkiran, M., Chapron, B., Cosme, E., . . . Zaron, E. D. (2019). Global Observations of Fine-Scale Ocean Surface Topography With the Surface Water and Ocean Topography (SWOT) Mission. *Frontiers in Marine Science*, *6*. doi: https://doi.org/10.3389/fmars.2019.00232

Müller, M., Cherniawsky, J. Y., Foreman, M. G. G., & von Storch, J.-S. (2014, Feb 01). Seasonal variation of the M2 tide. *Ocean Dynamics*, *64*(2), 159-177. doi: https://doi.org/10.1007/ s10236-013-0679-0

Munk, W., & Hasselman, K. (1964). Super-resolution of tides. *Studies on Oceanograhy*, 339-334. ((Hidaka volume))

Munk, W. H., & Wunsch, C. (1998). Abyssal recipes II: energetics of tidal and wind mixing. *Deep Sea Research Part I: Oceanographic Research Papers*, *45*(12), 1977-2010. doi: https://doi.org/10.1016/S0967-0637(98)00070-3

Munk, W. H., Zetler, B., & Groves, G. W. (1965). Tidal Cusps. *Geophysical Journal of the Royal Astronomical Society*, *10*(2), 211-219. doi: https://doi.org/10.1111/j.1365-246X.1965 .tb03062.x

Nash, J. D., & Moum, J. N. (2005). River plumes as a source of large-amplitude internal waves in the coastal ocean. *Nature*, *437*(7057), 400–403. doi: https://doi.org/10.1038/nature03936

National Oceanic and Atmospheric Administration. (2005). *Deep-Ocean Assessment and Reporting of Tsunamis (DART(R))*. NOAA National Centers for Environmental Information. (Accessed: 2021-03-21) doi: https://doi.org/10.7289/V5F18WNS

Nelson, A. D., Arbic, B. K., Zaron, E. D., Savage, A. C., Richman, J. G., Buijsman, M. C., & Shriver, J. F. (2019). Toward Realistic Nonstationarity of Semidiurnal Baroclinic Tides in a Hydrodynamic Model. *Journal of Geophysical Research: Oceans*, *124*(9), 6632-6642. doi: https://doi.org/10.1029/2018JC014737

NIST/SEMATECH e-Handbook of Statistical Methods [Computer software manual]. (2012). Gaithersburg, MD. doi: https://doi.org/10.18434/M32189

Ohlmann, C., White, P., Washburn, L., Emery, B., Terrill, E., & Otero, M. (2007). Interpretation of Coastal HF Radar–Derived Surface Currents with High-Resolution Drifter Data. *Journal of Atmospheric and Oceanic Technology*, *24*(4), 666-680. doi: https://doi.org/10.1175/JTECH1998.1

Onink, V., Wichmann, D., Delandmeter, P., & van Sebille, E. (2019). The Role of Ekman Currents, Geostrophy, and Stokes Drift in the Accumulation of Floating Microplastic. *Journal of Geophysical Research: Oceans*, *124*(3), 1474-1490. doi: https://doi.org/10.1029/2018JC014547

Paduan, J. D., & Rosenfeld, L. K. (1996). Remotely sensed surface currents in Monterey Bay from shore-based HF radar (Coastal Ocean Dynamics Application Radar). *Journal of Geophysical Research: Oceans*, *101*(C9), 20669-20686. doi: https://doi.org/10.1029/96JC01663

Parke, M. E., Stewart, R. H., Farless, D. L., & Cartwright, D. E. (1987). On the choice of orbits for an altimetric satellite to study ocean circulation and tides. *Journal of Geophysical Research: Oceans*, *92*(C11), 11693-11707. doi: https://doi.org/10.1029/JC092iC11p11693

Parker, B. B. (2007). *Tidal Analysis and Prediction*. NOAA NOS Center for Operational Oceanographic Products and Services, Silver Spring, MD. (NOAA Special Publication NOS CO-OPS 3) doi: http://dx.doi.org/10.25607/OBP-191

Pawlowicz, R., Beardsley, R., & Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Computers & Geosciences*, *28*(8), 929-937. doi: https://doi.org/10.1016/S0098-3004(02)00013-4

Pidgeon, E. J., & Winant, C. D. (2005). Diurnal variability in currents and temperature on the continental shelf between central and southern California. *Journal of Geophysical Research: Oceans*, *110*(C3). doi: https://doi.org/10.1029/2004JC002321

Polton, J. A., Lewis, D. M., & Belcher, S. E. (2005). The Role of Wave-Induced Coriolis–Stokes Forcing on the Wind-Driven Mixed Layer. *Journal of Physical Oceanography*, *35*(4), 444 - 457. doi: https://doi.org/10.1175/JPO2701.1

Ponte, A., & Klein, P. (2015, 02). Incoherent signature of internal tides on sea level in idealized numerical simulations. *Geophysical Research Letters*, *42*. doi: https://doi.org/10.1002/2014GL062583

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1988). *Numerical Recipes in C*. New York: Cambridge University Press.

Price, J. F., & Sundermeyer, M. A. (1999). Stratified Ekman layers. *Journal of Geophysical Research: Oceans*, *104*(C9), 20467-20494. doi: https://doi.org/10.1029/1999JC900164

Price, J. F., Weller, R. A., & Pinkel, R. (1986). Diurnal cycling: Observations and models of the upper ocean response to diurnal heating, cooling, and wind mixing. *Journal of Geophysical Research: Oceans*, *91*(C7), 8411-8427. doi: https://doi.org/10.1029/JC091iC07p08411

Radok, R., Munk, W., & Isaacs, J. (1967). A note on mid-ocean internal tides. *Deep Sea Research and Oceanographic Abstracts*, *14*(1), 121-124. doi: https://doi.org/10.1016/0011-7471(67)90035-6

Rainville, L., & Pinkel, R. (2006). Propagation of Low-Mode Internal Waves through the Ocean. *Journal of Physical Oceanography*, *36*(6), 1220 - 1236. doi: https://doi.org/10.1175/JPO2889.1

Ray, R. (1993). Global ocean tide models on the eve of TOPEX/Poseidon. *IEEE Transactions on Geoscience and Remote Sensing*, *31*(2), 355-364. doi: https://doi.org/10.1109/36.214911

Ray, R. D., & Cartwright, D. E. (2001). Estimates of internal tide energy fluxes from Topex/Poseidon Altimetry: Central North Pacific. *Geophysical Research Letters*, *28*(7), 1259-1262. doi: https://doi.org/10.1029/2000GL012447

Ray, R. D., & Egbert, G. D. (2004). The Global S_1 Tide. *Journal of Physical Oceanography*, *34*(8), 1922 - 1935. doi: https://doi.org/10.1175/1520-0485(2004)034⟨1922:TGST⟩2.0.CO;2

Ray, R. D., & Mitchum, G. T. (1996). Surface manifestation of internal tides generated near Hawaii. *Geophysical Research Letters*, *23*(16), 2101-2104. doi: https://doi.org/10.1029/96GL02050

Ray, R. D., & Zaron, E. D. (2011). Non-stationary internal tides observed with satellite altimetry. *Geophysical Research Letters*, *38*(17). doi: https://doi.org/10.1029/2011GL048617

Ray, R. D., & Zaron, E. D. (2016). M_2 Internal Tides and Their Observed Wavenumber Spectra from Satellite Altimetry. *Journal of Physical Oceanography*, *46*(1), 3-22. doi:

https://doi.org/10.1175/JPO-D-15-0065.1

Rio, M.-H., & Hernandez, F. (2003). High-frequency response of wind-driven currents measured by drifting buoys and altimetry over the world ocean. *Journal of Geophysical Research: Oceans*, *108*(C8). doi: https://doi.org/10.1029/2002JC001655

Roarty, H., Cook, T., Hazard, L., George, D., Harlan, J., Cosoli, S., . . . Grilli, S. (2019). The Global High Frequency Radar Network. *Frontiers in Marine Science*, *6*, 164. doi: https://doi.org/10.3389/fmars.2019.00164

Rodríguez, E., Bourassa, M., Chelton, D., Farrar, J. T., Long, D., Perkovic-Martin, D., & Samelson, R. (2019). The Winds and Currents Mission Concept. *Frontiers in Marine Science*, *6*. doi: https://doi.org/10.3389/fmars.2019.00438

Rotunno, R. (1983). On the Linear Theory of the Land and Sea Breeze. *Journal of Atmospheric Sciences*, *40*(8), 1999 - 2009. doi: https://doi.org/10.1175/1520-0469(1983)040⟨1999: OTLTOT⟩2.0.CO;2

Round, R. D. (1993). *Climatology and analysis of the Monterey Bay sea breeze* (Unpublished master's thesis). Naval Postgraduate School.

Savage, A. C., Arbic, B. K., Alford, M. H., Ansong, J. K., Farrar, J. T., Menemenlis, D., . . . Zamudio, L. (2017). Spectral decomposition of internal gravity wave sea surface height in global models. *Journal of Geophysical Research: Oceans*, *122*(10), 7803-7821. doi: https://doi.org/10.1002/2017JC013009

Savage, A. C., Arbic, B. K., Richman, J. G., Shriver, J. F., Alford, M. H., Buijsman, M. C., . . . Zamudio, L. (2017). Frequency content of sea surface height variability from internal gravity waves to mesoscale eddies. *Journal of Geophysical Research: Oceans*, *122*(3), 2519-2538. doi: https://doi.org/10.1002/2016JC012331

Savage, A. C., Waterhouse, A. F., & Kelly, S. M. (2020). Internal Tide Nonstationarity and Wave–Mesoscale Interactions in the Tasman Sea. *Journal of Physical Oceanography*, *50*(10), 2931 - 2951. doi: https://doi.org/10.1175/JPO-D-19-0283.1

Schmidt, R. (1986). Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, *34*(3), 276-280. doi: https://doi.org/10.1109/TAP.1986 .1143830

Scott, R. B., & Xu, Y. (2009). An update on the wind power input to the surface geostrophic flow of the World Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, *56*(3), 295-304. doi: https://doi.org/10.1016/j.dsr.2008.09.010

Shriver, J. F., Richman, J. G., & Arbic, B. K. (2014). How stationary are the internal tides in a

high-resolution global ocean circulation model? *Journal of Geophysical Research: Oceans*, *119*(5), 2769-2787. doi: https://doi.org/10.1002/2013JC009423

Terrill, E., Otero, M., Hazard, L., Conlee, D., Harlan, J., Kohut, J., . . . Lindquist, K. (2006). Data Management and Real-time Distribution in the HF-Radar National Network. In *Oceans 2006* (p. 1-6). doi: https://doi.org/10.1109/OCEANS.2006.306883

Thomas, A. C., & Weatherbee, R. A. (2006). Satellite-measured temporal variability of the Columbia River plume. *Remote Sensing of Environment*, *100*(2), 167-178. doi: https://doi.org/10.1016/j.rse.2005.10.018

Todd, R. E., Rudnick, D. L., Mazloff, M. R., Davis, R. E., & Cornuelle, B. D. (2011). Poleward flows in the southern California Current System: Glider observations and numerical simulation. *Journal of Geophysical Research: Oceans*, *116*(C2). doi: https://doi.org/10.1029/2010JC006536

Trenberth, K. E., Fasullo, J. T., & Kiehl, J. (2009). Earth's Global Energy Budget. *Bulletin of the American Meteorological Society*, *90*(3), 311 - 324. doi: https://doi.org/10.1175/2008BAMS2634.1

U.S. Integrated Ocean Observing System. (2016). Manual for Real-Time Quality Control of High Frequency Radar Surface Current Data: a Guide to Quality Control and Quality Assurance for High Frequency Radar Surface Current Observations. Version 1.0. [Computer software manual]. (58pp.) doi: https://doi.org/10.25607/OBP-1464

Van Trees, H. L. (2001). Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory. In (p. 54-63). John Wiley & Sons, Inc. (ISBNs: 0-471-09517-6 (Paperback); 0-471-22108-2 (Electronic))

Villas Bôas, A. B., Ardhuin, F., Ayet, A., Bourassa, M. A., Brandt, P., Chapron, B., . . . van Sebille, E. (2019). Integrated observations of global surface winds, currents, and waves: Requirements and challenges for the next decade. *Frontiers in Marine Science*, *6*. doi: https://doi.org/10.3389/fmars.2019.00425

von Storch, H., & Zwiers, F. W. (2003). *Statistical Analysis in Climate Research*. Cambridge University Press. (ISBN: 0-521-45071-3)

Walsh, J. E. (1974). Sea Breeze Theory and Applications. *Journal of Atmospheric Sciences*, *31*(8), 2012 - 2026. doi: https://doi.org/10.1175/1520-0469(1974)031⟨2012:SBTAA⟩2.0.CO;2

Wang, J., Fu, L.-L., Qiu, B., Menemenlis, D., Farrar, J. T., Chao, Y., . . . Flexas, M. M. (2018). An Observing System Simulation Experiment for the Calibration and Validation of the Surface Water Ocean Topography Sea Surface Height Measurement Using In Situ Platforms. *Journal of Atmospheric and Oceanic Technology*, *35*(2), 281-297. doi: https://doi.org/10.1175/

Wang, W., & Huang, R. X. (2004). Wind Energy Input to the Ekman Layer. *Journal of Physical Oceanography*, *34*(5), 1267 - 1275. doi: https://doi.org/10.1175/1520-0485(2004)034⟨1267: WEITTE⟩2.0.CO;2

Watanabe, M., & Hibiya, T. (2002). Global estimates of the wind-induced energy flux to inertial motions in the surface mixed layer. *Geophysical Research Letters*, *29*(8), 64-1-64-3. doi: https://doi.org/10.1029/2001GL014422

Weber, J. E. (1983). Steady Wind- and Wave-Induced Currents in the Open Ocean. *Journal of Physical Oceanography*, *13*(3), 524 - 530. doi: https://doi.org/10.1175/1520-0485(1983) 013⟨0524:SWAWIC⟩2.0.CO;2

Wunsch, C. (1996). *The Ocean Circulation Inverse Problem*. Cambridge University Press. (ISBN: 0-521-48090-6)

Wunsch, C. (1998). The Work Done by the Wind on the Oceanic General Circulation. *Journal of Physical Oceanography*, *28*(11), 2332 - 2340. doi: https://doi.org/10.1175/ 1520-0485(1998)028⟨2332:TWDBTW⟩2.0.CO;2

Yu, Z., Fan, Y., Metzger, E. J., & Smedstad, O. M. (2018). The wind work input into the global ocean revealed by a 17-year global HYbrid coordinate ocean model reanalysis. *Ocean Modelling*, *130*, 29-39. doi: https://doi.org/10.1016/j.ocemod.2018.07.009

Zaron, E. D. (2015). Nonstationary Internal Tides Observed Using Dual-Satellite Altimetry. *Journal of Physical Oceanography*, *45*(9), 2239 - 2246. doi: https://doi.org/10.1175/JPO-D -15-0020.1

Zaron, E. D. (2017). Mapping the nonstationary internal tide with satellite altimetry. *Journal of Geophysical Research: Oceans*, *122*(1), 539-554. doi: https://doi.org/10.1002/2016JC012487

Zaron, E. D. (2019). Predictability of non-phase-locked baroclinic tides in the Caribbean Sea. *Ocean Science*, *15*(5), 1287–1305. doi: https://doi.org/10.5194/os-15-1287-2019

Zaron, E. D. (2022). Baroclinic Tidal Cusps from Satellite Altimetry. *Journal of Physical Oceanography*, *52*(12), 3123 - 3137. doi: https://doi.org/10.1175/JPO-D-21-0155.1

Zaron, E. D., & Ray, R. D. (2018). Aliased Tidal Variability in Mesoscale Sea Level Anomaly Maps. *Journal of Atmospheric and Oceanic Technology*, *35*(12), 2421 - 2435. doi: https://doi.org/10.1175/JTECH-D-18-0089.1

Zetler, B. D. (1982). Computer Applications to Tides in the National Ocean Survey: Supplement to Manual of Harmonic Analysis and Prediction of Tides (Special Publication No. 98) [Computer

software manual]. La Jolla, California 92093.

Zetler, B. D., Schuldt, M. D., Whipple, R. W., & Hicks, S. D. (1965). Harmonic Analysis of Tides
from Data Randomly Spaced in Time. *Journal of Geophysical Research*, *70*(12), 2805-2811.

Zhao, Z., Alford, M. H., Girton, J., Johnston, T. M. S., & Carter, G. (2011). Internal tides around
the Hawaiian Ridge estimated from multisatellite altimetry. *Journal of Geophysical Research:
Oceans*, *116*(C12). doi: https://doi.org/10.1029/2011JC007045