

## Introduction

- We have developed a tidal analysis routine specialized for time series with tidal constituents that are relatively weak compared to their low-frequency, non-tidal components. This is an improvement over classical tidal harmonic analysis, such as that used in the well known `t_tide` package [4]. This routine is executed by the software package `red_tide`, named for its handling of spectrally red non-tidal signals. Notably, for weak tides, `red_tide` converges to tidal amplitude estimates over shorter record lengths than `t_tide` does.

## Background

- Classical tidal harmonic analysis has a long history of use in the field of tidal prediction [3][7]. Classical tidal harmonic analysis has well-known drawbacks [4], such as nodal modulations and limitations on frequency resolution. Nevertheless it has been widely adopted due to the relatively deterministic nature of tides, for which this method is suitable. Adjustments to the technique have been implemented, notably by Pawlowicz et al. [4] in the `t_tide` package, which correct for limitations of short time series and interference by tidal-frequency signals of incoherent origin.
- Nevertheless, weak tidal signals and/or relatively strong non-tidal components are difficult to describe with conventional tidal harmonic analysis, e.g. baroclinic tides.

## Relevance to SWOT

- SWOT will detect small-scale signals driven by tides, surface waves, and internal waves. We aim to build the modeling and assimilation capabilities to resolve these processes.
- Analysis of coastal high-frequency radar (HFR) surface current data in the California Current System (CCS), a SWOT cal/val region, has led to the development of a flexible tidal analysis package `red_tide` [2], designed for non-stationary and/or weak tidal signals with a spectrally red background field.
- Approximately 25% of baroclinic tidal SSH variance is non-stationary [6], marked by time-varying modulations to amplitude or phase. Therefore, high-wavenumber observations by SWOT may contain substantial non-stationary tidal energy for which such analysis is advantageous.

## Methods

- Harmonic analysis of select tidal constituents (e.g.  $O_1$ ,  $M_2$ ,  $S_2$ , etc.) and non-tidal frequencies, including low, non-tidal frequencies  $< O(1)$  cpd.
- Model weights are the Bayesian maximum *a posteriori* estimate assuming linear and Gaussian statistics.

Model [5]:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{r}$$

$\mathbf{y}$  = observations

$\mathbf{H}$  = model basis functions

$\mathbf{x}$  = amplitudes

$\mathbf{r}$  = unmodeled residual signal

$$\mathbf{x} \approx \hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{P}^{-1})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{y}$$

$$\mathbf{R} = \langle \mathbf{r} \mathbf{r}^T \rangle$$

$$\mathbf{P} = \langle \mathbf{x} \mathbf{x}^T \rangle$$

- $\mathbf{R}$  is assumed constant (uncorrelated, stationary noise) due to computational constraints of inverting non-diagonal matrices.
- $\mathbf{P}$  is prior guess for model, diagonal for computational efficiency. Model is a Fourier decomposition, thus  $\mathbf{P}$  is described by the power spectrum of  $\mathbf{y}$ :  
$$\mathbf{P} = \mathbf{S}_{yy}(f_{\text{modeled}}) \Delta f$$
- Parseval's theorem dictates energy conservation between physical and Fourier domains, so energy at unmodeled frequencies can dictate  $\mathbf{R}$ :

$$\mathbf{R} = \Delta f \sum \mathbf{S}_{yy}(f_{\text{unmodeled}})$$

- $\mathbf{R}$  sets the signal-to-noise allowance for model weights, and may be chosen freely. It accounts automatically for representation error (unmodeled energy) and can also be adjusted to account for expected observation error.
- Model dispersion about truth is an estimate for uncertainty of  $\hat{\mathbf{x}}$ :

$$\langle (\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T \rangle = (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{P}^{-1})^{-1}$$

- Classical harmonic analysis, including `t_tide`, employs a similar least-squares method, which does not include the prior statistical assumptions ( $\mathbf{P}$  and  $\mathbf{R}$ ) used above:

$$\hat{\mathbf{x}}' = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y}$$

## Methods (continued)

- Artificial time series are constructed from the inverse Fourier transform of idealized quasi-tidal spectra, with a phase randomly assigned to each frequency (figure 1).

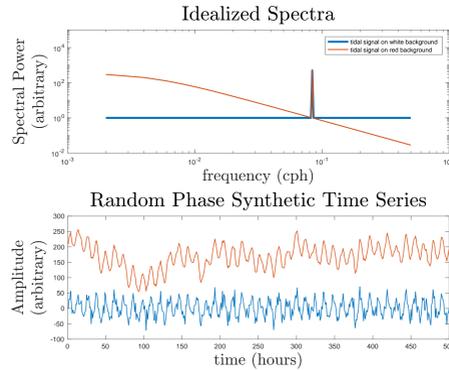


Figure 1. Simplified example of synthetic spectra and corresponding random phase time series used to evaluate `red_tide`.

## Results

- Analyzing artificial records at their prescribed frequencies reconstructs constituents, with minor inaccuracy due to the unmodeled allowance (figure 2). Uncertainty intervals do not depend on data, but rather on model assumptions and orthogonality of model basis functions.

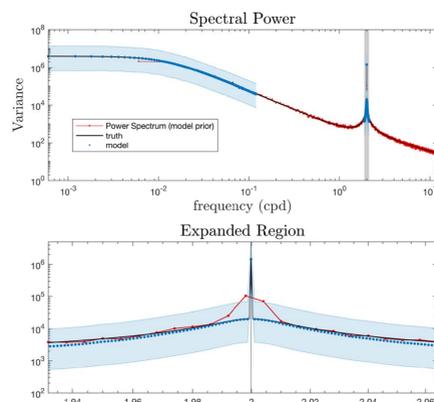


Figure 2. Synthetic spectrum (arbitrary units) and partial reconstruction using `red_tide`, where every frequency modeled in  $\mathbf{H}$  was also used to construct  $\mathbf{y}$ . The power spectral estimate of  $\mathbf{y}$  (here in red) was used to calculate the model prior covariance estimate  $\mathbf{P}$  and the error covariance  $\mathbf{R}$ .

- Analyzing a subset of artificial records, not necessarily at true constituent frequencies, results in ambiguity (figure 3), because the shorter record results in the prescribed frequencies not aligning exactly with those of the true parent spectrum. This mimics real data, which has energy at all frequencies but can only be modeled on a finite basis.

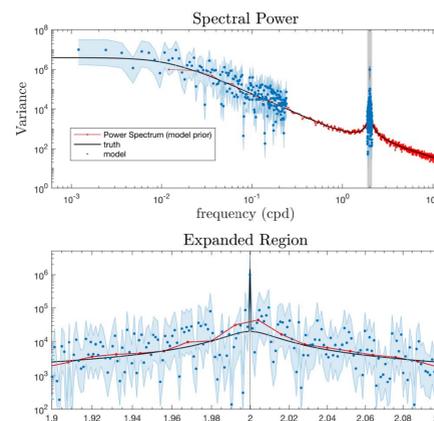


Figure 3. Analogous to figure 2, but with only half the record analyzed, thus not all modeled frequencies are true constituents of  $\mathbf{y}$ . The prior estimate of fractional residual variance is  $\mathbf{R}/\text{var}(\mathbf{y}) \approx 0.1$ , while the calculated fraction of residual variance of the model is  $\text{var}(\mathbf{y} - \mathbf{H}\mathbf{x})/\text{var}(\mathbf{y}) \approx 0.05$ . The estimates are distributed about the true spectrum roughly as expected, and the tidal peak is correctly estimated to be roughly two orders of magnitude stronger than the background.

## Results (continued)

### Effect of record length on accuracy

- By prescribing the exact spectral characteristics of random phase time series, the sensitivity of constituent estimates to various constraints can be examined. An hourly, 46 month-long pseudo-tidal process with three relatively weak tides, tidal modulation, and a strong background of red spectral slope ( $\propto \omega^{-2}$ ) is modeled at a set of low frequencies as well as narrow tidal bands about the constituents (figure 4).

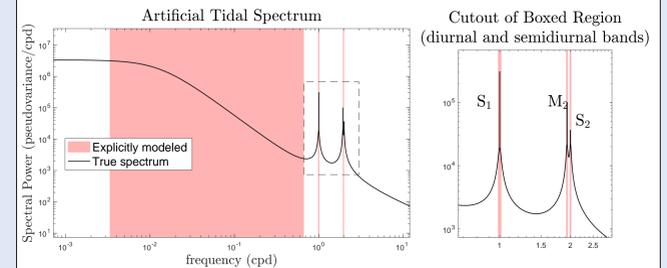


Figure 4. Partition of the true power spectrum into modeled (red) and unmodeled (white) bands, as performed by `red_tide`. Red bands are modeled explicitly by  $\mathbf{H}\mathbf{x}$ , and white signifies energy implicitly modeled as uncorrelated noise (or in the case of the lowest frequencies, a trend).

- The diurnal component is smaller than, but still comparable to, the non-tidal background, whereas the semi-diurnal components, while still prominent, are much less energetic than the lowest modeled frequencies. For 500 Monte Carlo simulations (figure 5), the pseudo- $S_1$  is modeled similarly by both `red_tide` and `t_tide`, though the latter underestimates amplitude when analyzing longer record lengths.
- The pseudo- $M_2$  and pseudo- $S_2$  (not shown) amplitudes are consistently overestimated by both methods, but there is much greater consistency between analyses of short and long segments of time series. Overestimation may be due to insufficient near-tidal frequencies being modeled.
- `t_tide`'s decrease of amplitude with record length therefore appears to be an artifact of the tidal analysis method. The observed decrease of tidal amplitudes in HYCOM output analyzed by Ansong et al. [1] may, therefore, not be the result of the circulation model but rather the tidal analysis.

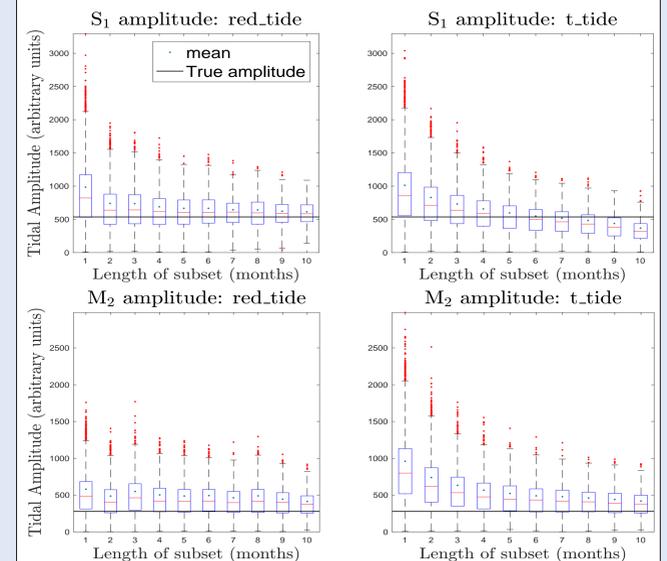


Figure 5.  $S_1$  (top) and  $M_2$  (bottom) as modeled by `red_tide` (left) and `t_tide` (right) over 500 Monte Carlo simulations (box and whisker plots are quartiles, red dots are outliers, mean and amplitude are labeled). Convergence for the weaker  $M_2$  signal appears to occur for shorter segments, which demonstrates the strength of this method: the prior information about the system's more energetic low-frequency band enforces variance to be partitioned approximately according to  $\mathbf{P}$ . This allows shorter records to yield amplitude estimates similar to those of longer records.

## Conclusion

- We have developed a software package `red_tide`, soon to be publicly available for use and improvement on GitHub at [https://github.com/lkach/red\\_tide](https://github.com/lkach/red_tide), which is intended to be used for oceanographic data with a strong, spectrally red background.
- Many processes in the ocean are not as strongly dominated by the barotropic tide as tide-gauge SSH, therefore `red_tide` may be preferable to `t_tide` in these cases. `red_tide` should also perform better on short time series, because its fitting scheme can account explicitly and realistically for frequencies lower than the fundamental frequency.
- SWOT will observe submesoscale processes for which non-stationary tides may be important, and it will also heavily alias them. Unlike the highly deterministic barotropic tide, the non-stationary component cannot be trivially removed. Understanding this component from *in situ* observations will be valuable for interpreting SWOT measurements.

## References

- Ansong, J.K., B.K. Arbic, M.C. Buijsman, J.G. Richman, J.F. Shriver, and A.J. Wallcraft (2015), Indirect evidence for substantial damping of low-mode internal tides in the open ocean. *Journal of Geophysical Research Oceans* 120, 6057-6071, doi:10.1002/2015JC010998.
- Kachelein, L., Gille, S.T., Cornuelle, B.D., Mazloff, M.R., 2019, in preparation
- Munk, W., and K. Hasselmann, 1964: Super-resolution of tides. *Studies on Oceanography*, 339-334, (Hidaka volume).
- Pawlowicz, R., R. Beardsley, and S. Lentz, 2002: Classical tidal harmonic analysis including error estimates in MATLAB using T TIDE. *Computers & Geosciences*, 28 (8), 929-937, doi: 10.1016/S0098-3004(02)00013-4.
- Wunsch, C., 1996, *The Ocean Circulation Inverse Problem*, Cambridge University Press
- Zaron, E.D., Ray, R.D., Egbert, G.D., Non-Stationary Tides and SWOT (Presentation) SWOT 2016 Science Team Meeting <https://swot.oceansciences.org/meetings.htm?id=1>
- Zetler, B. D., M. D. Schuldt, R. W. Whipple, and S. D. Hicks, 1965: Harmonic analysis of tides from data randomly spaced in time. *Journal of Geophysical Research*, 70 (12), 2805-2811.

## Acknowledgements

This work is supported by NASA award NNX16AH7G for the SWOT Science Team HFR Data provided by the Southern California Coastal Ocean Observing System (SCCOOS) of NOAA's Integrated Ocean Observing System (IOOS), available at <http://hfrnet-tds.ucsd.edu/thredds/catalog.html>

## Additional work

- The basis functions in  $\mathbf{H}$  are not required to be sinusoidal. An autoregressive model for modeling the non-tidal signals has been implemented to estimate background spectral slope (not shown).
- For time series that are expected to be forced by a different observable process (e.g. wind-driven currents), the response can be modeled as impulsive.
- Basis functions in  $\mathbf{H}$  are not required to be orthogonal; associated uncertainty accounts for resulting ambiguity.
- Analysis of HFR surface current data using these methods has yielded information about the energy and phase structure of tidal currents in the California Current System (meridional  $M_2$  velocity phase, right).

