JGR Oceans

RESEARCH ARTICLE
10.1029/2024JC021533

Special Collection:
Science from the Surface Water and Ocean Topography Satellite Mission

Tidal Corrections From and for SWOT Using a Spatially Coherent Variational Bayesian Harmonic Analysis

Thomas Monahan[1], **Tianning Tang**[1], **Stephen Roberts**[1], and **Thomas A. A. Adcock**[1]

[1]Department of Engineering Science, University of Oxford, Oxford, UK

Abstract The accuracy of global tidal models degrades significantly in coastal and estuarine regions. These models are important for correcting measurements from satellite altimetry and are used in numerous scientific and engineering applications. The new Surface Water Ocean Topography (SWOT) mission is providing measurements at unprecedented horizontal resolution in these regions. These data present both the opportunity and the necessity to quantify and correct the spatial variability in the model inaccuracies specific to these regions. We develop a variational Bayesian framework for tidal harmonic analysis which can be applied to SWOT, and is especially useful for exploting the data from the Cal/Val phase. The approach demonstrates superior robustness to different types of noise contamination in comparison to conventional least-squares approaches while providing full uncertainty estimation. By imposing a spatially coherent inductive bias on the model, we achieve superior harmonic constituent inference from temporally sparse but spatially dense data. Bayesian uncertainty estimation gives rise to statistical methods for outlier removal and constituent selection. Using our approach, we estimate a lower bound for the residual tidal variability for two SWOT Cal/Val passes (003 and 016) around the European Shelf to be 7% on average. We also show similar estimates cannot be produced using standard least-squares approaches. Tide gauge validation in the same region confirms the superiority of our empirical approach in coastal environments. Empirical corrections for the SWOT data products are provided alongside an open-source Python package, VTide.

Plain Language Summary Three decades of satellite altimetry has provided a rich data source for the study of oceanic processes. In order to use these data, and to study smaller processes, one must first remove larger oceanic signals: namely tides. Because the complexity of tides in coastal and estuarine regions, traditional modeling approaches often struggle. Although we know these models are not accurate in these regions, quantifying these errors remains a difficult task, as in situ measurements are sparse and often integrated into the models themselves. The new Surface Water Ocean Topography (SWOT) mission provides high-resolution measurements in these regions, presenting both the need and the opportunity to quantify and correct these inaccuracies. We develop a Bayesian approach for predicting tides in these regions capable of exploiting the spatial information in the data. Our method outperforms traditional modeling techniques in these regions and allows us to estimate errors in existing tidal models. These findings could be useful for analyzing the more than 30 years of historical satellite altimetry data and highlights the potential of our variational Bayesian method in other geophysical studies.

Key Points:

- Variational Bayesian harmonic analysis provides superior robustness to noise and uncertainty estimation of harmonic constituents
- A spatially coherent inductive bias further improves tidal and mean sea surface estimation
- SWOT data are useful for both quantifying and correcting errors in present geophysical models

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

T. Monahan,
thomas.monahan@eng.ox.ac.uk

Citation:

Monahan, T., Tang, T., Roberts, S., & Adcock, T. A. A. (2025). Tidal corrections from and for SWOT using a spatially coherent variational Bayesian harmonic analysis. *Journal of Geophysical Research: Oceans*, *130*, e2024JC021533. https://doi.org/10.1029/2024JC021533

Received 3 JUL 2024
Accepted 15 FEB 2025

Author Contributions:

Conceptualization: Thomas Monahan, Tianning Tang, Stephen Roberts, Thomas A. A. Adcock
Data curation: Thomas Monahan, Thomas A. A. Adcock
Formal analysis: Thomas Monahan
Funding acquisition: Thomas A. A. Adcock
Investigation: Thomas Monahan
Methodology: Thomas Monahan, Tianning Tang, Stephen Roberts, Thomas A. A. Adcock

1. Introduction

The NASA Surface Water Ocean Topography (SWOT) mission, launched in December 2022, is currently producing the highest spatial resolution satellite altimetric measurements of global oceans to date (Morrow et al., 2019). SWOT has the potential to greatly improve the field of coastal altimetry by providing high spatial and temporal resolution views into estuaries and coastal regions with measurements extending to within 50 m of coastlines. Although barotropic tides in deep oceans are essentially linear, interactions with coastal bathymetry, river outflow, and shallow water effects produce considerable nonlinearity. Because of data accuracy issues in historical altimetry and the increased nonlinearity present, the accuracy of global tide models tends to degrade in these regions with shelf and coastal errors in the derived M2 constituent ranging from 6%–10% to 7%–24% respectively (Stammer et al., 2014). Although, improved tidal analysis and prediction is, of course, useful for studying tides, arguably more important is the accurate removal of tidal variability from altimetric observations. This is a consequence of the fact that the tidal signal often dominates other mesoscale and sub-mesoscale processes. These processes are of high interest to SWOT researchers. Developing improved methods of tidal analysis

Project administration: Thomas Monahan, Thomas A. A. Adcock
Resources: Thomas A. A. Adcock
Software: Thomas Monahan, Stephen Roberts
Supervision: Tianning Tang, Stephen Roberts, Thomas A. A. Adcock
Validation: Thomas Monahan
Visualization: Thomas Monahan
Writing – original draft: Thomas Monahan, Stephen Roberts
Writing – review & editing: Thomas Monahan, Tianning Tang, Stephen Roberts, Thomas A. A. Adcock

and prediction is therefore critical to both the success of the broader SWOT mission and in utilizing the decades of satellite altimetry data that already exists.

A unique feature of the SWOT mission was the 102-day calibration and validation phase. Characterized by a fast-sampling orbit with near 1-day repeats, this provided the highest temporal resolution of ocean altimetry missions to date, capturing 8 full cycles of the aliased M2 signal. In addition to the improvements in temporal resolution, SWOT data provide sea surface height measurements at unparalleled horizontal resolution using two Ka-Band interferometers (KaRIn). Combined, these swaths cover nearly 120 kms and provide SSH data products at a resolution of (250 m $\times$ 250 m) globally between $\pm 77.6°$ latitude. This improved sampling can shed light on many short-scale (temporal and spatial) processes such as internal tides, storm surge, and the coupling of river discharge and tides in estuarine regions (Ray et al., 2020). Although this work does not make use of the 50 m $\times$ 50 m pixel cloud hydrology product, it should be noted that the methods developed herein can also be used on these data.

This work presents a variational Bayesian framework for spatially coherent tidal corrections which we apply to early stage ($\leq$1 year) of SWOT data. Our approach can be applied to any location, without prior knowledge of bathymetry or gauge constraints. Furthermore, our fully Bayesian framework provides accurate uncertainty estimation, allowing researchers to evaluate how much a given result should be trusted. This uncertainty captures both the intrinsic uncertainty from the data due to noise (aleatoric), as well as the uncertainty in the derived parameters due to insufficient information (epistemic). Quantifying epistemic uncertainty is particularly important for sparse reference series where insufficient data is a primary source of uncertainty. We develop several methods for utilizing the uncertainty information to remove outliers and determine whether tidal constituents should be included in a given analysis. This approach makes the incorporation of prior information simple while still generating predictions from the SWOT data. These features enable harmonic constituent "super resolution" (estimating constituents from reference series shorter/more sparsely sampled than required by Rayleigh Criterion/Nyquist. See Section 2.6). The proposed framework can function both as a tool for direct regional tidal correction and as a means of assessing the tidal errors present in current models. The latter is critical to address the spatial limitations of shallow-water and coastal tide gauge data sets used for validating global tide models (Arbic et al., 2015).

1.1. Empirical Correction Challenges and Philosophy

The unique spatiotemporal characteristics of the SWOT data and the challenges encountered when conducting tidal analysis on sparse data warrant the development of a specialized methodology. Here we look to briefly highlight some of these characteristics in order to motivate our methodological developments. The temporal sparsity intrinsic to satellite altimetry will always present a challenge for empirical correction of tidal phenomena due to aliasing of the harmonics. Although SWOT's Cal/Val orbit produces favorable alias periods for semi-diurnal constituents, with its near 1-day repeats, this phase lasted only 102 days. Thus, at best we are left with just over a hundred usable measurements per location. In practice, however, we encounter a second difficulty when performing tidal analysis. Although SWOT has extremely low measurement noise, early SWOT data has revealed several other sources of error (intertidal regions, roll errors, etc.), including the unavoidable "contamination" from non-tidal processes, especially storm surge (M. Hart-Davis et al., 2024; Lichtman et al., 2023). Even tidal processes can give rise to errors as a consequence of aliasing at certain frequencies (particularly diurnal constituents for SWOT) into the mean. These contributions are amplified by the limited amount of data. The resultant biasing of the derived constituents can lead to significant errors. Hence, the problem of deriving tidal harmonics from early-stage (<1 year) SWOT data are composed of two separate problems; true "aliasing" intrinsic to applying harmonic analysis to short and sparsely sampled reference series, and errors abounding from the interaction of tidal and non-tidal noise with the chosen estimator. Although little can be done for the former, this paper demonstrates that much can be done to address the latter.

The SWOT data have already been corrected with state-of-the-art tidal corrections from the assimilative finite element solution FES2014b (Lyard et al., 2021; Schaeffer et al., 2023). Our empirical correction philosophy is to treat this as a baseline, and only make corrections if the model's confidence in the derived parameters exceeds a given probabilistic threshold. Thus, if our model's confidence in the correction at a given location sits beneath this threshold, we can simply default to the standard geophysical corrections. Further, we also consider a soft fusion of geophysical corrections with empirical model based values. Here, a soft fusion refers to combining these data

probabilistically based on the model's confidence. Central to this philosophy is the ability to estimate when our model has gone wrong, which we accomplish by casting the problem within a Bayesian framework.

The objectives of this work are twofold:

1. Provide a framework to produce accurate uncertainty estimates for present tidal corrections.
2. Improve coastal and estuarine tidal estimates where possible.

The remainder of the paper is structured as follows. Section 2 provides a detailed overview of our approach. Next, we extensively test and validate the proposed method in Section 3 and compare it to conventional methodologies using synthetic data. Section 4 then evaluates the proposed method on the L2 2 km SWOT data product using a combination of in situ measurements from gauges and comparisons of SLA variance reduction. Estimates of the residual tidal errors are computed and presented with confidence intervals to inform researchers where current corrections are deficient. Finally, Section 5 discusses the limitations of the proposed method alongside avenues of future work and Section 6 states conclusions.

2. Methods

2.1. Harmonic Analysis

Harmonic analysis is the cornerstone of modern tidal analysis and prediction. The largely coherent response of tides to the periodic astronomical forcing which gives rise to them makes this procedure highly effective and has enabled accurate tidal predictions for more than a century (Parker, 2007). Modern harmonic analysis is carried out in the time-domain using least-squares estimation (Thomson & Emery, 2014), and in certain circumstances other weighted variants (Leffler & Jay, 2009). To achieve this, the harmonic analysis is formalized as a general linear model such that $y_i = w^{\mathsf{T}} x_i + \epsilon_i$. Here, y_i is the observed sea-level at time-step i, x_i is the i^{th} row of an $M \times N$ matrix of basis functions where M is the number of measurements and $N = 2n + 2$ with n equal to the number of constituents, w is a set of inferred weights, and ϵ_i is the nontidal residual. It is important to note that the assumptions made about the nature of ϵ have significant implications for the accuracy of the regression model used, which will be discussed later. The harmonic basis functions are defined within the design matrix X, with each column having the form

$$X = [1, t_i/t_{\max}, \sin\omega_0 t_i, \cos\omega_0 t_i, \ldots, \sin\omega_k t_i, \cos\omega_k t_i]^{\mathsf{T}} \tag{1}$$

where 1 is the mean sea level, $t_i/t_{\max}$ is the normalized linear sea-level trend, and the remaining entries corresponding to the quadrature basis functions for the k^{th} constituent. A nuanced discussion on how constituents are selected is provided in Section 2.6.1. Mathematical details of nodal corrections are omitted from the paper for simplicity but are implemented in the standard way (Le Provost, 2001). It is generally preferable to solve for the quadrature amplitudes A and B, though comparisons of tidal constituents are given in terms of the amplitude $C_k = \sqrt{A_k^2 + B_k^2}$ and phase $\phi_k = \arctan A_k/B_k$. We remark that A_k and B_k are elements of the weight vector w corresponding to the k^{th} set of quadrature basis functions. Note, the weights w correspond to the parameters of the model, and are different from the weights used in weighted least-squares. It is these amplitudes and phases which serve as the basis of how we study and describe tides globally.

2.1.1. Least Squares Parameter Estimation

Harmonic analysis constituent estimation is conventionally carried out using ordinary least squares (OLS). OLS estimation obtains parameter estimates which provide the most accurate description of the data in a least-squares sense. This imposes a strong prior assumption about the information content of the underlying data. In particular, conventional OLS assumes the residual term ϵ to be white Gaussian noise (WGN), with constant variance. In the case where this condition holds, the OLS parameter estimates are equivalent to those obtained via maximum-likelihood estimation (Myung, 2003) with weights w given by:

$$w = \left(X^{\mathsf{T}} X\right)^{-1} X^{\mathsf{T}} Y, \tag{2}$$

where $Y = [y_0, y_1, \dots, y_N]^T$ is a set of observations, and X is the design matrix above. Problems arise in parameter estimation when the noise content of the data deviates from these assumptions. Under these conditions, harmonic constituent estimates which yield the most accurate estimation of the data are not necessarily equal to the true harmonics of the system. Countless oceanic, hydrological, and meteorological processes give rise to nonstationary and temporally correlated noise, which violate these assumptions (W. Munk et al., 1965; Innocenti et al., 2022). These errors are propagated through the derived constituents and can yield significant biasing, such as the radiational double-counting problem (Williams et al., 2018). We note that although these issues have been partially addressed using weighted least-squares and other robust variants (Holland & Welsch, 1977; Leffler & Jay, 2009), these approaches encounter different challenges due to the extreme sparsity of the early Cal/Val SWOT data and the need to select appropriate regularization parameters. Finally, perhaps the most significant limitation of conventional harmonic analysis procedures with OLS and one of high salience to satellite altimetry is the poor estimation of model uncertainty (M. G. Hart-Davis, Piccioni, et al., 2021).

2.1.2. Least-Squares Uncertainty Estimation

It is worth remarking under what conditions OLS uncertainty estimation is accurate. A comprehensive discussion and validation of least-squares uncertainty estimation techniques for tidal estimation is provided in Innocenti et al., 2022, including IRLS estimators. Basic OLS confidence interval (CI) estimation is performed through comparison of the variance-covariance matrix of the regression such that

$$\mathrm{Var}(w) = \sigma^2 \left(X^{\mathsf{T}} X\right)^{-1}, \tag{3}$$

where $\sigma^2 = \sigma^2 I$ is the covariance matrix of errors given by

$$\sigma^2 = \frac{\epsilon \epsilon^{\mathsf{T}}}{M - N - 1}. \tag{4}$$

These estimates are accurate when the following assumptions are satisfied.

1. *Linearity:* The relationship between dependent and independent variables is linear and yields a zero-mean residual.
2. *Independent Variables:* There is no correlation between the estimated parameters and the errors.
3. *Homoscedasticity:* Errors have constant variance
4. *No temporal autocorrelation:* Errors are uncorrelated with each other.
5. *Gaussian Residuals:* Errors are normally distributed.

As will be discussed in Section 2.6, the limited Cal/Val data necessitates that only a subset of constituents be used in a given analysis. Thus, constituents excluded from an analysis are now treated as noise by the model potentially violating assumptions 2. As is seen in Equations 3 and 4, the OLS standard error does not explicitly consider the dependence of ϵ on X and can therefore lead to incorrect estimates. The uncertainty estimates computed in Figure A4 utilize this standard approach.

Several additional methods have been devised for confidence interval generation (Codiga, 2011; Leffler & Jay, 2009; Pawlowicz et al., 2002) and generally yield superior estimation in the presence of colored noise (Innocenti et al., 2022). However, because the focus of this work on application to extremely sparse SWOT Cal/Val data these are not applicable. For example, testing the Monte Carlo (MC) CIs in the UTide package yields SVD convergence errors. Similarly, the Linear CI estimation produces intervals containing NaNs due to the matrices used for inversion becoming rank-deficient (ill-conditioned) as the number of samples is reduced. In standard practice, these conditions are rarely encountered as the numbers of parameters is typically much smaller than the number of measurements.

Bootstrap methods, such as Moving Block Bootstraps and Semiparametric Bootstraps (only applicable to uniform sampling), generally yield superior uncertainty estimation to their analytical counterparts as they can better estimate the autocorrelation of the residuals (Innocenti et al., 2022). However, these approaches rely on separating residuals into many *sufficiently long* sub-series (blocks) of contiguous measurements. For the SWOT Cal/Val

data, the short and sparse records are insufficient for this procedure and thus bootstrap methods are not considered further.

Future work is needed to directly compare the uncertainty estimation between VBayes and conventional tidal confidence interval techniques using longer reference series with more regular sampling.

2.2. Bayesian Analysis

To combat the limitations discussed above and to achieve the objectives outlined in Section 1.1, we develop a fully Bayesian variant of harmonic analysis. The fundamental distinction between Bayesian and frequentist methodologies is the assumption that our parameters of interest are best represented by associated probability distributions, rather than as point values. This has the advantage that uncertainty associated with parameters in our models is explicitly represented.

Numerous studies have shown that tides can be modulated by various forms of external forcing leading to smearing of the tidal spectrum (W. H. Munk & Cartwright, 1966) and variability in tidal constituents (Matte et al., 2013; Mueller et al., 2014). As such, we argue modeling tidal constituents as distributions rather than discrete lines is (a) more physically accurate and (b) provides a more principled approach to handling this uncertainty. Under this assumption, we once again define the harmonic analysis problem as a general linear model $y_i = w^{\mathsf{T}} x_i + \epsilon_i$.

Using Bayes' theorem, we begin with a set of prior beliefs $p(\theta)$ about the distribution of all parameters (and hyper-parameters), which we denote θ. It is worth at this point noting that a *hyper-parameter* is itself a parameter, but one that governs the probability distribution of a *parameter*. For example, the mean and variance of a Gaussian are the hyper-parameters which define the distribution over another variable. By extension, the term hyper-hyper-parameter is then used to describe a variable that controls the distribution over a hyper-parameter. Bayesian inference rarely moves beyond this level of complication. The task of Bayesian inference is to update prior beliefs over the distribution of θ to a posterior distribution, given a set of observations $Y = \left[y_0, y_1, \ldots, y_N\right]^T$ and the design matrix X. This posterior distribution, which describes our final beliefs over θ, is given by:

$$p(\theta|Y) = \frac{p(\theta)p(Y|X,\theta)}{p(Y)}, \tag{5}$$

where θ represents the set of parameters and hyper-parameters of the model, $p(\theta)$ is the prior, $p(Y|X,\theta)$ the (data) likelihood, and $p(Y)$ the marginal likelihood, acting as a normalizing term in the inference (as it does not depend upon θ) (Wunsch, 2006).

2.2.1. Likelihood

The functional form of the (data) likelihood, $p(Y|X,\theta)$ in Equation 5, is consistent with the assumption of a least-squares error, thus is of Gaussian form. The squared residual, between observed y_i and model prediction, is weighted by hyper-parameter β, which represents a *precision* or inverse (co) variance. This leads to a Gaussian likelihood term of the form (and noting that the likelihood only depends upon w, β):

$$p(Y|X,\theta) = p(Y|X,w,\beta) = \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left\{-\frac{\beta}{2} E_Y(w)\right\}, \tag{6}$$

where we denote E_Y as the error functional, defined as:

$$E_Y = \sum_{i=0}^{N} \left(y_i - w^{\mathsf{T}} x_i\right)^2. \tag{7}$$

As described above, it is well documented that the tidal spectrum is super-imposed onto a red-noise background. To reduce the influence of this contamination, we introduce an automatic outlier determination procedure in Section 2.5 which exploits the assumptions of our noise prior. Testing using alternative noise priors, such as the

red noise covariance defined in Kachelein, 2023, while useful on high-frequency reference series, performed worse on SWOT Cal/Val data (0.99349 days sampling period). As such, this was not employed for this work.

Improving the handling of temporally correlated noise is an important area of future work, particularly in estuarine environments. In this vein, the VBayes framework can be modified to account for red noise by treating the residual as a mixture of Gaussians (S. J. Roberts & Penny, 2002). Modification with other noise priors such as a student-t or laplacian may also be beneficial.

2.2.2. Parameter and Hyper-Parameter Priors—$p(\theta)$:

The ability to place distributions over all parameters (and indeed all hyper-parameters) in Bayesian models is one of the great advantages of Bayesian inference. It requires selection of appropriate priors, which are updated to posterior distributions based on the data. We here detail the choice, and functional form, of these priors. Note that the prior over the parameters (weights) w depends only upon a set of hyper-parameters which govern the scale (or precision) of the (multi-variate) Gaussian over w. We denote this set of hyper-parameters $\alpha = \{\alpha_k\}$. Further, we note that the prior over θ factorizes as:

$$p(\theta) = p(w|\alpha)p(\alpha)p(\beta). \tag{8}$$

The first two terms factorize as $p(w|\alpha) = \prod_k p(w_k|\alpha_k)$ and $p(\alpha) = \prod_k p(\alpha_k)$. Each of the terms in Equation 8 are described below.

$p(w|\alpha)$—The model parameters (weights) w are drawn from a zero-mean Gaussian prior. For element w_k of w this is represented as follows:

$$p(w_k|\alpha_k) = \left(\frac{\alpha_k}{2\pi}\right)^{1/2} \exp\left\{-\frac{\alpha_k}{2}w_k^2\right\}. \tag{9}$$

In this way, each quadrature amplitude (and additional fit coefficients: sea-level trend, MSS, etc.) is modeled by its own Gaussian distribution with inferred precision α_k. This allows for the estimation of uncertainty over each weight w_k. This process is often referred to in the statistics literature as an *Automatic Relevance Determination* (ARD) prior (Bishop & Nasrabadi, 2006), as it induces a shrinkage of w_k that do not contribute to model-data fit. Shrinkage refers to constraining the size and number of coefficients, thereby favoring smaller models unless the data require additional parameters.

$p(\alpha)$—The weight precisions, α_k, are each drawn from a Gamma (hyper) prior with initial mixing hyper-hyper-parameters $a_0 = 10^{-2}$ and $b_0 = 10^{-4}$ such that

$$p(\alpha_k) = \Gamma(\alpha_k; a_0, b_0). \tag{10}$$

The values a_0 and b_0 are selected to yield a vague prior. This simply means that the prior distribution is broad. Often Bayesian statistics will refer to this as an *uninformative prior*. This provides natural parameter shrinkage without imposing a strong pre-experimental assumption about parameter values. (Ruanaidh & Fitzgerald, 2012). Furthermore, the selection of a Gaussian prior for w results in α being strictly positive, and thus a Gamma distribution is appropriate (S. Roberts et al., 2013).

$p(\beta)$—Finally, we model the precision (inverse variance) of the residual ϵ as a hyperparameter within the model. A Gamma prior is once again employed with mixing parameters $c_0 = 10^{-2}$ and $d_0 = 10^{-4}$ such that $p(\beta)$ is given by:

$$p(\beta) = \Gamma(\beta; c_0, d_0). \tag{11}$$

We note that a_0, b_0, c_0, d_0 are unitless as analysis can be run on normalized data, with final parameter estimates and predictions obtained by simply unnormalizing.

In the following section we briefly describe the principles behind our method of choice for performing inference: variational inference. This method allows us to update prior distributions over parameters of interest into the associated posterior distributions, conditioned on the observed data Y given the assumed model X.

2.2.3. Variational Inference

Our goal is to evaluate the posterior distribution $p(\theta|Y)$ for the harmonic model previously detailed. This evaluation requires us to infer distributions over the elements of θ, namely the (linear) weights w, the associated weight precisions α and the precision, β, of the residual between model and target. Even for a linear model, this set of inferences is not analytically tractable. We could use Markov-Chain Monte Carlo (MCMC) sampling, which is an effective approach to approximating the required distribution. However, MCMC approaches can be computationally intensive and scale poorly as the number of parameters increases (S. J. Roberts & Penny, 2002). This limitation becomes more salient when employing the spatially coherent harmonic analysis procedure we propose in Section 2.4. This is a consequence of the fact that the model simultaneously models the quadrature amplitudes for multiple sites which significantly increases the number of parameters. In our testing, the application of the Bayesian HA with MCMC sampling became computationally intractable beyond confined regions.

Instead, we advocate an approximate Bayes approach, referred to as *variational Bayes* (VBayes). Unlike, MCMC sampling, the VBayes framework offers a computationally tractable method of approximating the fully Bayesian inference of both parameters and hyper-parameters in the model. It achieves this by minimizing the Kullback-Leibler (KL) divergence between an approximate posterior $q(\theta|Y)$ and the true posterior distribution $p(\theta|Y)$. By imposing a functional form for $q(\theta|Y)$ which allows for computational tractability (Fox & Roberts, 2012), variational inference reduces the optimization of probability distributions to a problem of sequentially estimating the hyper-parameters of the distributions. This retains the benefits of fully Bayesian modeling without the need for sample-based approaches. This section is meant to serve as an introduction to the key concepts (and limitations) of variational inference—a detailed derivation of variational inference is provided in Fox & Roberts, 2012 and a more detailed exposition of the Bayesian linear model, used here, can be found in the appendix of S. Roberts et al., 2013.

As variational inference is an iterative optimization procedure, models are initialized using the maximum-likelihood (ML) solution given by

$$w_{ML} = \left(X^{\mathsf{T}}X\right)^{-1}X^{\mathsf{T}}Y, \tag{12}$$

where the n^{th} row of $X = x_n$ and has the standard form given in Equation 1 and $Y = \left[y_1, y_2, \ldots, y_N\right]^T$. The ML solution is then used to initialize the residual precision hyper-parameter, β, such that:

$$\beta^{-1} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - w_{ML}^{\mathsf{T}}x_i\right)^2 \tag{13}$$

Variational inference operates by defining a strict lower-bound to the posterior (but intractable) likelihood, $p(\theta|Y)$. VBayes introduces an *approximate posterior* $q(\theta|Y)$. The functional form of this approximation is not arbitrary, but is in accordance with the prior over θ, such that $q(\theta|Y)$ factorizes as:

$$q(\theta|Y) = q(w|Y)q(\alpha|Y)q(\beta|Y). \tag{14}$$

We then write the log of the data evidence, namely $\log p(Y)$, as the sum of two terms:

$$\log p(Y) = F(p(\theta|Y), q(\theta|Y)) + \mathrm{KL}(p(\theta|Y), q(\theta|Y)). \tag{15}$$

Equation 15 is the fundamental equation of variational inference. Understanding the two terms involved provides helpful intuition as to why this method is effective. The term $F(p, q)$ is known as the (negative) variational free-energy, or evidence lower bound (ELBO), and provides a *strict* lower-bound on the model evidence. Maximizing

$F(p,q)$ therefore results in the approximate posterior being as close as possible to the true posterior. The second term denotes the Kullback-Leibler (KL) divergence between the approximate and true posteriors over θ:

$$\mathrm{KL}(p(\theta|Y),q(\theta|Y)) = \int q(\theta|Y)\log\left(\frac{q(\theta|Y)}{p(\theta|Y)}\right)d\theta. \tag{16}$$

It can be shown that the KL term increases monotonically with the number of free parameters in θ and thus provides natural model shrinkage. Furthermore, the non-negativity of the KL divergence ensures the strictness of the variational bound.

Noting that the approximate posterior is factored (as per Equation 14), so $F(p,q)$ can be maximized by iteratively optimizing each of $q(\theta|Y), q(\alpha|Y), q(\beta|Y)$ separately. The update equations employed for this procedure create a series of sequential re-estimation equations with guaranteed convergence properties and are given in Penny & Roberts, 2002. Indeed, rather neatly, the ELBO of such a VBayes model is guaranteed to improve at each iteration. The more well-known expectation-maximization (EM) algorithm is a special case of the more general form (Dempster et al., 1977), in which the distributions over the variables of interest are delta functions (and so yield only point-value estimates, rather than full distributions).

2.2.4. Limitations and Considerations When Using VBayes

The VBayes approach is not without its limitations. First, we assume our model has conjugate priors. That is, the posterior $q(\theta,Y)$ will have the same form as the prior $p(\theta)$. In principle, this constitutes a strong pre-experimental assumption regarding the relationship between prior and posterior. We believe that our assumption of Gaussian distributed weights with Gamma distributed precisions is justified in the context of harmonic analysis after extensive analysis of constituent posteriors on synthetic and real data using a full Bayesian model in pymc-3 (a probabilistic programming language in Python) with MCMC sampling (Salvatier et al., 2016). Further the Gaussian and Gamma priors are also maximum entropy priors, and thus introduce minimal bias (Caticha & Preuss, 2004).

Second, VBayes (in the manner we use it) approximates joint distributions by products of marginals (Fox & Roberts, 2012). This is a consequence of the fact that we assume the factorized marginal densities of Equation 14 are separable under the mean-field approximation (Attias, 1999). The resultant parameter estimation is generally accurate, but can miss out on local maxima of $p(\theta,Y)$. This is illustrated in Figure 1 Panel A. By capturing the maxima of the posterior distribution our model provides more robust parameter estimates, at the expense of the full shape of the marginals, and thus potential under-estimation of uncertainty. As will be described in Section 2.6.2, we integrate the modeled noise variance β^{-1} into the final parameter level uncertainty estimates to reduce this underestimation. In altimetric tidal analysis, we are primarily interested in understanding how tidal constituents covary. VBayes explicitly considers the covariance between parameters rather than assuming independence as in least-squares estimators. Given the interaction and correlation (multicollinearity) between these constituents, the joint distribution offers insight into the global behavior of all parameters, and thus more accurate parameter estimation. This is the opposite of least-squares, which effectively models marginal distributions by assuming parameter independence. In doing so, least-squares approaches may overlook the complex relationships that exist between tidal constituents, potentially compromising the robustness of parameter estimates and the associated standard error estimates.

Comparisons of the marginals and joint distributions using MCMC sampling could provide interesting insights into "cross-talk" and aliasing between constituents. This represents an interesting area of future work, as these interactions encompass both physical nonlinear hydrodynamic interactions and spurious signal processing errors.

2.3. Comparisons With Other Estimators

The rationale for using VBayes becomes clearer when constrasted with other estimators commonly applied in tidal analysis. Here we compare our formulation with the Bayesian formulations of several common estimators employed for tidal analysis: OLS, iteratively-reweighted least-squares (IRLS) (Leffler & Jay, 2009), Lasso regression (L1 regularization), and ridge regression (L2 regularization).

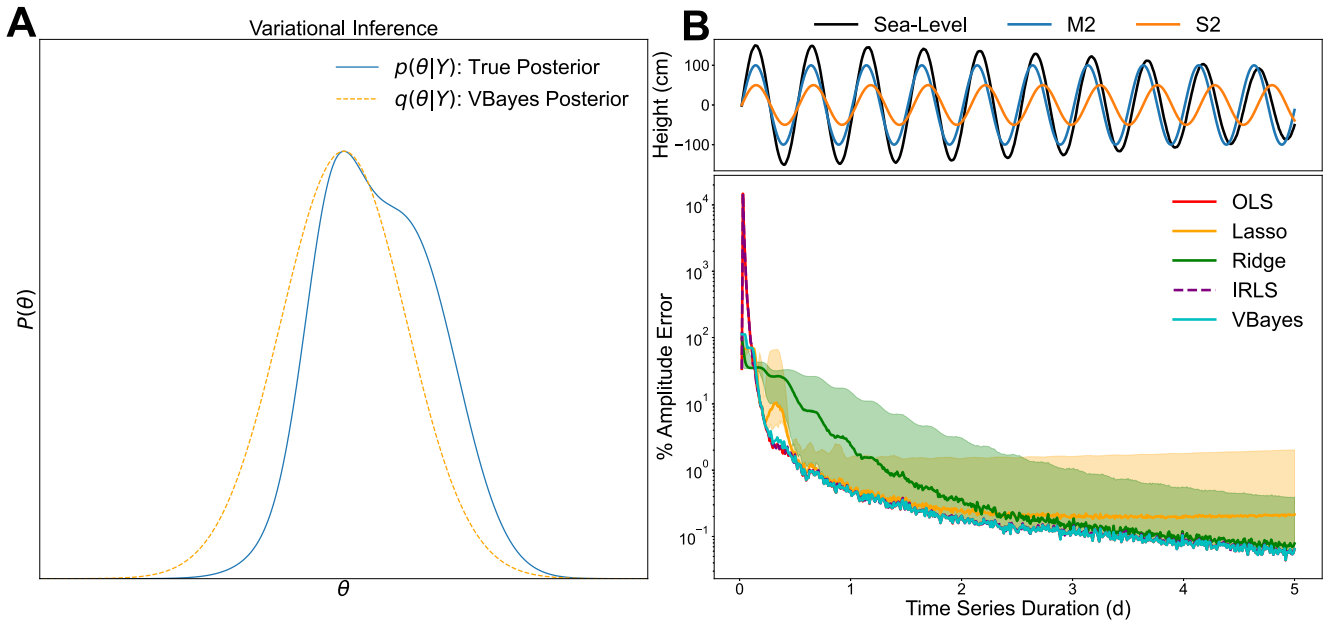

Figure 1. (a) Comparison between approximate VBayes posterior $q(\theta|Y)$ and true posterior $p(\theta|Y)$. (b) Comparisons of frequency resolution characteristics of different estimators. The synthetic signal is composed of two constituents M2 and S2 with amplitudes 100 and 50 cm respectively. The signal is contaminated with uniformly distributed noise in the frequency domain, as in (W. Munk & Hasselmann, 1964), with approximate SNR 1000. Amplitude errors for each estimator are shown as the duration of the time-series used for estimation is increased. Intervals for ridge and Lasso regression reflect the range of errors produced by varying λ between 0.01, and 1. Note errors are shown on log-axis.

VBayes can be understood as a regularized weighted least-squares regression, albeit with more flexibility and sophistication due to the way these features are incorporated via the prior and the likelihood.

Regularization: Within a Bayesian context, regularization is achieved through the selection of priors over the inferred weights w. The Bayesian equivalent OLS and IRLS estimators do not impose a prior over these parameters, and they are thus unregularized. For ridge regression, the prior is assumed to have the form of a zero-mean Gaussian such that $p(w) = \mathcal{N}(w\,|\,0, \lambda^{-1}I)$. This prior encourages smaller parameters, but does not force them to be zero. In contrast, Lasso regression assumes the pdf to be a laplacian (double-exponential) where $p(w) = \prod_j \frac{\lambda}{2}\exp(-\lambda|w_j|)$. This prior has the effect of forcing less-relevant parameters to be zero.

The "strength" of these priors is controlled by the regularization coefficient λ, which must be selected beforehand. This can significantly impact the efficacy of the estimator as shown in Figure 1 Panel B. Similar to Ridge regression, our VBayes approach also imposes a Gaussian prior over the weights w. However, rather than have a predefined λ, we instead model the weight precisions α_k as an internal parameter within the model. This removes the need to select a regularization parameter beforehand, and instead let's the data inform the degree to which regularization is necessary. This characteristic is valuable for application to satellite altimetry, where the diverse set of conditions and massive data sets prevent manual or iterative cross-validation based selection of λs.

As noted in Leffler & Jay, 2009 and can be seen in Figure 1 Panel B, IRLS and OLS estimators both exhibit a failure mode as the length-of-record is reduced. This failure mode leads to solutions containing significantly more energy than the original signal, and erroneous parameter estimation. Indeed, so too does the selection of an appropriate λ for Lasso and ridge regressions. These deviations are notable as the theory of tidal resolution, and indeed superresolution (W. Munk & Hasselmann, 1964), has focused on unregularized least-squares. The impact of the VBayes priors is readily seen, as the VBayes does not have the characteristic failure mode of OLS and IRLS while also avoiding bias introduced by the fixed regularization when sufficient data exist.

Weighting: Although regularization encourages less complex models by promoting sparsity or smaller parameter values, problems arise in the presence of outliers, which violate the assumption of Gaussian residuals. Even with

regularization, these outliers can significantly bias parameter estimates. The impact of outliers can be mitigated by *down-weighting* their influence. In a Bayesian context, this is achieved through the likelihood function, which can adjust for outlier effects in different ways.

As described above, the VBayes likelihood is given by

$$p(Y|X,\theta) = p(Y|X,w,\beta) = \left(\frac{\beta}{2\pi}\right)^{N/2} \exp\left\{-\frac{\beta}{2}\sum_{i=0}^{N}\left(y_i - w^{\mathrm{T}}x_i\right)^2\right\},$$

where β is inferred alongside w. The inverse precision, β, weights the influence of the individual data points, modeling the uncertainty based on the observed noise level. Hence, by fitting β within the model, the second term weights the observations based on the weighted residuals $\left(\beta^{-1}\right)$ such that observations with smaller residuals contribute more to the model fit and vice versa. We note that OLS, ridge, and Lasso regressions all assume a fixed β estimated directly from the residual, and thus do not down-weight outliers in this way.

For the case of IRLS, three primary differences exist. First, rather than model β within the model, a tuning parameter τ is used, which scales the influence of outliers on the model (e.g., small τ down-weights outliers more and vice versa). Just as with Lasso and ridge regressions, a weakness of this approach is the need to specify this parameter beforehand which can lead to over or under-fitting. Conversely, by modeling β explicitly, the equivalent parameter tuning in VBayes is accomplished implicitly through posterior inference. The second difference in IRLS is that the scaled residual is fed into a "weighting function", such as Huber's function or Tukey's bi-square. This is equivalent to modifying the assumed functional form of the residual in the likelihood to a more heavy tailed distribution. For these weighting functions, the most similar associated pdfs would be either a Laplace or Student's-t distribution. Third, IRLS introduces an additional weighting matrix, W, which explicitly weights each data-point. In a Bayesian context, this would be equivalent to introducing an additional set of latent weights into the model which weights each measurement within the likelihood function (Wang et al., 2017). In this way, although our VBayes approach down-weights the impact of outliers without requiring a tuning parameter, the more explicit down-weighting within the IRLS approach of Leffler & Jay, 2009 can be seen as more robust to outliers, if an appropriate tuning parameter can be selected. However, although this can increase model robustness, our testing found IRLS to struggle on the SWOT Cal/Val data. We attribute this to the extreme sparsity of the data. To increase the robustness of our approach to outliers, we introduce an automated outlier detection procedure in Section 2.5 using the intrinsic uncertainty estimation with our method.

2.4. Spatially Coherent Harmonic Analysis

Inspection of early SWOT data confirms our physical intuition that the oceanic response is spatially coherent. This fact, while interesting, cannot be fully exploited using single point measurements such as tide gauges. Empirical tidal models from satellite altimetry already capitalize on this, even if indirectly, through binning of measurements into fixed grids (Schrama & Ray, 1994; M. G. Hart-Davis, Piccioni, et al., 2021). Similarly, several works have attempted to use nearby gauges to further constrain solutions through interpolation (Matte et al., 2018). As will be discussed, SWOT data have revealed conventional binning, averaging, and interpolation procedures can mask short-scale features in the derived tides. Swath-based altimetry provides high-resolution spatial measurements whose spatial information can be exploited directly if an appropriate method can be developed. Here we develop a spatially coherent inductive bias, which can be imposed on the model. We note that an inductive bias differs from a prior, in that an inductive bias imposes an assumption about the relationship between data points (here we assume a spatial relationship), whereas a prior refers to our prior belief about the distribution our parameters come from.

Spatial coherence simply implies that there is information relevant to our point of interest in the surrounding points. Mathematically, we can define this as a single convolution between the central point $p_{0,0}$ and the surrounding points $p_{j,k}$. Recall that our general harmonic analysis model can be defined as the linear model $Y = w^{\mathrm{T}}X + \epsilon$. Once again, Y is the set of all measurements, X is an $M \times N$ matrix of basis functions where M is the number of measurements and $N = 2n + 2$ with n equal to the number of constituents, and w is the set of derived weights. To allow information to propagate to the central point we modify this procedure to

simultaneously regress the neighboring points by stacking the measurements, and associated weight and design matrices. This imposes the assumption that the sea surface at neighboring points $p_{j,k}$ is given by the sum of the sea level at the central point $p_{0,0}$ and an additional offset.

Measurements and design matrices are padded with zeros for neighboring points such that after dropping zero coefficients the sea level at $p_{j,k}$ is predicted by

$$Y_{j,k} = X_{0,0}w_{0,0} + \rho \cdot d \cdot X_{j,k}w_{j,k} \tag{17}$$

where d is the normalized distance between $p_{0,0}$ and $p_{j,k}$ taken to be (0,1) for uniformly spaced data, ρ is the probability that the observations $Y_{j,k}$ are correlated with $Y_{0,0}$. This is obtained by taking the Fischer transform of the Pearson's correlation coefficient r between $Y_{0,0}$ and $Y_{j,k}$, such that $z' = .5[\ln(1 + r) - -\ln(1 - r)]$, and $\rho = 1 - \Phi(z')$ with $\Phi = \mathcal{N}(0,1)$. This weighting imposes a simple physical assumption; the more similar the time-series is between two points, the closer their harmonic constituents should be. Hence, in the limit where both points are identical, $Y_{j,k} = X_{0,0}w_{0,0} + 0 \cdot 1 \cdot X_{j,k}w_{j,k}$. This formulation is equivalent to a first-order spatial autoregressive process, or SAR (Anselin, 2013). By simultaneously inferring weightings across points based on the central point, the central point can be better resolved. Although this procedure down-weights the entries of the design matrix for the neighboring points, the ratio of parameters to data-points is actually higher than in the single-point case. One can do better if one imposes more restrictive assumptions regarding the relationship between neighboring points.

For closely spaced points (the present approach has been tested up to 4 km spacing), we modify the above procedure based on the fact that the relationship between the quadrature amplitudes of neighboring points is approximately linear. This assumption allows us to describe the relationship between the center point and its neighbors using only two lines. The following procedure is designed for the structured grid-data produced for the L2 data products but is easily adapted to unstructured data. The lines are defined by

$$Y_{j,k} = \rho \cdot (d_x \cdot X_{j,0}w_{j,0} + d_y \cdot X_{0,k}w_{0,k}) + X_{0,0}w_{0,0} \tag{18}$$

where d_x and d_y are the normalized distances between points along the horizontal and vertical direction respectively, taken to be $(-1, 0, 1)$ for convenience. For SWOT, these can be thought of as representing the distance along the along-track and cross-track directions, respectively. A 3-d visualization of what this looks like for a set of points is provided in the Supporting Information S1. Using four neighboring points yields an increase in the effective number of measurements to parameters by nearly a factor of two (1.6–2.9). It is possible to incorporate more than four neighboring points on a nonuniform grid by weighting the distances appropriately. The results presented in this paper only utilize points separated by up to 4 km (e.g., only consider neighboring SWOT pixels). As will be shown in Section 3.2.1, this spatial coherence can significantly improve the accuracy of constituent estimation when using both sparse data and data with noise artifacts. The decision to use only nearest neighbors in this work reflects our objective of incorporating spatial coherence without imposing unrealistic assumptions about the spatial distribution of amplitudes. Although amplitudes may smoothly vary (in a linear sense) for much longer distances, the quadrature amplitudes are not necessarily linear and so care should be taken before applying this procedure to coarser data. Further validation is needed to determine where the breakdown point lies for the linearity assumption. This procedure can be leveraged using any conventional estimator, including OLS and its variants. Although we believe the VBayes procedure offers several intrinsic advantages, conventional OLS augmented with spatial coherence worked well in low-noise regimes. These results are given in the Figure S3 in Supporting Information S1.

2.5. Bayesian Pruning

Because of the presence of erroneous or contaminated measurements in coastal regions, statistical procedures for outlier removal typically fail to remove values contaminated by other nonstationary artifacts (e.g., storm surge, roll errors, satellite maneuvers). Although the VBayes approach is generally more robust to such artifacts than OLS, their presence will still cause biasing of the derived constituents. Implicit within the VBayes framework is uncertainty estimation over both the parameters and the data. This feature can be leveraged to flag and even remove points that exhibit behavior inconsistent with our model's expectation (e.g., stationary tidal signal with a

linear trend and mean sea-level). The procedure is remarkably simple. For each measurement, the model computes a predictive distribution with mean μ_i and standard deviation σ_i. The computed standard deviation is composed of the contribution $\sigma_i = \sigma_p + \sigma_{\text{noise}}(i)$ where σ_p is the sum of the standard deviations from the derived constituents and $\sigma_{\text{noise}}(i)$ is the contribution to the standard deviation from the i^{th} measurement. Measurements which sit greater than $3\sigma_i$ away from the mean are removed. The selection of this threshold was empirically determined using synthetic data. Testing of this procedure is provided in Section 3.2.2. We note this algorithmic pruning procedure echos the arguments provided in Leffler and Jay (2009) while removing the need for parameter tuning which creates complications for unsupervised application to satellite measurements.

2.6. Constituent Selection

Arguably, the most important consideration in any harmonic analysis is the choice of which constituents to include. These decisions have major implications for the accuracy and bias of the derived harmonics (Parker, 2007). Indeed, these decisions are made more salient when attempting super-resolution on sparsely sampled data, as is often the case with early stage SWOT data. Several methods have been proposed to determine which constituents can be obtained from a given reference series. The most well-known, and most often used, is the so-called Rayleigh criterion. Rayleigh defines the minimum uniformly sampled time-series duration (T) in order to separate two constituents ω_1 and ω_2 as

$$\left(\frac{T}{\omega_1 - \omega_2}\right) / R_{\text{min}} \geq 1, \tag{19}$$

where R_{min} is typically taken to be 1. Numerous works have shown that this criterion is often overly restrictive and that in practice one can resolve additional constituents depending on the signal-to-noise ratio (W. Munk & Hasselmann, 1964). As shown in Figure 1 Panel B, regularized estimators exhibit different frequency resolution characteristics from their unregularized counterparts, and can often perform better particularly in the presence of noise. As such, in practice we find even modified constituent selection approaches, which consider SNR to be insufficient. Further discussion on this is provided in Section 5. Rayleigh is often considered in conjunction with the Nyquist sampling theorem which states that for a given frequency ω to be resolvable, the associated signal must be sampled with sampling frequency $f \geq 2\omega$ (Landau, 1967). Because the tidal signal has a sparse Fourier basis, accurate constituent estimations can be obtained using sampling frequencies far below the Nyquist frequency (Mishali & Eldar, 2010) as is confirmed by the success of empirical tidal models from historical altimetry (M. G. Hart-Davis, Piccioni, et al., 2021).

We would like to emphasize that the selection of constituents goes beyond avoiding aliasing. Even for constituents which are resolvable according to Rayleigh, simple testing using OLS HA on synthetic data demonstrates that inclusion or exclusion of a given constituent can degrade/improve the accuracy of other constituents significantly. This is a consequence of how co-linearity, or cross-talk (Parker, 2007), interacts with different estimators. Simple comparison of different estimators yields different results. Hence, we argue that to maximize the effectiveness of a given analysis, the choice of constituents must be done in conjunction with the estimator employed for analysis.

To address this, we propose a data-driven approach which can efficiently explore and produce optimal combinations of constituents for any uniform or nonuniform sampling period. This approach is used to obtain a robust global optimal constituent solution for the SWOT Cal/Val orbit.

2.6.1. Evolutionary Optimization for Constituent Selection

Because of the massive data volume from the KaRIn, an iterative approach to constituent selection is computationally infeasible. However, based on the known sampling period (≈0.99349 days for Cal/Val), we can determine an optimal solution for our variational Bayesian harmonic analysis using synthetic data generated from the TICON-3 data set (M. Hart-Davis et al., 2022). The data set provides a diverse set of tidal regimes containing various levels of nonlinearity and is thus suitable for our purposes when searching for robust solutions. In order to efficiently explore potential solutions we utilize a genetic algorithm (GA) to perform evolutionary optimization. We define a candidate solution as a binary string with each digit corresponding to a constituent. A value of 1 means the corresponding constituent is used in the analysis, and a value of 0 means it is excluded. We restrict the

set of candidate constituents to the 40 contained within the TICON-3 data set (M. Hart-Davis et al., 2022). Although this set of constituents accounts for a majority of global tidal variance, the lack of consideration of additional constituents and potential influences of their co-linearity is a notable limitation.

Evolutionary optimization is performed using a multi-objective loss function and is tasked with simultaneously minimizing the RMS error of the M2 constituent, while maximizing the number of constituents with median percentage errors below 10%. The choice of median percentage RMS error is commensurate with our objective to produce a robust model, although we note that the optimal solution for median percentage RMS error is also the optimal solution for the mean RMS error. Interestingly, we also found that the optimal solution for M2 often coincided with the optimal solutions of several other constituents. By definition, the best parameter estimates will occur when the estimator is the least biased. This occurs when the residual is independently distributed with equal variance. Hence, the optimal set of constituents will be the set, which most closely satisfies this assumption. The TICON-3 data set is used to generate 2,569 pure tidal signals which are contaminated with moderate white Gaussian noise (100 SNR) and nonstationary noise (one event with signal-to-surge ratio of 0.1). Details for how signals are contaminated with white and nonstationary noise is provided in Sections 3.1.1 and 3.1.2 respectively. Additionally, the GA implementation can be found in the VTide GitHub repository. For each candidate solution, 300 stations are selected at random from the TICON-3 data set, time-series are then generated based on the SWOT Cal/Val sampling period and noise is applied. The evolutionary optimization is run for 100 generations (30,000 models) with the objective of obtaining a representative set of candidate solutions. This data set provides interesting insights into the relationship between individual and sets of constituents and the Cal/Val sampling period.

Visualization of the top solutions indicates strong positive and negative "lines" for different constituents. However, noise exists depending on which constituent we choose to optimize for. To account for this, we select the top 20 genomes for the primary constituents M2, N2, S2, and sum the total number of occurrences for each constituent using both median percentage RMS error and average RMS error. The resulting sums are divided by the total number of possible occurrences to yield a probability of inclusion. We select an empirically determined threshold for inclusion of 80% based on experimentation using both synthetic and real SWOT data. For the real data, inappropriate selection of constituents leads to visually "fragmented" phase estimates where interacting constituents will bias others. This interrupts the spatial coherence which is expected and observed when appropriate constituents are selected. It was found that the smoothest solutions were produced by the 80% threshold.

The global optimal solution found is comprised of M2, N2, S2, O1, M4, and MN4. We find that this deviates slightly from the optimal solution for semi-diurnal regions which also includes M6 and M8. In some instances, such as the Bristol channel, we have found that many additional minor constituents such as M10, MU2, 2N2, NU2, and MU2 can reliably be obtained, however, they are not necessarily reliable outside of macro-tidal regimes. The dominant diurnal constituent K1 was found to significantly degrade the accuracy of the analysis likely due to close proximity between the K1 period and the SWOT Cal/Val sampling period (1.0027 vs. 0.9934 days). Although the evolutionary procedure is imperfect, it provides an alternative to practitioner trial and error, the imperfections of Rayleigh, and provides us with a reliable solution for unsupervised applications. We note, however, that although this makes sense for SWOT with the fixed sampling period, it is computationally impractical for standard harmonic analysis. Under these conditions, the ARD priors, in conjunction with the superior uncertainty estimation, could aid in constituent selection but is not considered further due to the sparsity of the SWOT data.

2.6.2. Bayesian Constituent Uncertainty Estimation

It remains to be said how the quadrature amplitudes (μ_A, μ_B) and associated amplitude variances (σ_A^2, σ_B^2) can be related to the final uncertainty estimate. The following procedure closely relates to the linearized analysis in Pawlowicz et al. (2002).

First we compute the equivalent amplitude and phase $C = \sqrt{\mu_A^2 + \mu_B^2}$ and $\theta = \arctan 2(\mu_B, \mu_A)$. Next, we assume the contribution of the noise variance $\sigma_{\text{noise}}^2 = \beta^{-1}$ (aleatoric uncertainty) to the given constituent's uncertainty is proportional to the relative magnitude of the constituent $\upsilon = C / \sum_i^{2N} w_i$. It follows that the total

variance associated with each component is given by $\sigma^2_{A_{tot}} = \sigma^2_A + \upsilon \cdot \sigma^2_{\text{noise}}$, and $\sigma^2_{B_{tot}} = \sigma^2_B + \upsilon \cdot \sigma^2_{\text{noise}}$. Through a laborious derivation, the amplitude variance can be expressed as follows:

$$\sigma^2_C = \left(\frac{A^2}{A^2 + B^2}\right)\sigma^2_A + \left(\frac{B^2}{A^2 + B^2}\right)\sigma^2_B, \tag{20}$$

with the phase variance given by

$$\sigma^2_\theta = \left(\frac{B^2}{\left(A^2 + B^2\right)^2}\right)\sigma^2_A + \left(\frac{A^2}{\left(A^2 + B^2\right)^2}\right)\sigma^2_B. \tag{21}$$

These values can then be used to generate RMS uncertainty estimates using the RMS formula given in the Appendix A1). Robust uncertainty estimates in amplitude and phase are achieved by incorporating both epistemic and aleatoric uncertainties. This yields dynamic and locally accurate uncertainty estimates as shown in Section 4.1.

3. Simulated Results

3.1. Comparisons Between OLS and VBayes for SWOT Cal/Val Orbit

Prior to assessing the impact of the additional modifications outlined in Sections 2.4 and 2.5, it is worth first comparing the behavior of the OLS and VBayes methods under different conditions. To do this, we once gain utilize the TICON-3 data set (M. Hart-Davis et al., 2022). Although this cannot represent all global tides, the data provides a comprehensive distribution of tidal regimes (semi-diurnal, mixed, and diurnal), tidal ranges (micro, meso, and macro), and varying degrees of shallow-water effects. Synthetic reference series are generated for each tide gauge at the SWOT Cal/Val sampling period (0.99349 days) unless explicitly stated otherwise. This sampling does not represent crossover-points, which will have twice as many measurements. The two methods are then compared with variable types of noise contamination and sampling periods in the following sections. Models are evaluated on the root mean squared error (RMS) (defined in Appendix A1) of the primary constituents in the "optimal" GA solution ($M2, N2, S2, O1, MN4, M4$). The root-sum-of-squares (RSS) of all included constituents is also presented. The definition for the RSS is given in 7.1. The present analysis only compares OLS and VBayes as OLS is the relied upon method for present empirical tidal analysis of altimetry data (M. G. Hart-Davis, Piccioni, et al., 2021). Although this work focuses on conventional harmonic analysis, the VBayes estimator is similarly applicable to the response method (W. H. Munk & Cartwright, 1966; Andersen et al., 2023). As described in Section 2.3, OLS is not the only estimator employed for tidal analysis. In our testing, we found IRLS to perform significantly worse than both the VBayes and standard OLS on data with SWOT sampling periods. Because of the near 50% increase in RSS error we have not included it in the subsequent analysis. Additionally, given the need to specify a λ parameter beforehand, ridge and Lasso regressions are also not considered further.

3.1.1. White Noise

To test the two methods resilience to uncorrelated and normally distributed noise, the simulated reference series are contaminated with white Gaussian noise (WGN) at signal-to-noise ratios (SNR) ranging from 0 to 100 dB. For each SNR, trials are repeated 100 times for all simulated series with the mean and median values for the 100 trials presented in Figure 2 Panel(s) B. It can clearly be seen that for high SNR ratios, the OLS and VBayes are approximately equal. However, as the noise floor is increased, there is a pronounced divergence at approximately 40 SNR in which the OLS starts to perform significantly worse, whereas the VBayes remains relatively bounded. This behavior is consistent across the individual constituents and the RSS error. Interestingly, whereas the WGN satisfies the normality assumption of the error term by the OLS, the OLS still breaks down as the WGN is increased. This is a consequence of the fact that the OLS estimator's "breakdown point", defined as the minimum amount of contamination needed to cause an estimator to take on erroneous values (Donoho & Huber, 1983), is proportional to the inverse of the sample size (Rousseeuw & Leroy, 1987). Conversely, the compact support supplied by the priors in the Bayesian framework yield an estimator which in principle never breaks down (Donoho & Huber, 1983). The degree of this support, however, will of course depend on the strength of the priors.

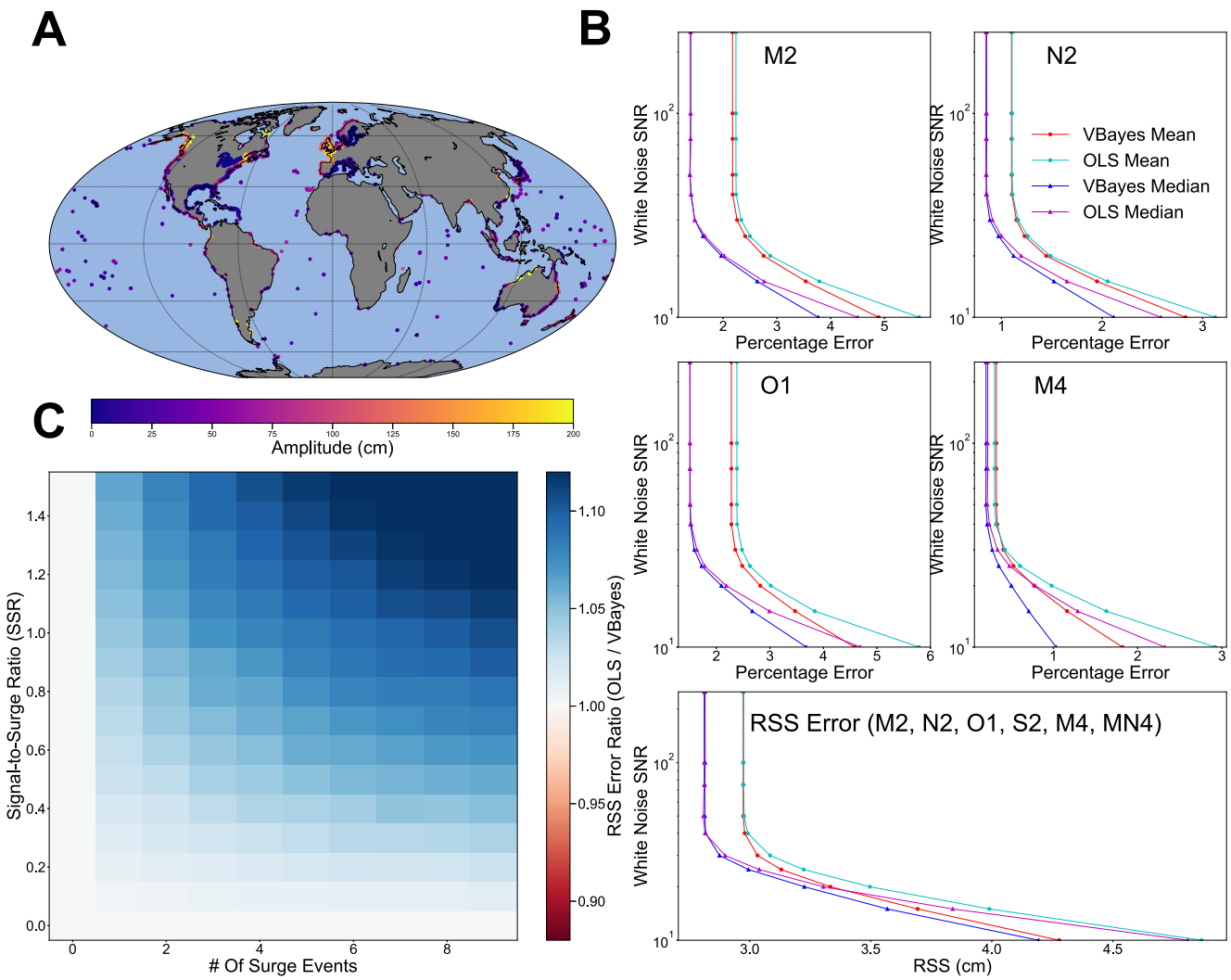

Figure 2. Comparisons of OLS and VBayes on synthetic TICON-3 data set when contaminated with different types of noise. Panel A shows the distribution of TICON-3 tide gauges used to generate the synthetic reference series employed for testing. Panel B compares OLS and VBayes when contaminated with increasing amounts of white gaussian noise. The mean and median values for the VBayes are given in red and blue. The mean and median values for the OLS are given in cyan and magenta. Panel C compares the ratio of OLS to VBayes RSS values at the associated signal-to-surge ratio and number of events (Blue, VBayes better; Red, OLS Better; White, Equal). For both Panels B and C the RSS is computed for the for the constituents (M2, N2, S2, O1, M4, and MN4).

3.1.2. Nonstationary Noise

To simulate nonstationary artifacts we contaminate the simulated tidal signals with several synthetic surge events. We utilize the formula outlined in Cartwright's work on surges (Cartwright, 1968) such that a synthetic surge is given by $\zeta_S = A \cdot (1 + \cos(2\pi t/l))$ where $6 \leq l \leq 96$ (hours) is the duration of the surge, defined from $0 \leq t \leq l$ and $0.1 \leq A \leq 1.5$ is the "surge-to-signal ratio" (SSR) of the surge defined by the ratio of the surge magnitude ζ_S to one half of the tidal range. Tests are repeated by varying both the number of surges N occurring within the 90-day Cal/Val period and the associated SSR. The range of SSR values and the frequency of events is based on empirical analysis of the residuals from the UK operational surge model (Williams & Horsburgh, 2013). In reality, both extremes (negligible noise, or extreme noise) are unrealistic. Hence, the objective of this case study is to show the range of possible scenarios to allow researchers to make informed decisions about the performance of each estimator. Surges are randomly added to the signal without overlap. For each SSR and number of occurrences, tests are averaged over the full TICON-3 data set (2,569 signals) and the mean error ratios defined as $RMS_{OLS}/RMS_{V\ Bayes}$ are plotted in heatmaps in Figure 2 Panel C. Signals are contaminated with a baseline 100 dB

of *WGN* to prevent exact inversion by OLS. Similar to the behavior shown in Section 3.1.1, Figure 2 Panel C shows that for minimal nonstationary contamination the two methods are statistically equivalent. However, as the number of surge events or surge intensity is even slightly increased, the two methods diverge with the VBayes yielding superior performance. This behavior conforms exactly to what is theoretically predicted by the two estimators when applied to data contaminated by non-normally distributed noise. Although the present analysis has used synthetic surges to simulate nonstationary contamination, similar results would be expected in the presence of other nonstationary artifacts (e.g., roll errors, signal dropout,).

3.1.3. Sampling Period

Because the global disparities in revisit time for locations sampled in the Science phase of the SWOT mission, and to demonstrate the salience of our method for application on historical altimetry data, we compare the performance of the two methods as the sampling period is changed. Three-year synthetic signals are generated with increasing sampling periods from 1 hr (Standard Tide Gauge Sampling Period) to 21 days (SWOT Science orbit repeat period). During the SWOT Science orbit, locations will be sampled multiple times within a given cycle, thus the sampling period will vary geographically. Once again a baseline 100 dB of *WGN* is added to prevent exact inversion. VBayes and OLS are run on the signals using all 40 harmonics in the TICON-3 data set in the analysis. A figure showing the median RMS error of several constituents and the associated RSS of each estimator and the associated constituents is provided in Figure A1 in Appendix A. Compared to the clear improvements in robustness shown in Sections 3.1.1 and 3.1.2, the performance characteristics of each estimator are far more complex when varying the sampling period. The duality of the time-frequency domain is, of course, a contributing factor—interactions between the sampling rate and constituent frequencies can induce aliasing and distortion effects. It can be seen that the two estimators behave differently for different constituents (e.g., sometimes OLS exhibits "aliasing" behavior, whereas the VBayes does not, and vice versa). In general, we observe that the VBayes is more robust to true aliasing. This "robustness" comprises two aspects: constraining the derived values of aliased constituents, and ensuring the errors these present do not propagate to other non-aliased constituents. Both of these aspects are supported by the ARD priors over the weights. This result is particularly promising for sun synchronous missions such as Envisat in which problems arrise if solar constituents are included in an analysis. Indeed, even when combining these orbits with non-sun-synchronous missions errors can be introduced using ordinary least-squares. Simulated results for several constituents (including S2) when sampled at the Envisat sampling period (35 days) are provided in the Supporting Information S1. The reduced biasing of the other constituents is a consequence of the regularization from the Bayesian ARD priors (Bishop & Nasrabadi, 2006) which reduce the influence of parameters that do not contribute to model-data fit. However, further study is needed to understand this behavior. We anticipate these characteristics could be further exploited to utilize both estimators in tandem to offset periods when one outperforms the other.

3.2. Simulated Testing of Spatial Coherence and Bayesian Pruning

Now that the disparities between conventional OLS and VBayes have been established, we now look to validate the spatially coherent harmonic analysis as well as the Bayesian pruning procedures proposed in Sections 2.4 and 2.5, respectively.

3.2.1. Spatial Coherence

To validate the spatially coherent harmonic analysis procedure, we extract harmonics from a year long 2-D depth averaged simulation of the tides in the Pentland Firth, UK. The model is forced with constituents M2, K1, K2, MU2, N2, NU2, O1, and S2 at the ocean boundary and has been tested and validated against in situ measurements in Adcock et al. (2013). The tidal heights in the Pentland Firth are largely semi-diurnal and are characterized by elevated shallow water effects induced by islands throughout the channel and some of the fastest tidal currents globally. The presence of simulated tidal stream turbines in the center of the channel amplifies these effects and produces relatively fast changes in tidal amplitudes due to the head drop as energy is extracted. The data offer a challenging stress test of the spatial coherence in complex coastal environments. The nonuniform grid-points are binned into a 2 km $\times$ 2 km grid to simulate the SWOT ocean data products and we restrict our study region to the center of the channel which is characterized by the greatest spatial variability. A visualization of the M2

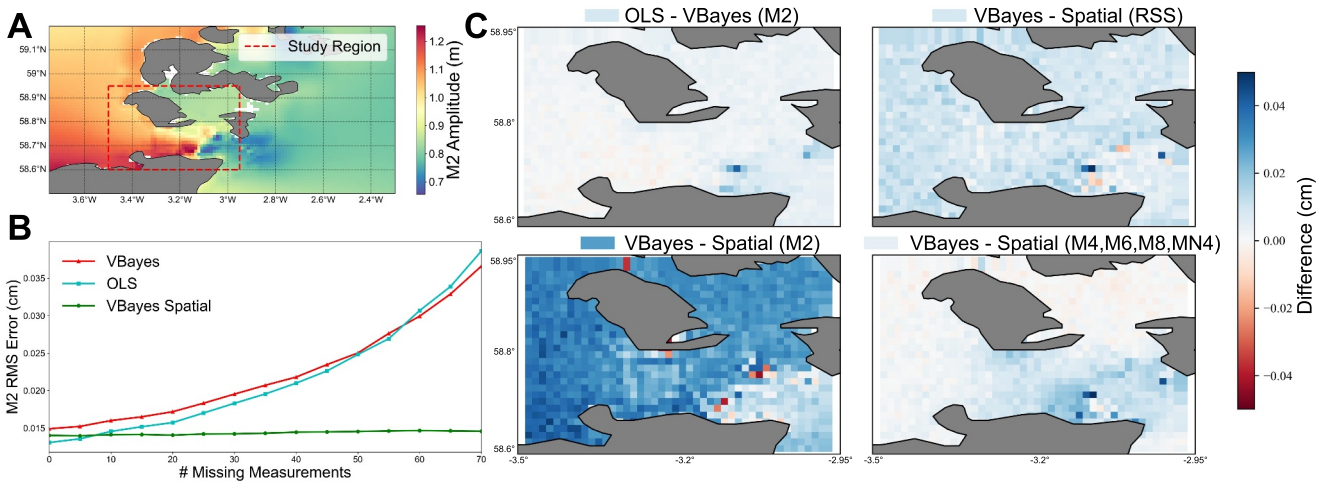

Figure 3. Validation of the VBayes spatial coherence. Panel A shows the simulated M2 amplitudes in the Pentland Firth and the study region used in Panels B and C. Panel B compares OLS, VBayes, and the spatially coherent VBayes as datapoints are randomly dropped (originally 102 points). Panel C compare the average difference in RMS error between the given models when contaminated with spatially incoherent noise (Red, first model better; White, same; Blue, second model better). The average difference is given next to the title of each plot.

amplitudes for the full simulation, and the 3,041 node study region is shown in Figure 3 Panel A. We test the proposed method under several conditions; random data dropout, spatially incoherent noise, and spatially coherent noise.

Early SWOT data have revealed the importance of model robustness under data-dropout which can arise from several sources including tidal flats, satellite maneuvers, and other roll errors. Although these data may be reprocessed in the future, presently in the Bristol Channel, for example, $\approx 10\%$ of the unsmoothed 250 m $\times$ 250 m measurement locations contain fewer than 70 measurements out of the 102 possible with the maximum number of measurements being 84 before outlier removal. Figure 3 Panel B compares the standard VBayes, OLS, and spatially coherent VBayes M2 RMS error as the number of randomly dropped measurements is increased. For this example, the spatial VBayes is only given access to a single neighboring point to capture "worst-case" performance. At least one neighbor is available in almost all cases for the SWOT data products. As before, 100 dB of WGN is added to prevent exact inversion by OLS and no nonstationary noise is added. As the number of missing measurements is increased the VBayes and OLS M2 RMS error appears to increase quadratically, with the OLS RMS error increasing by over 200% when the number of measurements available is reduced from 102 to 62 (0–40 missing data points in Figure 3). Conversely, the spatially coherent VBayes appears to increase linearly as measurements are dropped, with a relative increase of 18% using 32 data points (72 missing data points in Figure 3) from the central point. This is comparable to the OLS using more than double the number of measurements (72). In testing, the inclusion of additional neighbors further increased absolute model performance and robustness, though the improvements are less significant than those from the addition of the first point. As expected, both the spatially coherent and standard VBayes models perform marginally worse than OLS when no data dropout occurs. Under these conditions the assumptions of the OLS holds and it is thus equivalent to the maximum-likelihood estimator. It therefore outperforms any nonlinear unbiased estimator by definition (Myung, 2003). The Spatial coherence can also be implemented for the OLS. An analogous plot to Figure 3.2.1 Panel B is given in the Figure S3 of Supporting Information S1 and further confirms the benefits garnered by the Spatial coherence. The OLS is more sensitive to deviations from the linearity assumption and contamination from nonstationary noise and thus care should be taken if using the spatially coherent OLS.

Next, we evaluate the proposed method when applied to signals with both spatially incoherent noise and moderate dropout. Synthetic signals for the 3,041 nodes in the study region are contaminated with 100 dB of WGN, and are subjected to five spatially incoherent nonstationary events ranging from $0 - 1.0$ SSR. Signals also receive random dropout ranging from $0 - 50$ points leaving 77 measurements on average. These contaminants attempt to

emulate the conditions observed in early SWOT data in coastal regions which are characterized by much higher noise and erroneous measurements. The four plots in Figure 3 Panel C compare the average OLS, VBayes, and VBayes spatial differences by location and constituents over all trials (367,961 total simulations). As described in Section 2.6.1, it was found that in predominantly semi-diurnal regions it is also possible to obtain accurate estimates of the minor constituents M6 and M8. These were also included the in the analysis to test the impact of spatial coherence on minor constituent resolution. As shown previously, the standard VBayes garners a slight improvement in RSS over the OLS under moderate nonstationary contamination. Comparisons between the standard VBayes and the Spatially coherent VBayes show a significant improvement in RSS under the same conditions. The source of these improvements can be seen in the bottom two panels. The largest improvement is in the dominant M2 constituent shown in the bottom middle panel. The spatial VBayes on average demonstrates a relative improvement in M2 RMS of 44% and 40% over the OLS and standard VBayes respectively. Interestingly, estimation of other major constituents (N2, S2, and O1) remain relatively unchanged, though, critically they do not degrade. The other area of improvement produced by the spatial coherence is the resolution of the nonlinear and compound constituents (M4, M6, M8, and MN4). The difference in RSS for these constituents between the standard and spatially coherent VBayes is shown in the bottom right panel. Whereas the lowest RMS improvements for the M2 are found near the tidal turbine line in the center of the channel, the spatial VBayes demonstrates the largest improvements in minor constituent RSS in this region. Under significant noise contamination, these minor constituents can be lost beneath the noise floor. However, the spatial coherence helps to preserve the coherent information beneath the spatially incoherent noise. These improvements are significant as minor and compound constituents are critical to tidal correction accuracy in coastal regions and also drive many important coastal processes such as tidal asymmetry.

Finally, we evaluate the proposed method under spatially coherent noise. The same noise contamination utilized for the tests in Figure 3 Panel C is applied to the synthetic signals. However, this time nonstationary/white noise contamination are applied to all signals simultaneously. That is, if nonstationary noise is added to a point, it is added to all neighboring points at the corresponding time and with intensity drawn from a normal distribution given by $\mathcal{N}(SSR, SSR/10)$. Equivalent plots for the spatially coherent noise are provided in Figure A2 in the Appendix. Unsurprisingly, the spatial coherence performs worse when faced with this type of contamination, though it still yields superior performance to both the standard VBayes and OLS. As will be shown in the following section, the Bayesian pruning procedure can be used to offset these deficiencies.

3.2.2. Bayesian Pruning

In order to validate the Bayesian pruning procedure we must address two relevant questions. First, can our model correctly identify erroneous points? Second, due to the already limited amount of available Cal/Val data, does removing these datapoints improve performance? To assess these points we contaminate the data with nonstationary noise exactly as done in Section 3.1.2. Then, by applying the procedure outlined in Section 2.5, erroneous points are iteratively flagged by the model and removed. It can be seen that as outliers are pruned away, the VBayes prediction improves relative to the true signal and has greater confidence in the derived constituents (Figure 4 Panel(s) A). Pruning can be viewed as a simple binary classification problem. Hence, we assess the performance of our method using the F1-score which provides a balanced measure of precision versus recall for positive values (outliers). The formula to compute the F1-score is given in Appendix. A kernel density estimate of the distribution of F1-scores for the tests is shown in Figure 4 Panel C. The average F1-score of 0.860 and median score of 0.909 confirm the model's efficacy when classifying erroneous points. By selecting our threshold to be 3σ, we prioritize minimizing false positives. Because we are only including a subset of possible constituents in our analysis, minimizing false positives is crucial to not throwing away other tidal variability. This is reflected by the false positive rate of 1.451% and the true positive rate of 56.55%. It should be noted that the true-positive rate is likely underestimated because the synthetic surges range over multiple days, and even minor surge contributions are considered positive.

Now that we have verified that our model can accurately detect outliers we must evaluate how the derived constituents are impacted by their removal. Figure 4 Panel B compares the mean RSS of the OLS, standard VBayes, and VBayes with pruning when contaminated by 5 nonstationary surge events ranging for $0.0 \leq SSR \leq 1.0$ and WGN at 100 dB. Although the standard VBayes and OLS errors appear to increase monotonically with the added nonstationary noise, the pruned VBayes RSS error increases only marginally from the noise-free case. It can therefore be concluded that the pruning procedure is both robust and effective at removing

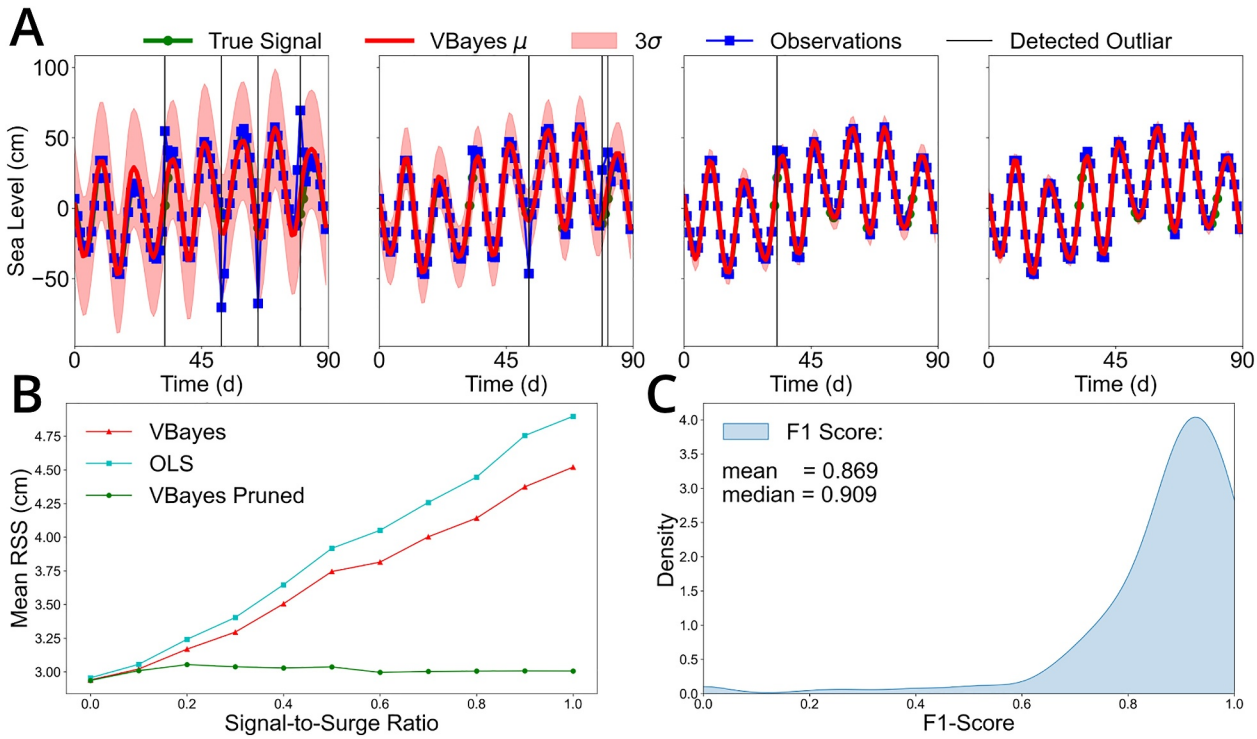

Figure 4. Evaluation of the pruning procedure outlined in Section 2.5. Panel A shows the iterative pruning procedure, with outliers being identified, removed, and the procedure repeated in each subsequent panel. Panel B compares the mean RSS error for the standard VBayes, OLS, and VBayes with pruning across M2, N2, S2, O1, M4, MN4, M6, and M8 when applied to TICON-3 synthetic signals contaminated with 100 dB WGN, SSR ratio ranging from 0 to 1.0, and number of events set to 5. A kernel density estimate of the distribution of F1-scores across all trials, defined by the binary classification of a point as being contaminated or not, is shown in Panel C.

such contaminants (e.g., surges, land contamination, roll errors). It is worth noting that and intentionally, this procedure will likely not remove systematic contaminants acting across longer time-scales, such as river outflow, and thus will not prevent biasing from these types of noise.

4. SWOT Data

Now that the VBayes methodology has been extensively validated using synthetic data, we apply our method to real SWOT data. As of writing this (May 2024), there are several versions of the SWOT Cal/Val data set of varying resolutions and processing levels. Here, we focus on the L2 2 km gridded ocean product (Processing level B) as this is intended to be the primary vehicle for scientific research and contains full geophysical corrections. We focus our analysis on the European continental shelf due to the abundance of tide gauges present within the Cal/Val orbit and the fact that two separate passes cross-over this region (003 and 016). We extract all available data from passes 003 and 016. Here all empirical analysis is conducted on the sea-level anomaly (SLA) unless explicitly stated otherwise. The SLA is derived by applying all default geophysical corrections including the CNES mean sea surface, the FES2014 tide model, and roll error correction. The validation of the methodology is carried out in two parts. First, we compare the VBayes derived harmonics directly with the 22 gauges in the TICON-3 data set. Next, we measure the associated SLA variance reduction for all locations in passes 003 and 016. All results are presented with uncertainty estimates and with examples of how these can be used to determine the confidence in a given correction.

4.1. Gauge Comparisons

Tide gauge validation is carried out on the subset of the TICON-3 gauges contained within passes 003 and 016 of the SWOT Cal/Val orbit. Because of the processing errors in the Bristol Channel, UK we do not make use of the tide gauges located in this region for validation. Gauges are selected based on several criteria. First, we locate all gauges which fall within 4 km of any SWOT KaRIn L2 2 km measurement location from passes 003 and 016. This

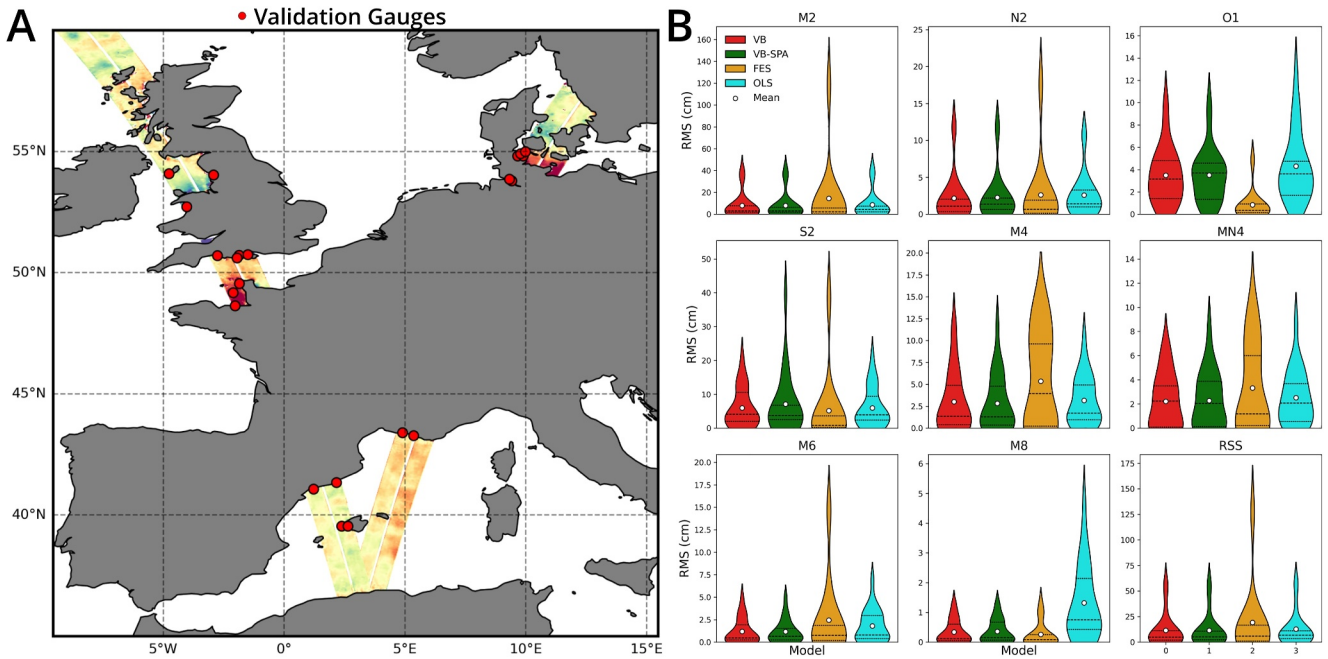

Figure 5. Tide gauge validation. Panel A shows example SSHA data for passes 003 and 016 over the European Continental shelf. The 22 TICON-3 tide gauges contained within these swaths are shown in red. Panel B compares the distribution of RMS across the 8 constituents identified in Section 2.6.1 and the total RSS between the VBayes (Red), VBayes Spatial (Green), and FES2014 (Orange) and OLS (Cyan) models. The mean values are denoted by the white circles.

yields an initial set of 45 gauges. For a location to be included we enforce that there must be at least 40 measurements present. We further restrict the set of measurements to those which have at least one neighbor in order to validate the spatial VBayes. Next, seven duplicate gauges were identified and removed to avoid over-representation from a given site. In total, 22 gauges were used for validation and are shown in Figure 5 Panel A. It is worth noting from the outset that due to the large tides in this region and the presence of several estuaries, the improvements garnered by empirical approaches will be larger than those globally.

During testing, it was found that certain constituents were more reliably estimated through residual analysis of the SLA data than direct analysis of the uncorrected data, and vice versa. We note the "residual" in residual analysis, refers to the fact that the analysis is carried out on the SLA, which is the residual signal after the applying the included tidal corrections. These differences appear to be systematic and can thus be exploited. Similar optimization has been done within the context of the inference of minor tides (M. G. Hart-Davis, Dettmering, et al., 2021). For the results shown in Figure 5 the M2 and N2 constituents are obtained via residual analysis with the remainder being obtained through direct analysis of the sea-surface height (SLA + FES Corrections). For the selected gauges it can be seen in Figure 5 Panel B that the spatial VBayes yields a 46% reduction in average M2 RMS relative to FES. This large improvement is explained by the fact that the distribution of FES RMS errors exhibits heavy tails due to poor performance in the Elbe estuary, GER. Estuaries remain an open challenge for global tidal modeling. Large and statistically significant mean improvements produced by the spatial VBayes over FES can also be found for the N2, M4, MN4, and M6 constituents with relative improvements in RMS error of 46%, 14%, 48%, and 32%. As discovered in the synthetic testing, the empirical methods struggle to resolve the S2 and O1 constituents. However, inclusion of S2 and O1 significantly improves the quality of the other derived constituents, particularly M2 and M4. These errors are nicely reflected by the VBayes uncertainty estimates, with 81% and 99% of the spatial VBayes estimates falling within the 0.95 CIs for S2 and O1, respectively. Further comparisons of the coverage of the equivalent VBayes and OLS CIs for the gauge validation are shown in Figure A4. It can be seen that OLS significantly underestimates the uncertainty of most constituents, particularly S2. The spatial VBayes marginally improves CI coverage relative to the standard VBayes. Although all methods can be seen to underestimate the uncertainty for M2 and N2, both the VBayes CIs provide a more than 30% increase in coverage relative to OLS for these constituents.

Finally, FES yields a marginal improvement in M8 resolution. However, this change is not statistically significant. With these results in mind, and in keeping with our empirical correction philosophy, our final model utilizes the spatial VBayes estimated M2, N2, M4, MN4, and M6, and combines them with the remaining FES constituents. The resulting RSS estimates are shown in the bottom right panel of Figure 5. Through supplementing the constituents derived by the spatial VBayes approach, the average FES RSS error in this region can be reduced on average by 40% with a median reduction of 12%. This improvement in median RSS indicates the spatial VBayes using Cal/Val SWOT data demonstrates systematic improvements in constituent resolution over FES in addition to just improving erroneous regions. We anticipate these results, and the number of additional constituents that can be recovered by spatial VBayes will increase as more SWOT data become available. It should also be noted that many locations contained only one neighboring point for the spatial VBayes and thus, improved performance is expected where additional measurement locations are available. Finally, we believe the statistical approach to estimate residual tidal variability presented in the next section can provide an unsupervised framework to only apply the empirical corrections where we are confident the present corrections are insufficient.

Inspection of the coastal tides derived from SWOT data, particularly using the L2 unsmoothed 250 m and pixel-cloud products, has revealed that tides exhibit rapid spatial variability in and around coastlines (M. Hart-Davis et al., 2024). This variability constitutes a non-negligible source of error when validating against tide-gauges. Classically, models are validated by either selecting the closest point to a gauge or performing linear interpolation. These approaches have been necessary but can constitute a large source of uncertainty in complex coastal regions (M. Hart-Davis et al., 2024). Our own analysis identified several gauges exhibiting significantly different behavior from the observed SWOT data. These gauges are listed in the Supporting Information S1. The identification and documentation of such gauges will be important for future tidal model validation. Quantifying the magnitude of these errors is important as they can yield biased estimates of model performance. This is particularly salient for models which assimilate gauges directly. To control for the relative error this introduces we propose a relative RMS metric ($RRMS_k$) which weights the RMS error based on the mean absolute difference between the tide gauge harmonic signal and the observed SWOT data. This takes the following form

$$\mathrm{RRMS}_k = \frac{\sqrt{\left[\left(A_{\mathrm{model}}\sin(\omega_k t) + B_{\mathrm{model}}\cos(\omega_k t)\right) - \left(A_{\mathrm{true}}\sin(\omega_k t) + B_{\mathrm{true}}\cos(\omega_k t)\right)\right]^2}}{\sum_{t=0}^{N}\left|y_{obs}(t) - y_{gauge}(t)\right|} \tag{22}$$

where the numerator is the conventional constituent RMS given in Stammer et al. (2014), and the denominator is the mean absolute error (MAE) between the SWOT measurements and the tide gauge harmonic prediction. We favor MAE over mean squared error (MSE), as MSE places greater weight on outliers. This metric quantifies the error of the tide model relative to the "uncertainty" introduced between gauges and adjacent measurements. Figure A3 in the Appendix presents the same analysis as Figure 5, but uses the RRMS metric. Although the spatial VBayes continues to yield superior mean RRMS for M2, N2, M4, and M6, the shape of the RRMS distributions changes significantly. The RRMS distributions for the empirical models exhibit significantly less kurtosis relative to the RMS. Because the RRMS down-weights errors associated with larger differences between the SWOT and gauge measurements, this reduction in kurtosis for the empirical models indicates that many of the erroneous RMS values can actually be attributed to discrepancies between the gauge location and the SWOT measurement location. By accounting for these discrepancies, we can conclude that any remaining large errors in the RRMS values are due to systematic modeling errors in whichever model is being assessed. This is certainly the case for the dominant FES outliers in the Elbe estuary which remain when using RRMS. Although we do not believe this metric should replace the standard RMS, it can add an additional dimension with which to understand the systematic errors intrinsic to the tide gauge validation of global tide models.

4.2. SLA Variance Reduction

Although in situ tide gauge validation allows for direct accuracy assessment of the derived constituents, point estimates cannot effectively quantify the spatial characteristics of model performance. Estimation of such errors is particularly salient in complex coastal regions were SWOT data have revealed standard linear interpolation to be deficient. Furthermore, several of the gauges employed for the previous analysis were assimilated into FES2014b and can thus yield biased estimates of global accuracy (Stammer et al., 2014). We look to approximate the spatial error characteristics of FES by applying the proposed methodology to correct the residual SLA. We estimate the

SLA variance difference according to $\mathrm{Var}_{diff} = 100 * \left(\frac{\mathrm{Var}_{FES} - \mathrm{Var}_{model} - N\sigma_{model}}{\mathrm{Var}_{FES}}\right)$ where N is the number of standard deviations of uncertainty to include. Note, OLS estimates do not include uncertainty. Inspection of the SLA data reveals several large errors likely due to satellite maneuvers and roll errors which lead to high SLA variance estimates. These errors are removed by filtering the initial SWOT data for values exceeding 3σ beyond the sample mean. The three methods are then applied to the filtered SLA with the results shown in Figure 6.

Panels A and B visualize the percentage SLA variance reduction in the Bristol Channel, UK, using the spatial VBayes under different uncertainty thresholds. Due to the massive tidal range and severe nonlinearity in this region global tide models typically perform poorly (M. Hart-Davis et al., 2024). A visualization of an SLA signal in this region can be seen in the second column in the bottom row of Panel C. Clearly, considerable tidal variability remains, as is evidenced by the reduced SLA variance after applying the OLS and the standard and spatial VBayes. It can also be seen that a majority of this variability exists beyond the $\pm 2\sigma$ confidence intervals of the spatial VBayes. Thus, we can say with 95% confidence that the original SLA which sits beyond this region can be attributed to residual tidal variability. Higher up in the Severn estuary it can be seen that the model is less confident. This is a combination of increased nonlinearity, river outflow, interactions with large sand deposits, and sensor dropouts which lead to large increases in aleatoric uncertainty (e.g., uncertainty from the noise intrinsic to the data itself). Even for confidence levels of 68% and 95% respectively, the spatial VBayes still yields median SLA reductions of 51.4% and 27.1% respectively (Figure 6 Panel B). Hence, the inclusion of the spatial VBayes corrections to these data can significantly improve the quality of the derived SLA for other studies. At the time of writing this, issues with the processing of the new SWOT low-rate, version B, data products have resulted in large gaps in coverage in the Bristol Channel. As such, the results shown in the Bristol Channel study makes use of the version A data-products and associated processing. Although our testing showed little to no change in SLA variance reduction between the A and B data products for the portion of the channel which was present in both data products, we stress that caution should be taken when interpreting these results.

Figure 6 panel(s) C provides comparisons of SLA variance reduction including the standard VBayes and OLS when used to correct data from pass 003 of the Cal/Val orbit in and around the United Kingdom. It can be seen that the OLS provides the greatest reduction in SLA when considered without uncertainty. This fact is unsurprising, as the estimator is simply finding parameter values which maximally describe the data set. Inspection of the standard VBayes and spatial VBayes as the confidence level is increased tells a different story. Although considerable tidal variability can be found in the SLA in and around the complex UK coastline, residual tides either do not exist or, more likely, cannot be determined from the available data in regions shown in blue. An example of this is shown in the bottom right panel ($50.222° N 1.022° W$). Without considering uncertainty, both OLS and VBayes tidal estimation suggest the observed SLA variance can be reduced by around 40%. Although some tidal variability may remain, it can also be seen that both estimators are biased by the nonstationary and clearly nontidal SLA spikes which were not filtered out. All values in the original SLA can be seen to fall within the spatial VBayes 2σ confidence and thus, we cannot conclude at the 95% confidence level that the signal contains any tidal variability. We believe this interpretation provides a simple and statistically grounded approach to estimate the residual tidal variability in the corrected SLA products. Using this approach, Table A1 estimates the residual tidal variability for the L2 2 km Cal/Val SLA in passes 003 and 016 between 50 and 60° latitude. Using this approach we estimate with a confidence level of 95%, on average 7.2% of the SLA variance in this region is attributable to tidal variability. The slight increase in estimated tidal variance between standard VBayes and spatial VBayes is due to the fact that spatial coherence yields a reduction in epistemic uncertainty (e.g., uncertainty which arises from insufficient information). This makes sense conceptually as when the information contained in adjacent points is congruous with the central point (spatially coherent), our model has greater confidence in its prediction. The 7.2% estimate should be viewed as a conservative lower bound as we are only fitting 8 harmonics and are largely neglecting diurnal tides. Regardless, we believe our approach when combined with the high-resolution KaRIn data is useful in estimating the spatial deficiencies in global tidal models.

For researchers employing the corrections directly, the final choice of confidence level can be selected depending on the nature of their usage. It is expected that additional data from the SWOT science phase will reduce the epistemic uncertainty and help estimate some of the smaller residual tidal signals. Errors in SWOT data processing can yield artifacts in both the estimated tides and the SLA variance reduction estimates. In particular, we see that large SLA variance reductions can be seen in and around the nadir gap. Reviewer comments have suggested this is a consequence of incorrect remapping from the SWOT measurement grid to the fixed reference

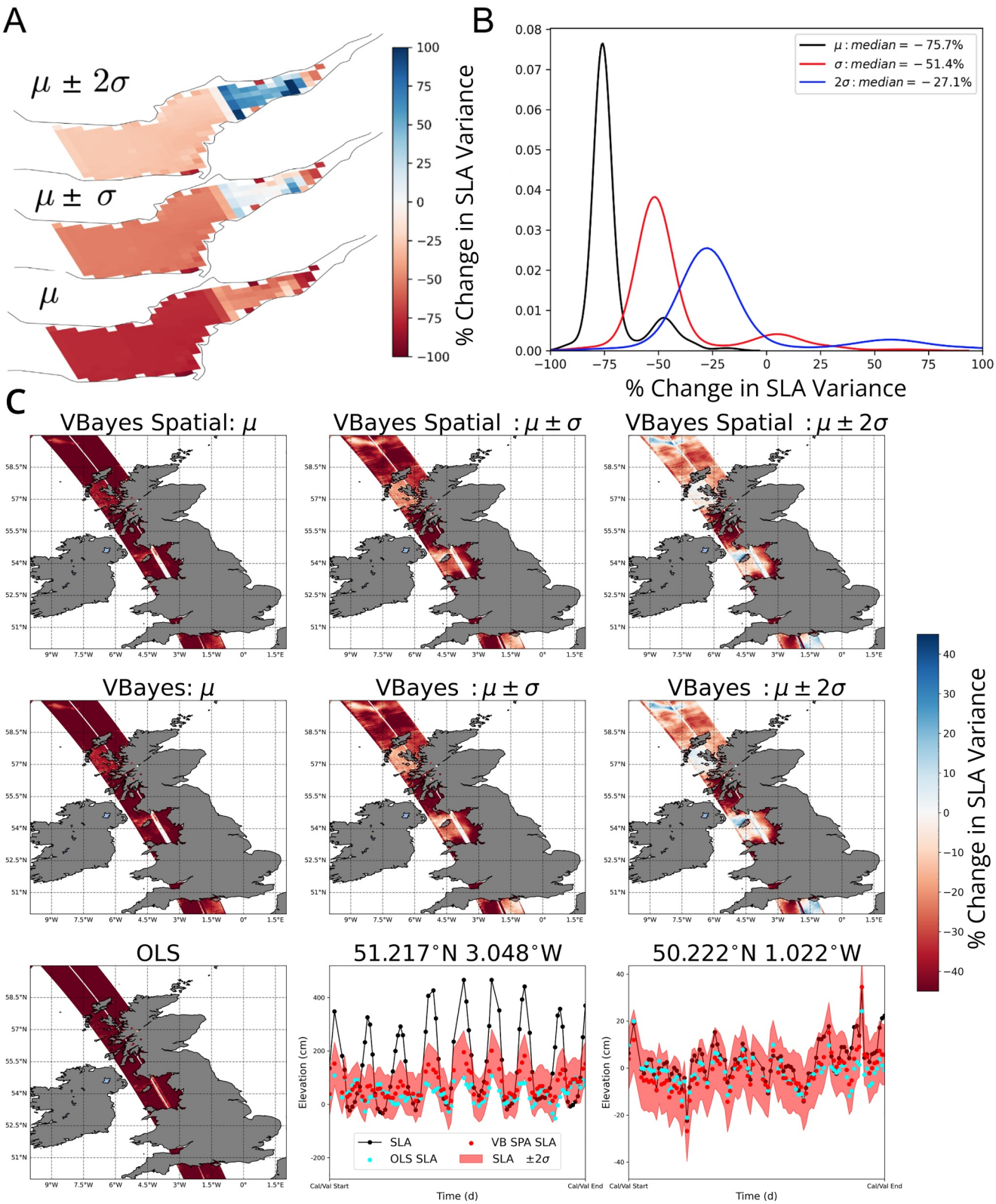

Figure 6.

grid in the data products. Irrespective, we find that the pruning procedure provided seems to effectively reduce these errors in the corrected SLA.

5. Discussion

Arguably, as important as correcting errors from present geophysical corrections is estimating where and how they are deficient. Despite their remarkable performance and reliability, many global tide models, FES included, do not provide uncertainty estimates (M. G. Hart-Davis, Piccioni, et al., 2021). The VBayes residual analysis of SWOT data provides a potential solution to this limitation by approximating these uncertainties directly, including those created by spatial interpolation. Critically, this provides a unique high-resolution view into both the short-scale spatial characteristics of tides and, by extension, the errors made by tide models. As demonstrated, equivalent estimates cannot accurately be obtained using conventional OLS harmonic analysis.

One important feature of our method is the novel incorporation of spatial coherence outlined in Section 2.4. This approach formalizes a long standing practice of binning altimetric measurements for empirical tidal models (Piccioni et al., 2021). Although our approach also weights measurements according to their distance to the central node we provide several further developments. First, we add an additional weighting based on the probability that the observations at the given point are correlated with those of the central point. This is particularly important in complex coastlines where greater spatial variability is present. Second, rather than assume the quadrature components of the central node describe all neighbors within the given radius, we instead allow a linear variation between these nodes. This new approach shows significant improvement for simulated and independently derived SWOT data. The inductive bias effectively increases the ratio of measurements to parameters by a factor of two when using just four neighbors and can scale with additional points. We have shown that this procedure can help reduce aliasing errors and improve the resolution of minor tides. Future work is needed to determine the breakdown point of the linearity assumption when the radius of neighboring points is increased beyond 4 km. Based on this, the spatial coherence can be applied to ungridded data from historical altimetry missions.

We wish to stress that no model is a silver bullet, and the spatial VBayes is not without its limitations. In low noise regimes, the OLS is by very definition, a superior estimator. Furthermore, assumptions made by the variational Bayesian approach and the epistemic uncertainty from the limited measurements can lead to underestimation of the true uncertainty. The construction of our methodology has made every attempt to mitigate these limitations. We advise that uncertainty estimates should be viewed as a lower bound on the true parameter uncertainty. As such, the final choice of confidence level is left to the end user to allow researchers to decide how confident they wish to be depending on their application. It is also worth remarking that for a majority of regions, assimilative tidal models such as FES2014b perform phenomenally well (Stammer et al., 2014). In these regions empirical correction of the SLA can actually introduce error into the signal. The procedure outlined in Section 4.2 helps to mitigate this by allowing a user defined threshold of residual tidal variability in order to apply corrections. For example, by simply requiring that a location's SLA contain at least 5% tidal variance at a 95% confidence level, testing with the gauges showed many of these errors can be avoided.

It is worth noting that the VBayes method does not explicitly consider temporal correlations in the residual, though it provides increased robustness to these errors compared to OLS. Because of the background red-noise intrinsic to the noise spectrum this can lead to inaccurate parameter estimation. Although several approaches were considered to account for this variability such as considering a red-noise covariance (Kachelein, 2023), the sparsity of the SWOT data significantly degraded the effectiveness of approaches which rely on spectral estimation of the residual. This represents an important area of future work when applying VBayes to more conventionally sampled reference series. In this vein, the VBayes estimator can be modified in several ways. Perhaps the easiest is to treat the residual as a mixture of Gaussians, which can overcome the assumption of

Figure 6. Comparisons of SLA variance change after applying empirical corrections under different confidence levels. Panel A shows the spatial VBayes percent variance change in the Bristol Channel, UK when considered with uncertainty. Panel B shows the associated kernel density estimates of the distributions of these changes. Panel C compares the spatial VBayes, standard VBayes, and OLS when used to perform residual analysis on data from pass 003 around the UK at different confidence levels. The bottom two panels show representative examples of the tides from the Bristol Channel, UK, and the English Channel, respectively. Blue, increase in variance; Red, reduction in variance.

homoskedastic residuals (S. J. Roberts & Penny, 2002). Additionally, to our knowledge, regularization has not yet been formally considered within the context of frequency resolution in tidal analysis. Although several authors have experimented with simple ridge regression models (Li et al., 2022; Pan et al., 2024), theoretical work is needed to formalize these improvements relative to the standard Rayleigh and Munk-Hasselman criterions and how it relates to the characteristic failure mode of unregularized least-squares as record length is decreased.

Future work will look to apply the spatial VBayes approach to the more than 30 years of historical altimetry measurements now available. As demonstrated in both the simulated and real validation, the reduced sensitivity to noise, outlier removal, and uncertainty estimation are useful properties within the context of SLA residual analysis. As the science phase of the SWOT mission nears completion of its first year, sufficient data will exist to produce analogous analyses at a global scale. With the approximately 21 days repeat orbit of this phase, the advantages garnered by the spatial VBayes will prove even more salient. Additionally, the VBayes estimator has shown very interesting properties regarding tidal super-resolution and overcoming "aliasing", as illustrated by the example in Figure 1, and the improved resolution of other constituents for the sun synchronous Envisat orbit in the Supporting Information S1. To be explicit, no-degree of super-resolution can overcome the aliasing of the S2 constituent for sun-synchronous missions. However, as is shown, unregularized estimators (OLS) can exhibit bias across other parameters if the aliased parameter is included in the analysis. Hence, the ARD priors of the VBayes help prevent the errors from S2 estimation from degrading the estimation of other constituents. Future work will look to formalize this result as it could be useful in maximizing the scientific return from historical sun synchronous missions.

Finally, it remains to be said how the SWOT data themselves can actually be used to further our scientific understanding of tides. Early analysis of the high-resolution data products in the Bristol Channel has revealed remarkable spatial characteristics of compound and nonlinear tides. This presents the opportunity to study features such as tidal asymmetry, and tidal interactions with river outflow. The unique 1-day repeat orbit of the SWOT Cal/Val orbit also allows for the study of relatively short-scale temporal processes. Future work will look to study the temporal and spatial characteristics of storm surges using these data. These applications are only scratching the surface of what is possible using wide-swath altimetry. However, we note that the success of these applications remains dependent on the quality of the tidal corrections themselves. It is our hope that the spatial VBayes approach can help to enable this success.

6. Conclusion

This work develops a spatially coherent harmonic analysis using variational Bayesian inference rather than conventional least-squares estimation. The method demonstrates superior robustness to several types of noise contamination common in altimetric tidal analysis. Several, techniques are developed using the Bayesian uncertainty estimation to automatically identify and remove erroneous measurements. The method demonstrates excellent performance relative to state-of-the-art models such as FES2014b when validated on tide gauges on the European shelf. The largest improvements can be found in the Elbe Estuary, GER. Further analysis of SLA reduction confirms these results and provides a first look at swath based uncertainty estimation for FES2014b using the SWOT Cal/Val data. We estimate a lower bound on the residual tidal variability in the corrected SLA with a confidence level of 95% of at least 7.2% on average for Cal/Val passes 003 and 016. Additionally, estimates of the MSS error highlight large uncertainties around complex coastlines arising from tidal flats, sediment deposits, and propagated errors from tidal corrections exceeding 1 m in regions such as the Bristol Channel, UK.

SWOT Cal/Val tidal estimates will be produced using the spatial VBayes methodology. These data sets will serve both as a source of corrections for early SWOT data, and as a unique and independent estimate of the spatial characteristics of the uncertainty in present geophysical corrections. An implementation of the variational Bayesian tidal analysis is provided as an open-source Python package (VTide) and can be employed for conventional tidal analysis and prediction (https://github.com/thomasmonahan/VTide).

Appendix A

A1. Metrics

Outlined below are the formulas for the metrics used for model validation.

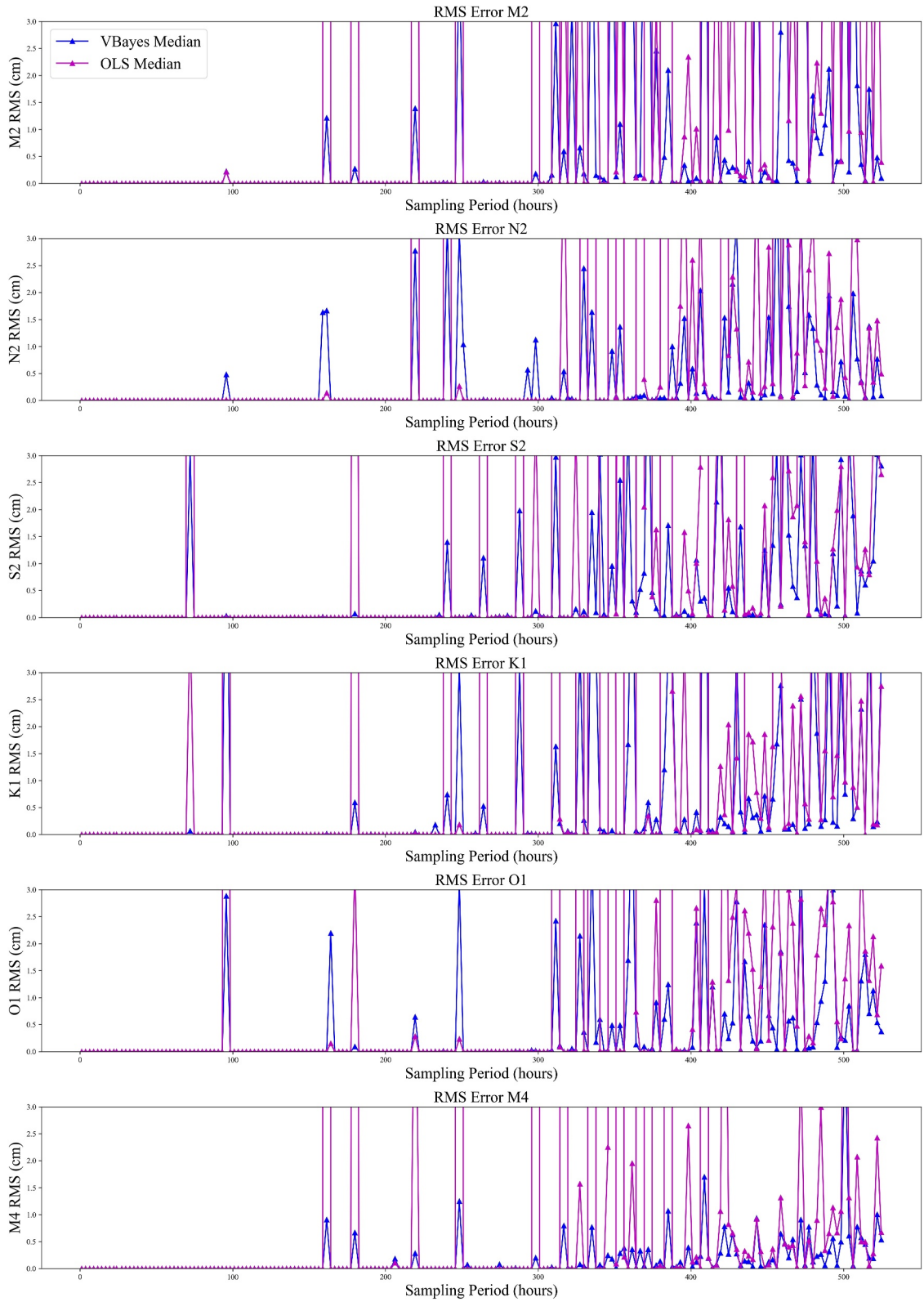

Figure A1. Comparisons of RMS Error for OLS and VBayes at different sampling rates. Median RMS error is computed for the entire TICON-3 data set based on a 3-year reference series spanning the SWOT Science Orbit sampled at the given sampling rate.

$$\mathrm{RMS}_k = \sqrt{\overline{[(A_{\mathrm{model}} \sin(\omega_k t) + B_{\mathrm{model}} \cos(\omega_k t)) - (A_{\mathrm{true}} \sin(\omega_k t) + B_{\mathrm{true}} \cos(\omega_k t))]^2}} \tag{A1}$$

$$\mathrm{RSS} = \sqrt{\sum_{k=1}^{n} (\mathrm{RMS}_k)^2} \tag{A2}$$

$$\mathrm{Var}_{diff} = 100 \times \left(\frac{\mathrm{Var}_{FES} - \mathrm{Var}_{model} - N\sigma_{model}}{\mathrm{Var}_{FES}} \right) \tag{A3}$$

$$\text{F1-Score} = \frac{\mathrm{TP}}{\mathrm{TP} + 0.5(\mathrm{FP} + \mathrm{FN})} \tag{A4}$$

$$\mathrm{SNR\ (dB)} = 10 \cdot \log_{10} \left(\frac{\sigma^2_{\mathrm{signal}}}{\sigma^2_{\mathrm{noise}}} \right) \tag{A5}$$

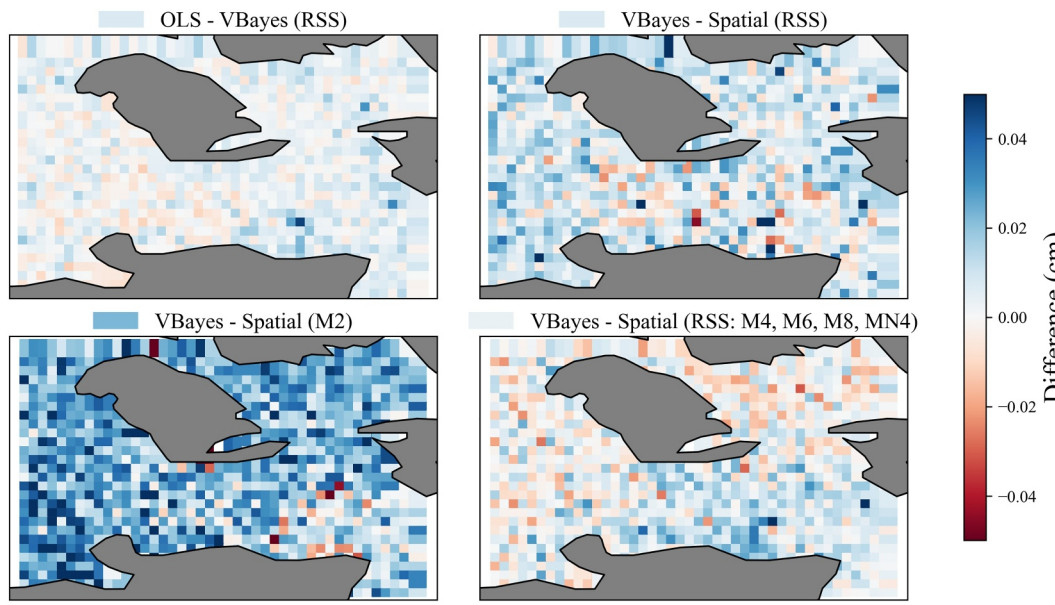

Figure A2. Comparisons of model performance under spatially coherent noise. Panels compare the average difference in RMS error between the given models when contaminated with spatially incoherent noise (Red, first model better; white, same; blue, second model better). The average difference is given next to the title of each plot.

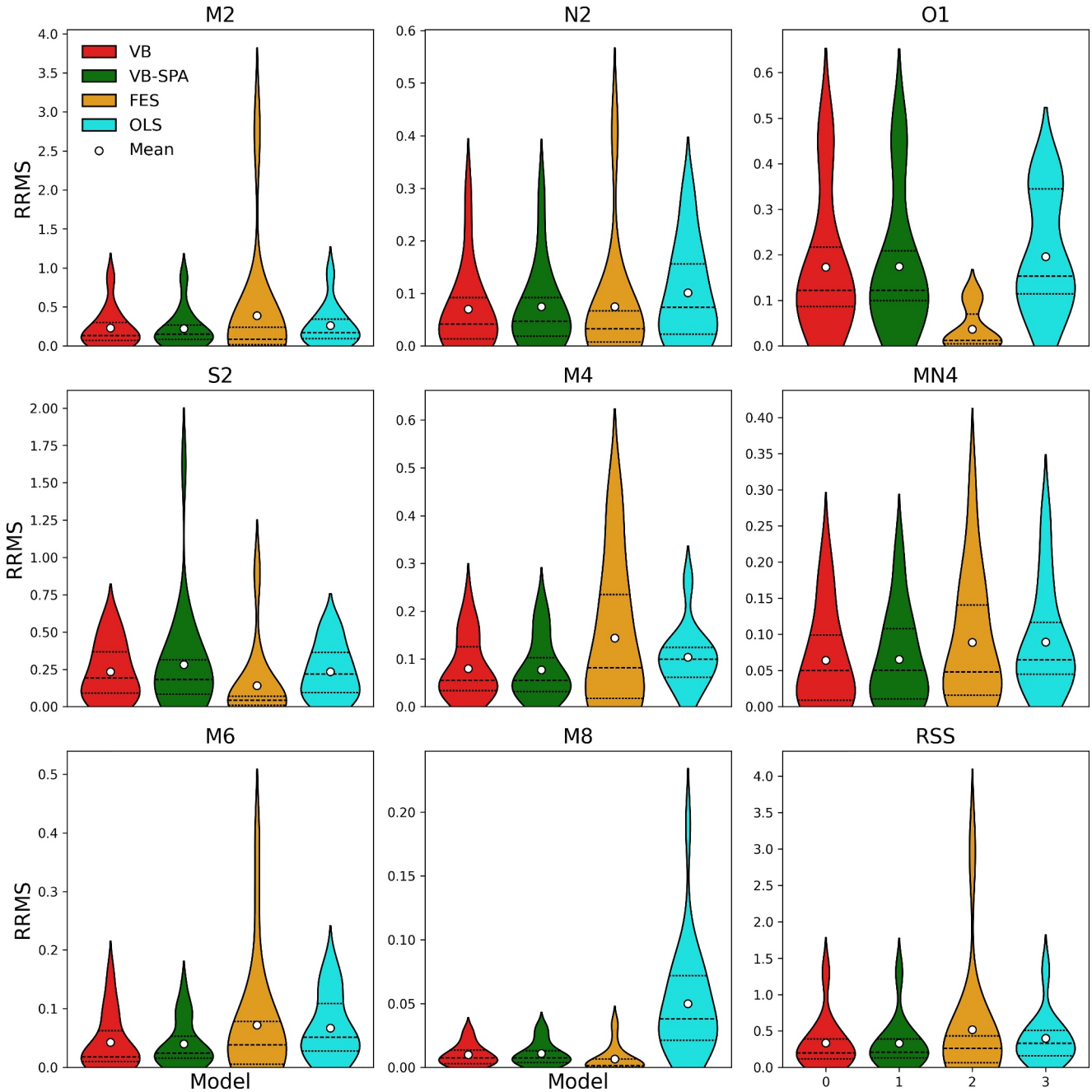

Figure A3. Comparisons of model relative RMS (RRMS) error for tide gauge validation.

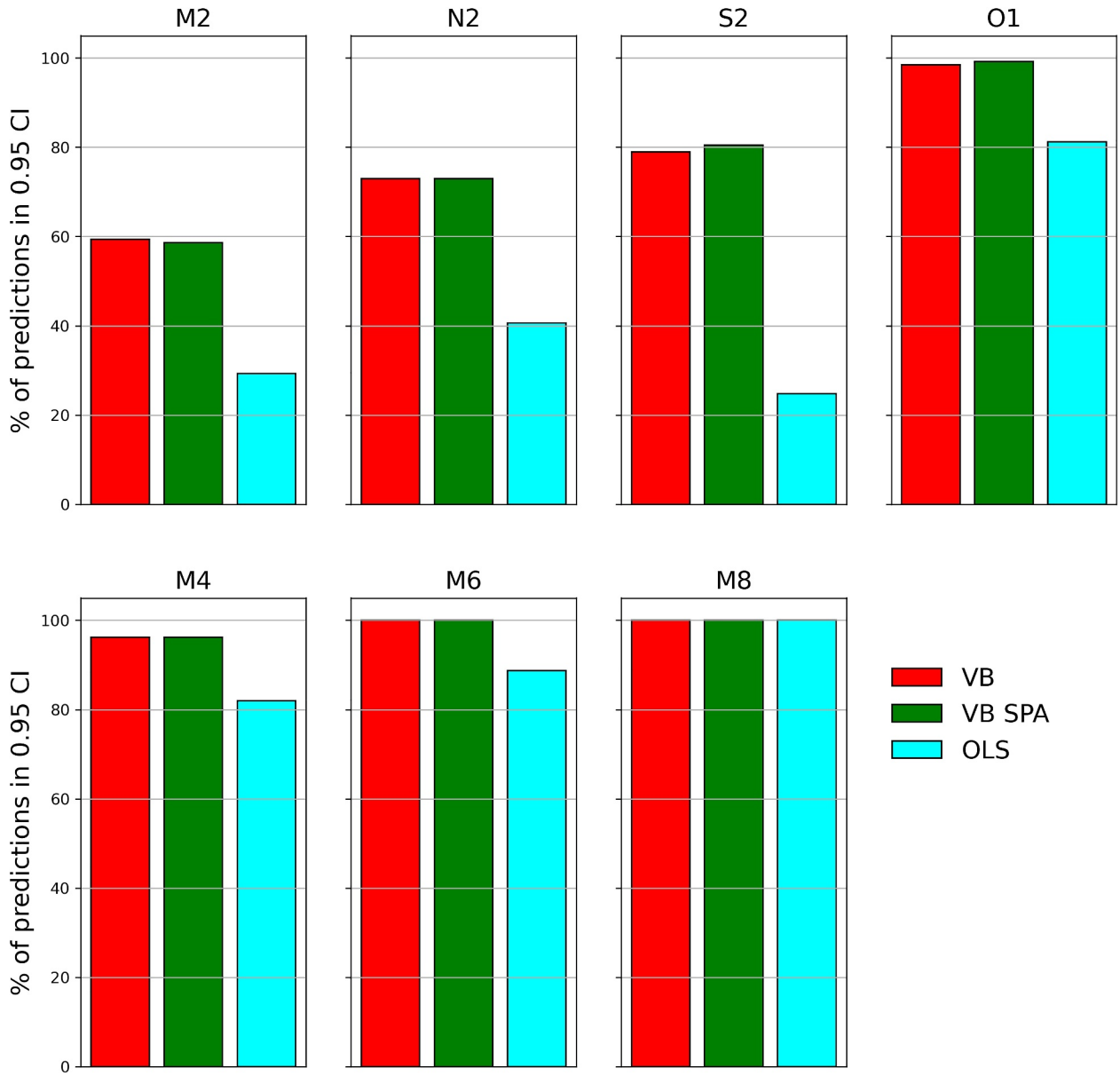

Figure A4. Comparisons of VBayes and OLS uncertainty estimates for Gauge Validation. The percentage of predictions which fall within the computed 0.95 Confidence Intervals (CI) are shown. A total of 133 SWOT measurements were used (all available measurements within 4 km of the gauges used).

Table A1
Average Estimate of Residual Tidal Variance for SWOT Cal/Val Orbit Passes 003 and 016 Under Different Confidence Levels. $\mu \pm \sigma = 68\%$, $\mu \pm 2\sigma = 95\%$

Method	μ	$\mu \pm \sigma$	$\mu \pm 2\sigma$
VBayes Spatial (%)	43.3	25.2	7.2
VBayes (%)	42.7	24.5	6.4
OLS	53.4%	–	–

Data Availability Statement

All low-rate sea surface height SWOT data are made publicly accessible via https://doi.org/10.24400/527896/A01-2023.015. The data used for validation in Sections 4.1 and 4.2 make use of the pre-validated L2_LR_SSH version 2.0 (version C). All results for the Bristol Channel, UK make use of the pre-validated L2_LR_SSH version 1.0 (version B) products. The FES2014b atlas used for tide-gauge validation can be accessed through Aviso (Lyard et al., 2021). We note that we utilized the "extrapolated" version which yields superior performance in coastal regions. The TICON-3 data set used for synthetic validation is described in (M. Hart-Davis et al., 2022) and can be accessed at: https://doi.org/10.1594/PANGAEA.951610. All code required to reproduce each analysis is available at https://doi.org/10.5281/zenodo.14892032. The code is also maintained as the VTide Python package: github.com/thomasmonahan/vtide.

Acknowledgments
TT acknowledges support from the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Futures Program. We would like to thank both anonymous reviewers for their helpful feedback. TM would like to thank Michael Hart-Davis for several helpful discussions and assistance validating the proposed method.

References

Adcock, T. A. A., Draper, S., Houlsby, G. T., Borthwick, A. G. L., & Serhadlıoğlu, S. (2013). The available power from tidal stream turbines in the Pentland Firth. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *469*(2157), 20130072. https://doi.org/10.1098/rspa.2013.0072

Andersen, O. B., Rose, S. K., & Hart-Davis, M. G. (2023). Polar ocean tides—Revisited using cryosat-2. *Remote Sensing*, *15*(18), 4479. https://doi.org/10.3390/rs15184479

Anselin, L. (2013). *Spatial econometrics: Methods and models* (Vol. 4). Springer Science and Business Media.

Arbic, B. K., Lyard, F., Ponte, A., Ray, R. D., Richman, J. G., Shriver, J. F., et al. (2015). Tides and the SWOT mission: Transition from science definition team to science team.

Attias, H. (1999). A variational baysian framework for graphical models. *Advances in Neural Information Processing Systems*, *12*.

Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning. *Springer*, *4*(4).

Cartwright, D. E. (1968). A unified analysis of tides and surges round north and east Britain. *Philosophical Transactions of the Royal Society of London—Series A: Mathematical and Physical Sciences*, *263*(1134), 1–55.

Caticha, A., & Preuss, R. (2004). Maximum entropy and Bayesian data analysis: Entropic prior distributions. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *70*(4), 046127. https://doi.org/10.1103/physreve.70.046127

Codiga, D. L. (2011). Unified tidal analysis and prediction using the UTide Matlab functions.

Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, *39*(1), 1–38. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Donoho, D. L., & Huber, P. J. (1983). The notion of breakdown point. *A festschrift for Erich L. Lehmann*, 157184.

Fox, C. W., & Roberts, S. J. (2012). A tutorial on variational bayesian inference. *Artificial Intelligence Review*, *38*(2), 85–95. https://doi.org/10.1007/s10462-011-9236-8

Hart-Davis, M., Andersen, O., Ray, R., Zaron, E., Schwatke, C., Arildsen, R., et al. (2024). Tides in complex coastal regions: Early case studies from wide-swath SWOT measurements. *Geophysical Research Letters*, *51*(20), e2024GL109983. https://doi.org/10.1029/2024gl109983

Hart-Davis, M., Dettmering, D., & Seitz, F. (2022). *TICON-3: Tidal constants based on GESLA-3 sea-level records from globally distributed tide gauges including gauge type information (data)*. Deutsches Geodätisches Forschungsinstitut.

Hart-Davis, M. G., Dettmering, D., Sulzbach, R., Thomas, M., Schwatke, C., & Seitz, F. (2021a). Regional evaluation of minor tidal constituents for improved estimation of ocean tides. *Remote Sensing*, *13*(16), 3310. https://doi.org/10.3390/rs13163310

Hart-Davis, M. G., Piccioni, G., Dettmering, D., Schwatke, C., Passaro, M., & Seitz, F. (2021b). EOT20: A global ocean tide model from multi-mission satellite altimetry. *Earth System Science Data*, *13*(8), 3869–3884. https://doi.org/10.5194/essd-13-3869-2021

Holland, P. W., & Welsch, R. E. (1977). Robust regression using iteratively reweighted least-squares. *Communications in Statistics—Theory and Methods*, *6*(9), 813–827. https://doi.org/10.1080/03610927708827533

Innocenti, S., Matte, P., Fortin, V., & Bernier, N. (2022). Analytical and residual bootstrap methods for parameter uncertainty assessment in tidal analysis with temporally correlated noise. *Journal of Atmospheric and Oceanic Technology*, *39*(10), 1457–1481. https://doi.org/10.1175/jtech-d-21-0060.1

Kachelein, L. (2023). *Bayesian harmonic analysis of tidal and wind-driven currents in the California current system*. University of California.

Landau, H. (1967). Sampling, data transmission, and the Nyquist rate. *Proceedings of the IEEE*, *55*(10), 1701–1706. https://doi.org/10.1109/proc.1967.5962

Leffler, K. E., & Jay, D. A. (2009). Enhancing tidal harmonic analysis: Robust (hybrid L1/L2) solutions. *Continental Shelf Research*, *29*(1), 78–88. https://doi.org/10.1016/j.csr.2008.04.011

Le Provost, C. (2001). Ocean tides. In *International geophysics* (Vol. 69, pp. 267–303). Elsevier. https://doi.org/10.1016/s0074-6142(01)80151-0

Li, Y., Peng, G., Chen, P., Chen, K., Li, R., & Song, Z. (2022). Harmonic analysis of short-term tidal level prediction model for tidal reaches. *Arabian Journal of Geosciences*, *15*(6), 473. https://doi.org/10.1007/s12517-022-09757-1

Lichtman, D., Banks, C., Calafat, F., Gommenginger, C., & Bell, P. (2023). Towards validation of SWOT in the coastal zone: A radar altimetry and water level gauge case study in the Bristol Channel and severn river-estuary system. In *2023 ocean surface topography science team meeting*. 123.

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. https://doi.org/10.5194/os-17-615-2021

Matte, P., Bernier, N., Fiset, J.-M., Fortin, V., & Secretan, Y. (2018). Constrained harmonic analysis: A new method for mapping estuarine tides based on in-situ and remotely-sensed data. In *Poster presented at the 2018 ocean sciences meeting* (pp. 11–16).

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. https://doi.org/10.1175/jtech-d-12-00016.1

Mishali, M., & Eldar, Y. C. (2010). From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals. *IEEE Journal of selected topics in signal processing*, *4*(2), 375–391. https://doi.org/10.1109/jstsp.2010.2042414

Morrow, R., Fu, L.-L., Ardhuin, F., Benkiran, M., Chapron, B., Cosme, E., et al. (2019). Global observations of fine-scale ocean surface topography with the Surface Water and Ocean Topography (SWOT) mission. *Frontiers in Marine Science*, *6*, 232. https://doi.org/10.3389/fmars.2019.00232

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

Munk, W., & Hasselmann, K. (1964). Super-resolution of tides. *Studies on oceanography*, 339–344.

Munk, W., Zetler, B., & Groves, G. (1965). Tidal cusps. *Geophysical Journal International*, *10*(2), 211–219. https://doi.org/10.1111/j.1365-246x.1965.tb03062.x

Munk, W. H., & Cartwright, D. E. (1966). Tidal spectroscopy and prediction. *Philosophical Transactions of the Royal Society of London—Series A: Mathematical and Physical Sciences*, *259*(1105), 533–581.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*(1), 90–100. https://doi.org/10.1016/s0022-2496(02)00028-7

Pan, H., Xu, T., & Wei, Z. (2024). Improved tidal estimates from short water level records via the modified harmonic analysis model. *Ocean Modelling*, *189*, 102372. https://doi.org/10.1016/j.ocemod.2024.102372

Parker, B. B. (2007). Tidal analysis and prediction.

Pawlowicz, R., Beardsley, B., & Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Computers & Geosciences*, *28*(8), 929–937. https://doi.org/10.1016/s0098-3004(02)00013-4

Penny, W., & Roberts, S. (2002). Bayesian multivariate autoregressive models with structured priors. *IEE Proceedings—Vision, Image and Signal Processing*, *149*(1), 33–41. https://doi.org/10.1049/ip-vis:20020149

Piccioni, G., Dettmering, D., Schwatke, C., Passaro, M., & Seitz, F. (2021). Design and regional assessment of an empirical tidal model based on FES2014 and coastal altimetry. *Advances in Space Research*, *68*(2), 1013–1022. https://doi.org/10.1016/j.asr.2019.08.030

Ray, R., Egbert, G., & Kelly, S. (2020). Barotropic and baroclinic tide models for and from SWOT.

Roberts, S., McQuillan, A., Reece, S., & Aigrain, S. (2013). Astrophysically robust systematics removal using variational inference: Application to the first month of Kepler data. *Monthly Notices of the Royal Astronomical Society*, *435*(4), 3639–3653. https://doi.org/10.1093/mnras/stt1555

Roberts, S. J., & Penny, W. D. (2002). Variational Bayes for generalized autoregressive models. *IEEE Transactions on Signal Processing*, *50*(9), 2245–2257. https://doi.org/10.1109/tsp.2002.801921

Rousseeuw, P., & Leroy, A. (1987). *Robust regression and outlier detection*. John Wiley& Sons.

Ruanaidh, J. J. O., & Fitzgerald, W. J. (2012). *Numerical Bayesian methods applied to signal processing*. Springer Science and Business Media.

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, *2*, e55. https://doi.org/10.7717/peerj-cs.55

Schaeffer, P., Pujol, M.-I., Veillard, P., Faugere, Y., Dagneaux, Q., Dibarboure, G., & Picot, N. (2023). The CNES CLS 2022 mean sea surface: Short wavelength improvements from CryoSat-2 and SARAL/Altika high-sampled altimeter data. *Remote Sensing*, *15*(11), 2910. https://doi.org/10.3390/rs15112910

Schrama, E., & Ray, R. (1994). A preliminary tidal analysis of TOPEX/POSEIDON altimetry. *Journal of Geophysical Research*, *99*(C12), 24799–24808. https://doi.org/10.1029/94jc01432

Stammer, D., Ray, R., Andersen, O. B., Arbic, B., Bosch, W., Carrère, L., et al. (2014). Accuracy assessment of global barotropic ocean tide models. *Reviews of Geophysics*, *52*(3), 243–282. https://doi.org/10.1002/2014rg000450

Thomson, R. E., & Emery, W. J. (2014). *Data analysis methods in physical oceanography*. Newnes.

Wang, Y., Kucukelbir, A., & Blei, D. M. (2017). Robust probabilistic modeling with bayesian data reweighting. In *International conference on machine learning* (pp. 3646–3655).

Williams, J., & Horsburgh, K. (2013). *Evaluation and comparison of the operational Bristol Channel model storm surge suite (research and consultancy Report No. 38)*. National Oceanography Centre.

Williams, J., Irazoqui Apecechea, M., Saulter, A., & Horsburgh, K. J. (2018). Radiational tides: Their double-counting in storm surge forecasts and contribution to the highest astronomical tide. *Ocean Science*, *14*(5), 1057–1068. https://doi.org/10.5194/os-14-1057-2018

Wunsch, C. (2006). *Discrete inverse and state estimation problems: With geophysical fluid applications*. Cambridge University Press.