Analytical and Residual Bootstrap Methods for Parameter Uncertainty Assessment in Tidal Analysis with Temporally Correlated Noise

SILVIA INNOCENTI, a PASCAL MATTE, a VINCENT FORTIN, a and NATACHA BERNIER a

Meteorological Research Division, Environment and Climate Change Canada, Montreal, Quebec, Canada

(Manuscript received 11 May 2021, in final form 15 July 2022)

ABSTRACT: Reconstructing tidal signals is indispensable for verifying altimetry products, forecasting water levels, and evaluating long-term trends. Uncertainties in the estimated tidal parameters must be carefully assessed to adequately select the relevant tidal constituents and evaluate the accuracy of the reconstructed water levels. Customary harmonic analysis uses ordinary least squares (OLS) regressions for their simplicity. However, the OLS may lead to incorrect estimations of the regression coefficient uncertainty due to the neglect of the residual autocorrelation. This study introduces two residual resamplings (moving-block and semiparametric bootstraps) for estimating the variability of tidal regression parameters and shows that they are powerful methods to assess the effects of regression errors with nontrivial autocorrelation structures. A Monte Carlo experiment compares their performance to four analytical procedures selected from those provided by the RT_Tide, UTide, and NS_Tide packages and the robustfit.m MATLAB function. In the Monte Carlo experiment, an iteratively reweighted least squares (IRLS) regression is used to estimate the tidal parameters for hourly simulations of one-dimensional water levels. Generally, robustfit.m and the considered RT_Tide method overestimate the tidal amplitude variability, while the selected UTide and NS_Tide approaches underestimate it. After some substantial methodological corrections the selected NS_Tide method shows adequate performance. As a result, estimating the regression variance-covariance with the considered RT_Tide, UTide, and NS_Tide methods may lead to the erroneous selection of constituents and underestimation of water level uncertainty, compromising the validity of their results in some applications.

SIGNIFICANCE STATEMENT: At many locations, the production of reliable water level predictions for marine navigation, emergency response, and adaptation to extreme weather relies on the precise modeling of tides. However, the complicated interaction between tides, weather, and other climatological processes may generate large uncertainties in tidal predictions. In this study, we investigate how different statistical methods may lead to different quantifications of tidal model uncertainty when using data with completely known properties (e.g., knowing the tidal signal, as well as the amount and structure of noise). The main finding is that most commonly used statistical methods may estimate incorrectly the uncertainty in tidal parameters and predictions. This inconsistency is due to some specific simplifying assumptions underlying the analysis and may be reduced using statistical techniques based on data resampling.

KEYWORDS: Tides; Spectral analysis/models/distribution; Uncertainty; Regression analysis; Coastal meteorology

1. Introduction

The astronomical nature of the tide-generating potential has led to a deterministic representation of the ocean tides as a sum of harmonic waves with known frequencies (Godin 1972; Doodson 1921). The classical harmonic analysis (HA) and its extensions are the approaches most widely used to assess tidal water level fluctuations at specific locations (Parker 2007; Thomson and Emery 2014; Guo et al. 2018). Despite some limitation due to its underlying assumptions (e.g., the temporal stationarity of water levels), HA remains the most reliable method for medium-to-long-term predictions (e.g., Hoitink and Jay 2016) and can be conveniently applied to a variety of situations, including detiding satellite altimetry data, assessing temporal changes in sea level, evaluating flood forecasting systems based on coupled tide-surge models, and specifying boundary conditions of hydrodynamical models (Cherniawsky et al. 2001; Talke et al. 2018; Beck et al. 2020; Wang et al. 2021).

For one-dimensional water level series, regression HA models typically represent the tidal heights as a superposition of sinusoidal curves characterized by the amplitudes $\beta_{s,k}$ and $\beta_{c,k}$ at specific frequencies $\omega_k$, $k = 1, 2, \ldots, K$, i.e. (Foreman and Henry 1989),

$$h_t = h_0 + \sum_{k=1}^{K} \beta_{s,k} \sin(\omega_k \tau) + \beta_{c,k} \cos(\omega_k \tau) + \epsilon_t,$$  \hspace{1cm} (1)

where $\tau$ denotes the time, $h_t$ the surface elevation evaluated at time steps $t = 1, 2, \ldots, T$, and $\epsilon_t$ the error term which represents the only stochastic (unresolved) component of the signal. Equivalent formulations of Eq. (1) express $h_t$ as the sum of $K$ positive- and negative-frequency sinusoids in the complex plan (e.g., Codiga 2011).
Nodal and astronomical corrections are typically included in the linear system to account for long period variations in constituent parameters (Foreman et al. 2009; Schureman 1958). Similarly, significant temporal trends and other meaningful signal components can be considered in the model as additional regressors. The $2 \times K + 1$ unknown parameters ($h_0$ and the $\beta$) of Eq. (1) are then estimated from the data with the ordinary least squares (OLS) or their generalizations (e.g., Leffler and Jay 2009; Strutz 2016) and inherit the statistical properties of the corresponding multiple regression model (e.g., Foreman and Henry 1989; Stephenson 2016; Davidson and MacKinnon 2004). For each constituent $k$, the tidal amplitude and phase are retrieved from the regression coefficients as $A_k = \sqrt{\beta_k^2 + \phi_k^2}$ and $\phi_k = \text{arctan}(\beta_{k,t}/\beta_{c,k})$.

From the spectral analysis theory (Jenkins and Watts 1969), the Fourier decomposition of the deterministic tidal signal used in Eq. (1) is well defined. The estimated regression parameters converge to the true astronomical values as the available series length increases (van Haren 2016). Longer records and higher sampling frequencies allow for more tidal constituents to be resolved as suggested by the Nyquist theorem and the Rayleigh criterion in its original formulation (Godin 1972). In common situations, however, sparse sampling and finite series length preclude a sharp distinction between the deterministic tidal signal and the stochastic fluctuations present in the tidal records. Many hydrological and meteorological phenomena (e.g., river discharge and storm surges; Hotink and Jay 2016; Maudsley and Haigh 2016; Arns et al. 2020) interact with tides and leave substantial residual energy in the spectrum of the recorded signal at frequencies close to the principal tidal spectral lines (e.g., hourly to seasonal frequencies; Munk et al. 1965; Munk and Cartwright 1966; Jay and Flinchem 1997). Similarly, estuarine and ocean stratification, baroclinic motions, and other natural (oceanic and climatic) and anthropogenic factors introduce variability, from chaotic high-frequency perturbations to long-term trends and cycles (Berntier and Thompson 2015; Haigh et al. 2020; Talke and Jay 2020). This may induce HA parameter modulations from short (e.g., subdaily to decadal temporal scales. Finally, errors associated with measurements (e.g., clock errors and instrument blockage; Marcos et al. 2015; Zaron and Jay 2014; Thompson et al. 2019) and inhomogeneities in data (e.g., datum shifts or changes in tide gauge location; Haigh et al. 2020; Talke et al. 2018) represent additional sources of uncertainty which further preclude a reliable estimation of some HA parameters, especially for constituents characterized by the smallest amplitudes.

Due to these unresolved sources of variability the $\epsilon_k$ error term of Eq. (1) does not follow the white noise assumption used in the OLS regression analysis. The residuals obtained after performing HA are temporally correlated. Ignoring this autocorrelation can cause severe underestimation of the uncertainty of the fit and parameter variability (Bos et al. 2014; Foster and Brown 2015), leading to incorrect assessments of the significance of one or more predictors (Sherman et al. 1998). Therefore, subsequent investigations based on such tidal reconstructions may be biased by a poor selection of the tidal constituents. Monte Carlo (MC) resampling and parametric bootstraps are often used to evaluate the propagation of background energy present in HA residuals to multivariate statistics of the estimated regression parameters. For instance, Pawlowicz et al. (2002) (T_Tide), Codiga (2011) (UTide), and Matte et al. (2013) (NS_Tide) used MC simulations to determine the $A_k$ and $\phi_k$ standard errors from the estimated $\beta_{k,t}$ and $\beta_{c,k}$ variances. These resamplings, however, rely on the precise knowledge of the $\epsilon_k$ probability distribution and autocorrelation structure, which are in practice unknown and difficult to estimate. As a result, the MC based on a priori error models ignore part or most of the autocorrelation expected for tidal residuals.

Residual bootstraps (Efron and Tibshirani 1994; Davison and Hinkley 1997) are simple resampling techniques widely used in geophysical disciplines for the quantification of the uncertainty in correlated time series (Vogel and Shallcross 1996; Mudelsee 2013; Hirsch et al. 2015; Ebtehaj et al. 2010). Contrary to parametric MC simulations, residual bootstraps make no assumption about the error term distribution to produce random replicates of the observed residuals (Kunsch 1989; Politis and Romano 1994). Despite this advantage, residual bootstraps still have limited application in tidal analysis [see Sherman et al. (1998), Ezer and Corlett (2012), and Vainiková et al. (2020) for some exceptions].

In the present study, we explore residual bootstrap methods as an alternative to the analytical procedures used in four MATLAB packages for estimating the precision of the tidal HA regressions. Specifically, the HA uncertainty assessment from one nonparametric and one semiparametric bootstrap are compared to some of those produced by four MATLAB tools widely used in applications: the “robustfit.m” function (Holland and Welsch 1977), the UTide (Codiga 2011) and NS_TIDE (Matte et al. 2013) packages, and the RT_Tide extension of the T_Tide package (Pawlowicz et al. 2002; Leffler and Jay 2009).

The aim is to assess each method’s statistical properties and their implications in terms of the regression statistical inference. The study has three specific objectives:

(i) Review the statistical definition of each analytical and resampling technique, identify the corresponding underlying assumptions, and evaluate their theoretical consistency with the HA model in Eq. (1) and the least squares (LS) framework.

(ii) Investigate the ability of each method to account for the colored nature of noise signals typically observed in tidal records; to this end, the parameter variances estimated by each method are compared over a large set of simulations with known tidal parameters and residual autocorrelation structure. Considering two limiting-case noise structures (i.e., white noise and strongly autocorrelated residuals), the study compares situations involving strong tidal and nontidal background energy and those in which the noise is purely random, independently of the hydrodynamical setting that produces the deterministic known signal.

(iii) Evaluate the validity of the confidence intervals resulting from each uncertainty assessment and the corresponding tidal constituent selection.
The paper is organized as follows. Section 2 presents an example of the uncertainty estimation methods available in the T_Tide, RT_Tide, UTide, and NS_Tide packages and motivates our focus on the packages’ default options in the remaining of the analysis. Section 3 reviews in detail the statistical basis of the selected methods using a common and unified notation. Section 4 introduces two bootstrap algorithms for HA. Section 5 describes the observed records of hourly water levels and simulations used in the analysis. Section 6 defines the steps and metrics used to compare the six uncertainty estimation methods. The impact on regression statistics of ignoring the residual autocorrelation structure is presented in section 7 by comparing the bootstrap and analytical methods for hourly water level series. Finally, section 8 provides a summary.

2. Preliminary method screening based on the regression model hypotheses

This section presents a preliminary comparison of the uncertainty assessment produced by the T_Tide and RT_Tide (Pawlucz et al. 2002; Leffler and Jay 2009), UTide (Codiga 2011), and NS_Tide (Matte et al. 2013) MATLAB packages when constructing the tidal amplitudes confidence intervals (CIs). The goal is to provide an example of the different $A_k$ variability estimates produced by the four packages for the same tidal model, LS estimator, and observed data (hourly water levels at the Halifax station, presented below, for the 7 November 2001–7 November 2002 period). Based on this example, this section motivates the use of the iteratively reweighted least squares (IRLS) approach for estimating Eq. (1) and the selection of only the default CI computation method of each package. These default options and the OLS and IRLS frameworks are discussed in detail in section 3. Furthermore, Table 1 summarizes the relevant steps and quantities involved in the CI construction for all the packages and options.

To compare the package CI options, Fig. 1 displays the median absolute deviations (MAD) [also referred to as median average deviation in Codiga [2011, Eq. (76)] of the tidal amplitudes estimated using the OLS (left) and the IRLS (right) approaches. This figure highlights two contradictions in the CI construction within the four MATLAB packages. First, the four packages allow the OLS estimation of the tidal parameters when the regression model assumes non-white noise errors (Fig. 1b, left). However, assuming a colored noise for the regression model corresponds to presuming serially correlated and heteroskedastic errors. In this case the regression errors cannot be independent and identically distributed (i.i.d.) as assumed by the OLS (see Table 2 for more details on LS models). Similarly, the use of the white noise error options for IRLS regressions (Fig. 1a, right) leads to an inconsistency since the IRLS estimator implicitly assumes a weighted least squares (WLS) model and heteroscedastic error terms (Table 2, second row). Finally, all the linear methods presented in Fig. 1b (T_Tide and RT_Tide linear, and UTide linear colored) produce constant $A_k$ MAD values over a fixed set of 9 frequency intervals centered on the major tidal constituents frequencies. Aiming at satisfying the assumption of locally white noise spectra [see Pawlowicz et al. (2002), Codiga (2011), and section 3 for more details], this result is inconsistent with the LS framework for both the OLS and IRLS methods. Moreover, the analytical derivation of the parameter standard errors involved in these linear methods is laborious and practically impossible to extend to regression models with additional terms and increasing complexity. Hence, the present application does not consider the package options based on the linearization of the amplitude and phase variances.

As second contradiction, Fig. 1 shows that the $A_k$ MAD estimated for the IRLS approach is generally lower than the corresponding OLS values. However, by construction the OLS approach yields the lowest empirical variance among the LS estimators since it consists in minimizing the residual standard error (Hastie et al. 2009; Davidson and MacKinnon 2004). While the result of OLS minimum variance is strictly valid only for the residuals and $\beta$ parameters, one may expect larger IRLS $A_k$ MAD values for part or most of the constituents. Intuitively, we can motivate this result by the fact that the IRLS require estimating the weights and, hence, a larger number of parameters than the OLS. In practice, it could also be shown that many IRLS algorithms force the weights to be smaller than or equal to 1.

Considering the theoretical inconsistency of the other package options, sections 3 and 7 investigate the possible misestimation of the IRLS parameter uncertainty only for the RT_Tide cboot, UTide MC colored, and the NS_Tide correlated methods presented in Fig. 1b. Methodological results and some application examples are also introduced for OLS regressions and T_Tide to support the discussion. The present study, however, does not present a systematic comparison between the OLS and IRLS parameter uncertainty since the two approaches are intended to be used in different experimental conditions: the OLS when dealing with homoscedastic observations and residuals, and the IRLS when the $h_i$ errors have variances that change in time, as in the context of tidal analyses. In the latter case, although the OLS tidal parameter estimates are not unrealistic on average (the estimated regression coefficients are still unbiased), they can only be used for descriptive purposes. All the regression uses that involve the parameter standard errors (e.g., selecting the significant tidal constituents or constructing prediction CIs) are invalid since the OLS regression variance–covariance estimate is biased. Conversely, if the estimated weights are approximately proportional to the inverse of the error variances, the IRLS approach is more efficient (i.e., it provides, on average, more accurate $\beta$ values and better standard error estimates) than the OLS. In the opposite situation of i.i.d. observations, the IRLS coefficients and their standard errors should closely agree with the OLS estimates.

3. Regression variance–covariance estimation in the T_Tide, RT_Tide, UTide, and NS_Tide packages

This section presents the statistical definition of the T_Tide and RT_Tide cboot, UTide MC colored, and NS_Tide correlated options for IRLS regressions and discusses their properties within the LS framework. Table 2
supposes this discussion with the detailed list of regression model hypotheses and relevant LS estimator expressions.

The OLS solutions of Eq. (1) are based on the assumption that HA tidal residuals are i.i.d. realizations of a Gaussian error term with variance $\sigma_\epsilon^2$. As such, the regression residuals are assumed to be generated by a stochastic process with a flat power spectrum (i.e., a white noise). Rewriting Eq. (1) in matrix form as $h = \Omega \beta + \epsilon$, where $h$, $\Omega$, and $\beta$ respectively represent the $T \times 1$ vector of observed water level values, the $T \times p$ matrix of the basis functions of the linear system, and the $p \times 1$ vector of $\beta$ parameters, $p = 2K + 1$, the OLS estimator of the regression is $\hat{\beta}_{OLS} = (\Omega^T \Omega)^{-1} \Omega^T h$. Accordingly, the variance–covariance matrix of $\hat{\beta}_{OLS}$ is expressed as (Table 2, first row)

$$\Sigma_{\beta}^{OLS} = (\Omega^T \Sigma_{\epsilon}^{-1} \Omega)^{-1} = \sigma_\epsilon^2 (\Omega^T \Omega)^{-1},$$

where $\Sigma_{\epsilon} = \sigma_\epsilon^2 I$ is the covariance matrix of the $T \times 1$ vector of i.i.d. residuals and $I$ is the identity matrix. Hence, $\Sigma_{\epsilon}$ is a $T \times T$ diagonal matrix with elements equal to $\sigma_\epsilon^2$. The
matrix $\Sigma_\epsilon$ reproduces the autocorrelation structure of a homoscedastic and uncorrelated noise, $\epsilon_t$.

When temporal trends or other nontidal signal components are included in Eq. (1), $\Omega$ and $\beta$ would respectively contain additional regressors and coefficients. The corresponding $\Sigma_\beta$ elements would thus relate to the variance–covariance structure of these additional signal components.

In tidal regressions, the variability of tidal amplitudes and phases is often estimated from $\Sigma_\beta$ by generating a series of constituent replicates using multivariate Gaussian MC simulations [see Pawlowicz et al. (2002), Codiga (2011), and Matte et al. (2013) for detailed descriptions of this MC approach]. Other methods that do not rely on the OLS hypothesis of Gaussian errors have also been used in other contexts (e.g., Munk and Cartwright 1966).

However, significant nontidal background energy and serially correlated disturbances are expected for tidal signals. Using $\hat{\beta}_{\text{OLS}}$ with such nonflat HA residual spectra (i.e., non-i.i.d. realizations of $\epsilon_t$) results in larger parameter estimator variances, which are underestimated by Eq. (2) (Mudelsee 2013). For this reason, several modifications to the OLS estimation of Eqs. (1) and (2) have been proposed in the literature.

First, Pawlowicz et al. (2002) proposed a $\Sigma_\beta$ estimation based on a piecewise representation of the $\Omega$ spectrum. Assuming a locally white tidal residual spectrum, the Pawlowicz et al.’s T_Tide MATLAB package (Pawlowicz et al. 2002) estimates the $\Sigma_\epsilon$ elements as the averages over a fixed set of nine frequency bins of the residual power spectral density (PSD). This results in up to nine distinct values $\sigma_1^2, \sigma_2^2, \ldots, \sigma_9^2$, with equal $\sigma_j^2$ attributed to all the constituents belonging to the same frequency bin. Then, T_Tide constructs one variance–covariance matrix, $\Sigma_{\beta_j}$, identical for all the constituents in the $j$th bin, using only the $\sigma_j^2$ value. For Eq. (1), $\Sigma_{\beta_j}$ has constant diagonal elements and possibly nonzero

FIG. 1. Parameter uncertainty estimated by the T_Tide, RT_Tide, UTide, and NS_Tide packages: amplitude median absolute deviations (MAD) used as standard errors for constructing the confidence intervals (CIs) of the (left) OLS and (right) IRLS parameter estimators when assuming (a) white noise and (b) colored noise residuals.
covariances between the $\beta$ values corresponding to the same tidal constituent; all other off-diagonal elements in $\Sigma_\beta$ are zero. Hence, the sine and cosine parameters of each constituent have the same variance, and all the $\beta$ pairs corresponding to different tidal constituents are uncorrelated. This strategy has three major drawbacks. First, the $T_{\mathrm{Tide}}$ approach is inconsistent with the OLS hypothesis of homoscedastic residuals. Second, it neglects the $(\Omega^T \Omega)^{-1}$ matrix that defines $\Sigma_{\beta}$ in Eq. (2). Considering that $\Omega$ is the sine-cosine basis function of the tidal model, the corresponding $\Sigma_\beta$ tends to overestimate the analytical variance of $\hat{\beta}_{\text{OLS}}$. Finally, $T_{\mathrm{Tide}}$ sets to 0 the residual PSD at frequencies close to the analyzed ones before the binning. This operation, named “line decimated method” in $T_{\mathrm{Tide}}$ scripts, results in the improper reduction of the residual energy and underestimated residual variances.

To mitigate the influence of outliers and heteroscedastic uncorrelated residuals, Leffler and Jay (2009) proposed a robust (IRLS) estimation of the $\beta$ values (Holland and Welsch 1977; Huber and Ronchetti 2011). Based on a WLS formulation of Eq. (1), the IRLS uses a weighting matrix $W$ to estimate the HA regression coefficients: $\hat{\beta}_{\text{IRLS}} = (\Omega^T W \Omega)^{-1} \Omega^T W h$ [see Table 2, second row, Leffler and Jay (2009), and Del Pino (1989) for more details]. Within the IRLS procedure the $W$ elements are estimated iteratively based on the studentized residuals and may involve predefined weighting functions (Holland and Welsch 1977).

The corresponding WLS variance–covariance matrices are expressed as (Table 2, second row) $\Sigma_{\text{WLS}} = \sigma^2 e W^{-1}$ and $\Sigma_{\beta}^{\text{WLS}} = \sigma^2 (\Omega^T W \Omega)^{-1}$,

$$\Sigma_{\beta}^{\text{WLS}} = \sigma^2 (\Omega^T W \Omega)^{-1}, \quad \Sigma_{\beta} = \sigma^2 (\Omega^T \Omega)^{-1}.$$

where $W^{-1}$ is proportional to the observation variances to account for the heteroscedasticity of the regression error terms. When all the observations have the same variance, $W = I$ and WLS becomes OLS. However, in some applications and software, the IRLS estimator of $\Sigma_\beta$ does not take into account the data heteroscedasticity. This is the case, for instance, of the robustfit.m MATLAB function, which defines the IRLS variance–covariance matrix as $\Sigma_{\beta}^{\text{robustfit.m}} = \sigma^2 e W^{-1}$ and $\Sigma_{\beta}^{\text{robustfit.m}} = \sigma^2 (\Omega^T W \Omega)^{-1}$,

$$\Sigma_{\beta}^{\text{robustfit.m}} = \sigma^2 e W^{-1} \quad \text{and} \quad \Sigma_{\beta}^{\text{robustfit.m}} = \sigma^2 (\Omega^T W \Omega)^{-1},$$

where $\sigma^2 e$ is a robust statistic for the residual scalar variance that involves the weight matrix $W$.

The RT_{Tide} extension of $T_{\mathrm{Tide}}$ uses the robustfit.m MATLAB function to estimate $\hat{\beta}_{\text{IRLS}}$. Then, the package applies the “line decimated band-average method” on the weighted residuals $e_W = e^T W$ to compute the $\sigma^2 e_w$ values. Finally, RT_{Tide} cboot uses the RT_{Tide} procedure to define a block diagonal $2K \times 2K$ variance–covariance matrix, i.e.,

$$\Sigma_{\beta}^{\text{RT}_{\text{Tide}}} = \hat{\Sigma}_{\beta} e_W \quad \text{and} \quad \Sigma_{\beta}^{\text{RT}_{\text{Tide}}} = \text{diag}[\Sigma_{\beta_1}, \ldots, \Sigma_{\beta_1}, \ldots, \Sigma_{\beta_K}],$$

where $\Sigma_{\beta_1}$ is the band-averaged power spectrum of $e_W$ and each block $\Sigma_{\beta_1} = \Sigma_{\beta_1}$ for all the constituents in the $j$th frequency bin. When using OLS regressions, the RT_{Tide}
resumes to T_Tide since the two approaches only differ in the use of the IRLS weights in the residual spectrum calculation and minor computational details (i.e., RT_Tide allows for different MATLAB functions to compute $P_e$).

Similarly to T_Tide, RT_Tide assumes zero correlations between the $\beta$ values of different tidal constituents and neglects $\Omega$ in Eq. (5), while the quadratic term $(\Omega^{2}W\Omega)^{-1}$ should be used as in Eq. (3). Accordingly, $\Sigma_{\beta}^{RT_Tide}$ is expected to overestimate the $\hat{\beta}_{IRLS}$ variances for typical tidal HA applications of Eq. (1).

Consolidating and building from Pawlowicz et al. (2002), Leffler and Jay (2009), and Foreman et al.'s (2009) developments, Codiga (2011) proposed to adjust the OLS and IRLS variances–covariance matrices with the residual spectral power. However, while T_Tide uses the PSD to estimate the OLS $\hat{\Sigma}_b$ (i.e., its diagonal), Codiga’s default $MC$ colored strategy in the UTide package uses the spectrum of the weighted residuals to scale the—previously computed—regression variance matrices as

$$\Sigma_{\epsilon,UTide} = \sigma^2_{\epsilon,UTide} W^{-1} P_e$$

and

$$\Sigma_{\beta,UTide} = \sigma^2_{\beta,UTide} (\Omega^{2}W\Omega)^{-1}. $$

(6)

The $P_e$ overbar indicates that the spectrum is approximated as a staircase function over nine frequency intervals (as in T_tide and RT_Tide) and estimated from the line-decimated (i.e., energy-reduced) band-averaged spectrum of $\epsilon_w$. Note that Eq. (6) is strictly valid only for univariate series treated with the real model formulation of Eq. (1). Using a complex formulation of the harmonic regression model, Codiga (2011) included in $\Sigma_{\beta,UTide}$ nonzero covariances between the $\beta$ parameters related to the same tidal constituent. These off-diagonal elements are not represented in Eq. (6) since they are null for real series.

The weight matrix $W$ appears three times in Eq. (6): once within the $\Sigma_{\beta,UTide}$ matrix, once in the estimation of the $T_{\epsilon_w}$, and once in the estimation of $\sigma^2_{\epsilon}$, which is computed as the corrected weighted covariance between the observed water levels and the residuals, i.e.,

$$\sigma^2_{\epsilon,UTide} = \frac{1}{(T - 2 K)} \sum_{t=1}^{T} \epsilon_t W_t h_t = \epsilon^T W h / (T - 2 K).$$

(7)

This corresponds to a triple smoothing of the residual variance since the weights in $W$ are sample estimates of the reciprocal $\epsilon_t$ variances (Davidson and MacKinnon 2004; Faraway 2002). This improperly reduces the parameter variability below the WLS and OLS variances and, consequently, underestimates the $\beta$ standard errors. The reader can refer to sections II.B and II.C [Eqs. (51)–(80)] of the UTide documentation for the detailed derivation of Eqs. (6) and (7) in complex formulations of the tidal HA regression.

An estimation based on a generalized least squares (GLS; see Table 2, third column) representation of $\Sigma_p$ was proposed by Matte et al. (2013) and included in the NS_Tide package. The GLS allows for colored noise by explicitly modeling the autocorrelation structure of the observations (Davidson and MacKinnon 2004). The NS_Tide correlated approach uses Fourier transformations to compute the observation spectrum and the average power of the weighted residuals and $\beta$ parameters (i.e., their variances). According to this spectral representation, Eq. (3) can be rewritten as

$$\Sigma_{\epsilon,NS_Tide} = P_e$$

and

$$\Sigma_{\beta,NS_Tide} = [F_{W1} P_e W_{1}^{-1} F_{W1}]^{-1}. $$(8)

where $F_{W1}$ is the Fourier transform of the weighted observations $W_1$ and $P_e$ represents the weighted residual power evaluated at $t = 1, 2, \ldots, T$ [a more concise notation than the original Eqs. (14)–(16) of Matte et al. (2013) is used here]. The absence of the overbar indicate no band averaging of the residual power. A detailed description of the GLS method and the analytical derivation of the NS_Tide equations are given in appendix A.

The NS_Tide approach has two shortcomings. First, Eq. (8) modifies the GLS variance–covariance expression by using the Fourier transforms of the weighted residuals and observations instead of the unweighted $P_e$ and $F_{W1}$ quantities. This leads to the underestimation of the $\beta$ variances, since $W \propto \Sigma_{\beta}^{-1}$. Section 7c shows that this NS_Tide defect is avoided by simply replacing the weighted residual by $e$ when computing the Fourier transforms in Eq. (8). In that case, $\Sigma_{\beta,NS_Tide}$ would correspond to the variance–covariance matrix of a GLS regression [see Eq. (A3)]. Second, while Eq. (8) acknowledges the heteroscedastic and correlated nature of tidal residuals, Matte et al. (2013) did not include the $F_{W1}$ and $P_e$ spectra in the regression model [i.e., the authors omitted Eq. (8) from the LS estimator of the $\beta$ parameters as the GLS model would require; see Table 2, third row, and appendix A]. Hence, $\Sigma_{\beta,NS_Tide}$ and $\Sigma_{\epsilon,NS_Tide}$ are not consistent with the $\hat{\beta}_{IRLS}$ used for estimating the regression coefficients.

Figure 2 graphically summarizes the workflow of parameter variability estimations for the approaches that will be considered in the rest of the analysis: robustfit.m, RT_Tide $eboot$, UTide $MC$ colored, and NS_Tide $correlated$, and the bootstrap approaches presented in the next section. Figure 2 and the remaining sections do not report the package option names for simplicity.

4. Bootstrap methods for tidal records

Let $e = [e_1, e_2, \ldots, e_T]$ be the residuals from an HA model fit, computed by subtracting the astronomical tide prediction from the observed water levels. Residual bootstraps are a class of methods which reuse the $e$ series to construct many replicates of the water level series and base the statistical inference on these replicates. Each replicate (or resample) has the same length as the original water level record and, if well constructed, reproduces the statistical properties of the observed phenomenon. This section presents two residual bootstraps adapted to HA regressions: a moving block bootstrap (MBB) and a semiparametric bootstrap (SPB). Both methods are intended to produce similar results with comparable accuracy. While the MBB can be applied in virtually all sampling conditions, it is more computationally demanding than SPB, especially for long records.
On the other hand, the SPB is faster than the MBB but cannot be directly applied to irregularly sampled records, since it involves a Fourier transform of residuals.

a. MBB

The MBB relies on the principle that the autocorrelation of the HA errors can be reproduced by sampling with replacement many sufficiently long subseries of contiguous residuals (i.e., blocks), while residuals belonging to different blocks are nearly independent. MBB constructs the replicates via the following steps:

1) Draw with replacement $B$ uniform random values $i_b$ from the set of equally spaced indices $\{1, 2, \ldots, N_T\}$, where $N_T$ is the number of time steps separating $e_1$ and $e_T$; each $i_b$ represents the index of the starting time of a block of residuals to be sampled from $e$.

2) For each $b$, construct the $b$th block of $l$ residuals as $e = \{e_{i_b}, e_{i_b+1}, \ldots, e_{i_b+l-1}\}$, with $i_b$ being the residual time index that satisfies $i_b \leq i_b + l - 1$; the present application considered $l = 30$ (days), as this block length allows to capture most autocorrelation expected for tidal HA residuals series. Preliminary analyses also showed a little gain in the use of blocks of variable length and low sensitivity for $l$ ranging in $(25, 35)$ (not shown). Detailed discussions about the choice of $l$ and the use of MBB with random block lengths can be found in Mudelsee (2013, 2019), Beyaztas and Firuzan (2016), and references therein.

3) Concatenate the $B$ blocks as $e_1, e_2, \ldots, e_B$ to construct a series of $N^*_B$ bootstrapped residuals, with $T^*_b = \sum_{b=1}^{B} i_b \geq T$. Discard the last $T^*_b - T$ bootstrapped values, namely, the residuals for indices that exceed the original series length.

4) Define the surface elevation resample as the sum of the original HA model fit and the bootstrapped residuals constructed at step 3.

5) Repeat steps 1–4 as sufficient high number of time ($R = 1000$ in the present application).

Note that the MBB algorithm can be safely applied to irregularly spaced records since it resamples blocks of time indices (i.e., noncontiguous time intervals). Specifically, it avoids the oversampling of the residuals from periods with denser sampling by using equally spaced time indices at step 1 (Hirsch et al. 2015; Lahiri and Zhu 2006).

b. SPB with Fourier transform of residuals

The SPB resampling considered in the present study is obtained via the following steps:

1) Estimate the spectrum or PSD of the $e$ series using a non-parametric estimator, such as the fast Fourier transform (FFT) method.

2) Generate $R$ independent and equally probable replicates of the observed residuals based on the PSD estimated at step 1. In our application, uniform random phases were added to the residual PSD amplitudes to generate noise replicates in...

FIG. 2. HA uncertainty estimation in analytical (robustfit.m, RT_Tide cboot, UTide MC colored, and NS_Tide correlated) and bootstrap (MBB and SPB) methods. In the figure, the package option names have been omitted for simplicity.
the frequency domain; then, we applied an inverse FFT transformation to recover residual resamples in the time domain.

3) Finally, construct $R$ surface elevation resamples by summing the $R$ simulated series of residuals to the original HA model reconstruction.

In the present application, an FFT and inverse FFT algorithms were used for $h_1$ records, as they are fast and accurate estimators under perfect sampling conditions. One-off missing values in the residual series can be replaced via linear interpolation of the values at neighboring time steps. However, this interpolation may introduce nonnegligible biases in the FFT estimate of the PSD in the presence of a large number of data gaps and missing values (Munteanu et al. 2016). It would be important to evaluate the performance of the proposed SPB in such situations.

The adaptation to multivariate series (e.g., tidal currents) and different HA model formulations would only require to consider the vector of multivariate residuals for the produced regression fit.

c. Bootstrap and Monte Carlo estimators

For both methods, the bootstrapping considered $R = 1000$ resamples and plug-in estimators of the HA model statistics (Efron and Tibshirani 1994). Based on the plug-in principle the standard error of the tidal amplitude, $A_k$, of the $k$th constituent was estimated as the sample standard deviation of the $R$ replicates $\{A_k^{(r)}\}_{r=1}^R$, where $A_k^{(r)}$ is the value estimated on the $r$th bootstrap resample:

$$\hat{A}_k^2 = \frac{1}{R} \sum_{r=1}^R [A_k^{(r)} - \bar{A}_k]^2, \quad k = 1, 2, \ldots, K,$$

(9)

where $\bar{A}_k^2$ (cm) is the arithmetic mean of the $A_k^{(r)}$ values. For the same constituent, the circular variance ($\hat{\nu}_k$) of the tidal constituent phase was computed as

$$\hat{\nu}_k = 1 - ||\phi_k^{(r)}||, \quad k = 1, 2, \ldots, K,$$

(10)

where $\phi_k = \tan^{-1}(\beta_{k1}/\beta_{k2})$ is the phase of the $k$th constituent and the mean bootstrap phase $\phi_k^{(r)}$ is computed by averaging the cosine and sine components of the $R$ phase replicates, i.e. (Berens 2009),

$$\bar{\phi}_k^{(r)} = \tan^{-1}\left[ \frac{1}{R} \sum_{r=1}^R \sin \phi_k^{(r)}, \frac{1}{R} \sum_{r=1}^R \cos \phi_k^{(r)} \right].$$

Within the circular statistics framework (Pewsey 2004), the norm of the mean resultant vector of a set of angles is interpreted as a measure of their circular dispersion. The closer $||\phi_k^{(r)}||$ is to 1, the less dispersed are the bootstrap phases around their circular mean [Eq. (11)], while $||\phi_k^{(r)}|| \approx 0$ when the $\phi_k^{(r)}$ angles are uniformly distributed over the unit circle (i.e., for maximum variance $\hat{\nu}_k = 1$). The reader can refer to Berens (2009) and Jammalamadaka and Sengupta (2001) for further details on circular statistics.

Following the plug-in approach, the robustfit.m, RT_Tide, UTide, and NS_Tide variance-covariances [i.e., Eqs. (4)–(8)] were used to produce $R$ replicates of the HA regression parameters through a multivariate Gaussian MC. Accordingly, the uncertainty of tidal amplitude and phase estimates was computed based on the variability of the bootstrap or MC $\beta$ replicates.

5. Data and simulation setup

Eleven years of data were extracted from the Fisheries and Oceans Canada (2021) archives for two gauge stations in eastern Canada: Halifax, St.1 = (44.66°, −63.58°), and Saint Francois Orleans Island, St.2 = (48.45°, −68.52°). Hourly water levels for the 1 December 1998–30 November 2009 period were used. These records are long enough to cover a large variety of ocean, storm surge, and river-flow conditions, since the two stations are respectively located on the North Atlantic coast of Canada (St.1, not influenced by fluvial processes), and within the upper Saint Lawrence River estuary (St.2). These records involving many unsolvable subtidal and nontidal disturbances on major astronomical tides, they are representative of the typical serial correlation of the noise affecting tidal signals. The reader can refer to Matte et al. (2014a, 2017) and Thompson et al. (2009) for comprehensive descriptions of the hydrodynamics in these areas.

A Monte Carlo experiment was then used to construct a large number of water level series with known properties, namely, the number of tidal constituents generating the deterministic tidal signals, their amplitudes and phases, and the structure of the stochastic error generating the HA model residuals. Several sets of simulations were used to differentiate the effects of the neglected residual autocorrelation from those arising from two other major sources of uncertainty: 1) the interannual variability of tidal constituents due to the interactions with the local hydrology, natural climate variability, and satellite (nodal) modulations; and 2) the shortness of the series with respect to the period needed to separate some tidal constituents. Cycles at scales shorter than one year, on the other hand, were explicitly included in the HA model by using low-frequency and monthly tidal constituents (e.g., SA and SSA; see below), although part of the hydrological variability may not be fully captured at nontidal frequencies. Figure 3 schematically represents the Monte Carlo experiment workflow.

To construct realistic simulations, the $\beta_{k1}$ and $\beta_{k2}$ parameters were first estimated on the 11-yr records using UTide considering 68 tidal constituents and embedded nodal corrections (Foreman et al. 2009). For the exhaustive list of constituents see Fig. 7 (y axes). The 68 estimated constituents were used to generate 100 hourly water level simulations over periods of length between 6 months and 3 years. Each period had random initial time $t_0$ drawn between 1 January 1999 and 31 December 2006 to ensure the availability of 3 years of data for each simulation. For simulating residuals that are representative of the wide range of noise structures that may affect the observed records two noise models were considered (Fig. 3):

1) a white noise (WN) signal with mean and variance equal to the corresponding empirical statistical moments of the original 11-yr residuals.
2) an autocorrelated resampled noise (RN) constructed by selecting $T = 100$ contiguous residuals from the original 11-yr reconstruction over the 1999–2006 period.

Note that the $T$ initial times considered for this selection differ from those used for the tidal reconstructions. These residuals were added to each water level reconstruction to represent two opposite situations: the one representing the (unrealistic) situation of purely random residual energy and the other involving substantial energy from unresolved signal components. Further, the original record residuals observed at the $t_0$ used for reconstructing the tidal signals were also extracted. Accordingly, the series constructed with these residuals correspond to the original records observed for 6-month to 3-yr periods that start at the $t_0$. They are hereinafter referred to as OBS series.

As intended, the OBS and RN series display comparable residual spectra, while the WN residual spectrum is flat (see, e.g., the PSDs shown in Fig. S1 in the online supplemental material). At St.2, the mean and standard deviation of the residual PSD is larger than at St.1, most likely due to its estuarine location where the influence of river flow on the tidal signal starts to appear. The preliminary analysis of these residual spectra motivated the choice of limiting the number of constituents to 68 since, otherwise, the 11-yr regression residuals would have been unrealistic for shorter series. Also, the regressions on short simulations (e.g., the 6-month series) would have been algebraically intractable with more constituents.

Despite the use of nodal corrections in estimating Eq. (1), one can expect a certain temporal variability of the $A_k$ and $\phi_k$ values due to the effects of satellites on resolved constituents (Godin 1986; Foreman and Henry 1989). To differentiate and isolate the uncertainty of the HA estimates linked to the unexplained residual variability from the satellite modulations of tidal constituents, 500 additional WN and RN simulations were produced starting on 11 November 2001 for periods of 6 months, and 1 and 3 years. These simulations (hereinafter “fixed start simulations” indicated as $h^*$) thus correspond to equally probable tidal signals that could have been observed over the 11 November 2001–10 November 2004 period without the effects of the astronomical temporal oscillations of the tidal constituent parameters (Fig. 3).

Table 3 summarizes the characteristics of the MC experiments, highlighting the uncertainty sources controlled in each specific set of simulations.

6. Metrics for comparing bootstrap and analytical methods

For comparing the uncertainty assessments from robustfit.m, UTide, RT_Tide, NS_Tide, and the two bootstraps the analysis proceeded with the following steps: 1) estimate Eq. (1) on each MC simulation, 2) compute the $\beta$ variability using the six uncertainty estimation methods for each MC simulation, and 3) construct the tidal amplitude CIs corresponding to the estimated standard errors and evaluate their adequacy. This section details the methods used in each step.

For each simulation, the HA parameters for the 68 tidal constituents were estimated by the robust IRLS procedure with Cauchy weight function and tuning constant equal to 2.385 since they represent the UTide and robustfit.m default (Codiga 2011). The IRLS applies a WLS for heteroscedastic residual regressions and have been shown to reduce the noise influence on HA parameter estimates. These estimates were obtained with an adaptation of the UTide package (Codiga 2011) for treating the real formulation of the tidal model presented in Eq. (1). They were then used for all bootstrap and analytical uncertainty estimation methods. This choice was motivated by the need to have the same $\beta$ accuracy for all the produced uncertainty assessments.
from the same $\beta$ values allows attributing the $\beta$ standard error differences only to the different variance estimation methods. The default nodal corrections and astronomical arguments computed by UTide were included in the analysis.

For robustfit.m, UTide, RT_Tide, and NS_Tide, a multi-variate normal distribution with variance-covariance matrix, $\Sigma_f$, [Eqs. (4)–(8)] was used to simulate $R = 1000$ MC realizations of the HA regression parameters and corresponding $A_k$ and $\phi_k$ values, as described, for instance, in Codiga (2011) and Matte et al. (2013) (see also Fig. 2). The HA uncertainty assessments produced by the six methods were first compared in terms of the $A_k$ standard errors [$\sigma^2_{A_k}$, Eq. (9)] and $\phi_k$ circular variances [$\nu_{\phi_k}$, Eq. (10)]. Since the methods share the same $\beta_{IRLS}$, the differences in $\sigma^2_{A_k}$ and $\nu_{\phi_k}$ values can be directly attributed to the different estimation of the regression uncertainty. Specifically, the differences in estimated standard errors are a direct result of the different statistical assumptions involved in each method, as the simulations in each MC experiment have equal known residual structure.

The standard errors computed for each method were then used to construct CIs for the 68 tidal amplitudes based on a Gaussian approximation of the $A_k$ distributions. By definition, each CI approximates a range of plausible values for an unknown parameter that should contain the true parameter value with known probability ($1 - \alpha$). Hence, for each constituent and $\alpha$ value, the effective probability coverage of a given amplitude CIs can be approximated using a high number of equiprobable realizations of the data generating process (Zwiers 1990; Schall 2012). Since the fixed start $h_k$ simulations represent sets of equiprobable water level series that could be observed at a given location for a given period, they were used to estimate the effective CI probability coverage, as the proportion of intervals CI($\alpha$, $A_k$) which recover the true $A_k$ value (i.e., the value used for simulating the tidal signal in the MC experiment). Specifically, for $\alpha = 0.01, 0.05, \text{and} 0.1$, the nominal ($1 - \alpha$) confidence levels were compared to the effective CI probability coverage, ($1 - \alpha_{\text{eff}}$), estimated as

$$1 - \alpha_{\text{eff}} = \frac{1}{M} \sum_{m=1}^{M} I(\hat{A}_k \in \text{CI}_{\alpha}([\alpha, A_k])),$$  (12)

where $I(\hat{A}_k \in \text{CI}_{\alpha}([\alpha, A_k]))$ takes value 1 if the true $A_k$ value is in the CI computed for the $m$th simulation, 0 otherwise. For MBB and SPB, several alternative CI definitions based on the parameter bootstrap distributions exist and could be used to improve the effective interval coverage, i.e., to have $(1 - \alpha_{\text{eff}})$ closer to $(1 - \alpha)$ (Efron 1987). However, preliminary analysis revealed only small CI coverage improvements for percentile, bias corrected, and accelerated bias-corrected bootstrap methods. More robust methods such as the studentized bootstrap (Dixon 2006) should be considered to improve MBB and SPB CIs but they would involve substantial increase in computational costs.

7. HA uncertainty estimates on simulated data

This section compares the four analytical estimation methods with the two bootstraps in terms of the computed parameter variability (section 7a) and the correctness of the corresponding CIs (section 7b).

a. Amplitude standard errors and phase circular variances

Figure 4 shows, for each method and station, the median of the 100 $\hat{\sigma}_{A_k}$ [Eq. (9)] estimated for annual OBS, WN, and RN series. The amplitude standard errors at St.2 are larger than the corresponding estimates at St.1 since at the estuarine location the overall residual energy from unresolved signal components is larger. For the OBS and RN series (Figs. 4a,c), the
Ak values estimated by all methods except robustfit.m decrease with increasing frequency. This reflects the characteristics of the red and pink spectra used for perturbing the tidal signals in the MC simulations (see, for instance, Fig. S1). Instead, robustfit.m ignores the colored nature of the simulated residuals and returns standard errors uniform over the frequency spectrum, as if the series were perturbed by white noises. As a result, the robustfit.m Ak values are either too small (for low-frequency components) or too large (for the highest frequencies). Specifically, the robustfit.m estimates are close to the OLS standard errors since the method uses the IRLS weights only for computing the residual scalar variance estimate, σ²_e,0. The complementary example provided in appendix B supports this analytical conclusion (e.g., Figs. B1a,b). In this regard, note however that using blocks of length 30 days may also lead to the MBB underestimation of the Ak variability for high-frequency constituents. In contrast, RT_Tide and NS_Tide reproduce the nonflat relationship between the Ak and the constituent frequency. However, for constituents with periods shorter than 2 days the RT_Tide amplitude standard errors are generally higher than the corresponding bootstrap values, while NS_Tide Ak values are typically lower than the MBB and SPB standard errors. For constituents with periods longer than 2 days, the RT_Tide Ak values are close to the corresponding MBB values. The discussion of section 3 explains these results: on the one

**FIG. 4.** Amplitude standard errors [Eq. (9)] for each tidal component (x axis) at (left) St.1 and (right) St.2: median over 100 annual series for (a) OBS, (b) WN, and (c) RN series.
hand, an overestimation of the parameter uncertainty is expected for RT_Tide due to the omission of the regressor quadratic term in \( \Sigma_{RT,Tide} \) [Eq. (5)]; on the other hand, the weighting of \( \Omega \) and \( \epsilon \) in the Fourier representation of \( \Sigma_{NS,Tide} \) [Eq. (8)] induces the underestimation of the \( \sigma_{Ak} \). The complementary OLS results in appendix B further show that the absence of the IRLS weights increases the RT_Tide(T Tide) overestimation but improves the NS_Tide performance (e.g., Fig. B1b). In fact, these analytical methods use the \( W \) matrix to decrease the values of the \( \Sigma_p \) diagonal elements.

For the OBS at St.2 (Fig. 4c, right) and the RN series (Fig. 4c), MBB and SPB are approximately equivalent and consistently convert the residual energy into the amplitude standard errors, as the variation of their \( \sigma_{Ak} \) over the frequency spectrum reproduces the shape of the simulated residual spectra. However, SPB shows low \( \sigma_{Ak} \) values for the St.1 OBS (Fig. 4c, left), especially at low frequencies, possibly due to the larger number of missing values at this location (Munteanu et al. 2016). Also, at the lowest frequencies, the greater emphasis given to the residual FFT, may induce underestimating the cumulative energy of the residual spectrum, especially for relatively short series (Munteanu et al. 2016; Troncosi and Pesaresi 2019; Kihm et al. 2013).

As expected, the WN \( \sigma_{Ak} \) were approximately constant for all \( k \) and methods and both locations (Fig. 4b). Good robustfit.m performance was anticipated for these series since the HA regression satisfies the OLS assumptions of i.i.d. Gaussian residuals. Interestingly, RT_Tide shows results close to robustfit.m since the power spectrum of the weighted residuals well approximates the diagonal of an OLS variance–covariance matrix for WN residuals [i.e., \( \mathbf{p}_{\mathbf{e}_w} \) is close to the \( \Sigma_{OLS}^{D} \) diagonal since the IRLS weights are inversely proportional to \( (\Omega^T\Omega)^{-1} \)].

For WN series, the block bootstrap may underestimate the total level of uncertainty, as suggested by the fact that the MBB \( \sigma_{Ak} \) range between the NS_Tide (underestimated) and robustfit.m standard errors. Equally important, for the WN series the SPB produced the lowest \( \sigma_{Ak} \) values for many constituents (Fig. 4b). Attributing this underestimation only to the use of the FFT is less obvious since other methods (e.g., NS_Tide) also involve the Fourier transform. Further analyses on spectral-based simulation methods are needed to identify the causes of the poor SPB performance on WN series.

The PSD band averaging used in UTide visibly alters the slope (i.e., autocorrelated) structure of residuals, returning less \( \sigma_{Ak} \) values than the number of constituents in the least squares fit. Moreover, UTide substantially underestimates the amplitude variability for all series types (Fig. 4) and generally displays small \( \sigma_{Ak} \) spread over the 100 simulations (Fig. S2 shows an example). The double smoothing of \( \mathbf{p}_{\mathbf{e}_w} \) with both the IRLS weights and the line-decimated spectrum \( \mathbf{p}_{\mathbf{e}_w} \) of the weighted residuals [Eq. (6)] explains these results. In fact, when using the UTide MC white option, this underestimation vanishes, as the method is closer to the WLS (for an example, see Fig. 1a, right). For the same reason, MC white and MC colored options for OLS regressions (Figs. B1a, b in appendix B) return higher \( \sigma_{Ak} \) values and results generally closer to the bootstrap methods.

The analysis of the median \( \hat{v}_{Ak} \) [Eq. (10)] presented in Fig. 5 confirmed these findings. In particular, UTide and NS_Tide returned lower phase circular variances than the other methods, while the RT_Tide estimated the largest \( \hat{v}_{Ak} \) values. robustfit.m underestimates (overestimates) the \( \phi_k \) variability estimated by the other methods for low (high)-frequency constituents. However, contrary to what is observed for \( \sigma_{Ak} \), the median \( \hat{v}_{Ak} \) values present a flat structure over the frequency domain (Fig. 5). This is expected since spectral methods make no explicit use of the information about phase variability. Also, the robustfit.m and RT_Tide \( \hat{v}_{Ak} \) dispersion over the 100 series were generally lower than for the bootstraps, while UTide and NS_Tide returned the highest dispersion (Fig. S4).

Noteworthy is the similarity of the results presented for St.1 and St.2 in Figs. 4 and 5. While the St.2 median \( \sigma_{Ak} \) are overall higher than the corresponding St.1 estimates, comparing the six uncertainty estimation methods leads to the same overall conclusions at the two locations. For this reason, the rest of the manuscript will mainly present results for St.1, far from fluvial influence. Unless explicitly shown and mentioned, equivalent results were obtained for St.2.

Figure 6 summarizes the differences between the six methods by showing the distribution over the 100 OBS, WN, and RN simulations of the following mean statistics:

\[
\delta_{\sigma_A} = \frac{1}{K} \sum_{k=1}^{K} \sigma_{Ak}^{MBB} \quad \text{and} \quad \delta_{v_A} = \frac{1}{K} \sum_{k=1}^{K} v_{Ak} - v_{Ak}^{MBB},
\]

where the MBB statistics are used as reference since, differently to other methods, no theoretical or empirical inconsistency has been evidenced in previous sections.

Results for series spanning periods from 6 months to 3 years are consistent with what was observed for the annual simulations in previous figures. Specifically, Fig. 6 shows that the differences between MBB and all other methods except RT_Tide tend to decrease with the series length. For most of the series perturbed by colored noise (OBS and RN), RT_Tide returned the largest discrepancies, with \( \sigma_{Ak} \) 1.5–3 times larger, on average, than MBB standard errors and \( \delta_{v_A} \) approximately between 0.1 and 0.2 (Figs. 6a–c). For WN series, RT_Tide showed good performance that does not heavily depend on the available series length. On the other hand, UTide presents the largest underestimation for both amplitudes and phases and all series types. For instance, UTide \( \sigma_{Ak} \) are up to 50% lower than MBB for series shorter than one year (Figs. 6a–c), while UTide \( \delta_{v_A} \) take values approximately between −0.1 and −0.025 (i.e., the largest median phase differences reported in Figs. 6d–f). NS_Tide showed only slightly better performance than UTide, with \( \delta_{\sigma_A} \approx 0.4 \) and \( \delta_{v_A} \approx 0.1 \) for OBS and RN series. Reminding that \( v_{Ak} \) ranges between 0 and 1, these \( \delta_{v_A} \) values stress the important underestimation of circular variances when using UTide and NS_Tide.

The differences between robustfit.m and MBB show that the analytical method tends, overall, to overestimate the HA parameter variability, with both \( \sigma_{Ak} \) and \( v_{Ak} \) being generally larger than the corresponding MBB estimates. For St.1, \( \delta_{v_A} \) diminishes when increasing the series length, approaching 1 for 3-yr series. This suggests that the \( \sigma_{Ak} \) underestimation for low-frequency components tends to be compensated by
overestimating the amplitude standard errors for high frequencies. For St.2., the robust fit.m $\delta_{\nu}$ are small and negative for OBS and RN series (Fig. S5).

At St.1, SPB underestimates the MBB $\sigma_{\phi}$ of roughly 10% and the MBB $\nu_{\phi}$ by approximately 0.1 (Fig. 6). At St. 2 the two methods are mostly equivalent for all series lengths and correlated noises ($\delta_{\phi}$ and $\delta_{\nu}$ are close to 1 and 0, respectively, for OBS and RN series; e.g., Fig S5). For both locations, the average difference between SPB and MBB $\sigma_{\phi}$ estimates increases with the series length for WN, and SPB underestimates by approximately 50% the MBB variability for 3-yr WN simulations (e.g., Fig. 6b). This advises for the possible inconsistency of SPB estimators, mostly detectable for WN, as increasing the series length further increases the differences between SPB and the other methods.

The ensemble of these results implies that the MBB method would estimate higher variability than the UTide, NS_Tide, and SPB methods for the water level predictions and derived statistics. For instance, it could be shown that UTide underestimation of parameter variability results in standard errors of hourly water levels that are 50% to 60% lower than the corresponding MBB estimates (for an example on annual RN series at St.2, see Fig. S8). In our examples, UTide underestimates by roughly 3 cm the MBB standard errors and by 12 cm the 95% CI width of reconstructed hourly $h_{t}$ value. Such differences in the uncertainty

![Figure 5](https://example.com/fig5.png)

**Fig. 5.** Phase circular variances [Eq.(10)] for each tidal component (x axis) at (left) St.1 and (right) St.2: median over 100 annual series for (a) OBS, (b) WN, and (c) RN series.
FIG. 6. Comparison of parameter variability estimates at St.1: distribution over 100 simulated series of (a)–(c) the mean ratio between amplitude standard errors and (d)–(f) the mean difference between phase circular variances for series of different lengths. MBB standard error and circular variances are used as reference.
assessments may be significant in applications that need centimeter-scale or higher accuracy water level reconstructions (e.g., for verifying altimetry products or estimating sea level trends; Foreman et al. 1995; Stammer and Cazenave 2017; Li et al. 2021).

b. Confidence intervals

To definitively evaluate the adequacy of the six uncertainty estimation methods, Fig. 7 displays the probability coverage [Eq. (12)] of the amplitude CIs constructed at the 0.95 confidence level with each method using 500 WN and RN $h^*$ simulations. The darker the color, the lower is the fraction of CIs containing the true $A_k$ value used for simulating the tidal signals. Despite encompassing the true tidal amplitude, the intervals with too high coverage (gray crossed boxes) indicate an overestimation of the parameter uncertainty and the loss of discriminant power for the corresponding hypothesis testing (e.g., reduced power of the tests for evaluating the constituent significance). The CI $(1 - \alpha)$ level is reached when all the assumptions about the HA regression model and data distribution are met, including those related to the stochastic process generating the error term. Conversely, CIs do not reach the nominal confidence level if the hypotheses about the parameter distributions are violated, and, most often, this is due to neglecting the data autocorrelation (Zwiers 1990).

As expected from previous results, UTide presents $\alpha_{eff}$ close to the nominal confidence level only for few tidal constituents and underestimates the width of amplitudes CIs (Fig. 7a).

NS_Tide shows slightly better performance by producing CIs with effective coverage close to $(1 - \alpha)$ for some diurnal and higher-frequency constituents, especially for WN series (Fig. 7d). For frequencies lower than approximately 1 cycle per day, the robustfit.m CIs are also narrower than those constructed with the other methods for RN series, resulting in $1 - \alpha_{eff} < 0.3$ for long-term constituents. Conversely, the robustfit.m overestimated the CI for constituents with period shorter than 12 h (Fig. 7b). This confirms the robustfit.m tendency to underestimate the spread of the amplitude distributions for low-frequency components and overestimate the high-frequency constituent uncertainty, when water level series contains autocorrelated noise. However, when the residuals are not serially correlated (WN series), the robustfit.m mostly produces CIs with coverage close to the nominal level. Comparing these results with the OLS example in appendix B emphasizes that the UTide and NS_Tide underestimation is linked to a misuse of the IRLS weights, while robustfit.m applies weak modifications to the OLS $\sigma_{Ak}$ values (e.g., Fig. B1d).

For RN series, RT_Tide generated too wide CIs for most of the constituents except for three frequency bands: long-period constituents (Sa–Mf) having $0.7 \leq 1 - \alpha_{eff} < 0.95$; a group of constituents around $M_2$ ($\gamma_2-L_2$) with $0.4 \leq 1 - \alpha_{eff} < 0.6$; and the $M_4$ frequency band (MN4–SK4) mostly presenting adequate CI coverage. Conversely, for WN series, the RT_Tide CI intervals have $\alpha_{eff} \approx \alpha$ for most of the constituents since the method adequately assesses the constituent standard errors for uncorrelated noise regressions (Fig. 7c).

To confirm that RT_Tide benefits of the IRLS weights as an estimate of the $(\Omega^T\Omega)^{-1}$ diagonal, note that RT_Tide largely overestimates the CI coverage for both WN and RN series also when considering OLS regressions (see appendix B).

These findings are particularly relevant since some previous studies used the CI width to justify the use of the IRLS in tidal HA regressions or to evaluate the optimal weighting function and tuning parameter (Leffler and Jay 2009; Codiga 2011; Matte et al. 2014b). In this regard, further investigations should reevaluate the impacts of the IRLS optimization parameters when the residual and CI calculations correctly use the estimated weights. Such optimization parameters also affect the convergence properties of the IRLS estimators.

For most constituents and both RN and WN series, SPB and MBB have $\alpha_{eff} \approx \alpha$ (Figs. 7e,f), besides SPB produced CIs with adequate coverage for fewer constituents, especially for WN simulations (Fig. 7e). These results are particularly interesting since they differ from what may be expected from parametric and nonparametric resampling. While bootstrap methods typically overestimate the CI coverage when the resampling procedure does not account for data serial correlation adequately (Zwiers and von Storch 1995), in our application, the MBB shows coverage close to the nominal confidence level for most of the constituents, and the SPB shows better performances for RN than for WN series.

Moreover, at both St.1 and St.2, some specific constituents presented rather low CI coverage for RN series and all considered methods. At St.1, $1 - \alpha_{eff}$ was lower than 0.5 for a group of semidiurnal constituents around $M_2$ (Fig. 7). At St.2, $M_8, L_2, 2N_2$, and $OQ_2$ had $1 - \alpha_{eff} \leq 0.5$ and $(1 - \alpha_{eff}) \leq 0.37$ for $M_2$ CIs, even for RN series longer than 1 yr (Fig. S5). In these cases, a large proportion of the point $A_k$ estimates differ substantially from the $A_k$ parameter used for simulating the fixed start $h^*$ series in the MC experiment (see Fig. S9 for an example). For this reason, the corresponding CIs do not recover the true amplitude values, independently of the method considered for estimating $\sigma_{Ak}$. One possible explanation is that, while some of the methods (partly) account for the residual heteroscedasticity and autocorrelation for estimating $\Sigma_P$, the IRLS make little use of this information to estimate $\beta_{IRLS}$—Figure B1d (appendix B) corroborates this conclusion showing similar results for OLS regressions. Hence, some important uncertainties affect the point estimates of the regression coefficients, even for constituents associated with large tidal energy and relatively long records (e.g., about 3 years). For some constituents (e.g., $M_2$) such uncertainty may relate to the nodal cycle which is poorly resolved in the 11-yr fit and flow into the residuals used in the MC simulations. Coping with this issue would involve integrating the residual variance–covariance in the estimator of the regression coefficients. Specifically, the use of a GLS estimator for Eq. (1) could be an effective way to eliminate this uncertainty affecting the tidal model parameters [see, e.g., appendix B and Christoffersson (1997) and Davidson and MacKinnon (2004)].

Independently from the considered regression model, the true CI coverage $(1 - \alpha)$ is typically a function of the sample size since more observations should lead to more accurate point
FIG. 7. Probability coverage of the amplitude CIs at St.1: percentage of annual fixed start h’ simulations for which the true $A_k$ value is included in the CI constructed for the $k$th constituent at the 0.95 confidence level for (a) UTide, (b) robustfit.m, (c) RT_Tide, (d) NS_Tide, (e) MBB, and (f) SPB.
and interval parameter estimations. To evaluate the change of the six method performance in assessing the HA parameter uncertainty when increasing the number of observations available, the effective coverage of $A_k$ CIs was estimated for $h^*$ simulations of different lengths and nominal confidence levels. Accordingly, Fig. 8 displays the fraction (%) of the 68 tidal constituents with effective CI coverage equal and higher than the nominal $(1 - \alpha_{eff})$ level (solid and shaded bar, respectively), for each method and 6-month, 1-yr, and 3-yr series.

As expected, the difference between the nominal and effective CI coverage tend to decrease when increasing the series length. Interestingly, however, while robustfit.m coverage recovered the nominal $(1 - \alpha)$ level for 20%–50% of the constituents for the WN simulations (Fig. 8a), the bootstraps produced CIs with adequate coverage for a higher proportion of tidal constituents in the presence of autocorrelation (i.e., for RN series; Fig. 8b). Similarly to robustfit.m, RT_Tide exceeded the nominal CI coverage for most of the $A_k$ for the RN series, with only few constituents CI presenting $\alpha_{eff} = \alpha$. UTide and NS_Tide never reached the adequate coverage for WN and only few constituents have $\alpha_{eff} \geq \alpha$ for 3-yr RN simulations. This confirms that, while the amplitude standard
errors are better assessed by the robustfit.m and RT_Tide methods for WN series, in the presence of autocorrelated noise, MBB produces better uncertainty estimates. For methods and constituents that do not reach the nominal CI coverage, a component-wise analysis of the β significance would fail in identifying the relevant tidal constituents.

c. NS_Tide correction

Figure 9 compares the MBB uncertainty assessment with the corresponding NS_Tide statistics when avoiding using the IRLS weights for computing the regression variance–covariance [i.e., by replacing \( F_{\text{IRLS}} \) and \( P_{\text{IRLS}} \) with \( F_1 \) and \( P_1 \) in Eq. (8)]. Equation (A3) in appendix A provides the exact NS_Tide unweighted variance–covariance formulation. This correction distinctly rectifies the NS_Tide underestimation and improves the method performance. Specifically, NS_Tide unweighted and MBB become approximately equivalent at St.1 and St.2 in terms of the estimated \( \sigma_{\beta_1} \) values for RN series (Fig. 9a) since both methods respect the colored noise assumption of the LS regression. Accordingly, NS_Tide unweighted shows coverage close to the nominal confidence level for most constituents at St.2 and for 3-yr series at St.1 (Fig. 9b) allowing to better assess the constituent significance. For the same reason NS_Tide performs well when applied to OLS regressions (see, e.g., Fig. B1d), for which NS_Tide and NS_Tide unweighted coincide. This confirms that Eq. (8) can efficiently account for the LS hypothesis of colored noise needed for tidal HA applications once corrected for unweighted residuals. In these circumstances, using a GLS estimator of the regression coefficients could further improve both NS_Tide and MBB by explicitly integrating the residual variance–covariance matrix in the \( \hat{\beta}_{\text{GLS}} \) estimator (see appendix A).

8. Summary and discussion

The present study uses resampling methods to illustrate some limitations in the uncertainty assessments produced by four statistical tools for tidal analysis: the RT_Tide, UTide, and NS_Tide packages and the robustfit.m MATLAB function applied to harmonic regressions. The study focuses on the default RT_Tide colored bootstrap, UTide Monte Carlo colored, and NS_Tide correlated methods for IRLS regressions since the other package options conflict with the least squares framework assumptions. Although the hypothesis of uncorrelated homoscedastic errors is unrealistic for most tidal
records and leads to suboptimal regression estimators, complementary results were also produced for OLS regressions to further support our conclusions. OLS regression is the basis of most of the historical literature on tidal analysis.

The first study objective was to evaluate the theoretical differences among the methods by reviewing their statistical properties. Rewriting each method variance–covariance formula with a common matricial notation, it has been shown that none of the approaches adequately reproduces the variance of the IRLS estimator. Specifically, RT_Tide neglects the regressors and IRLS weight quadratic product in the coefficient variance–covariance, and inherits from T_Tide some inconsistencies in the residual PSD estimation (e.g., the averaging over fixed frequency bands and removal of some spectral lines). Robustfit.m implicitly uses the OLS hypothesis of uncorrelated residuals to compute the IRLS variance–covariance, while UTide down-weights the IRLS variance–covariance diagonal with the weighted residual PSD (computed as in RT_Tide). Similarly, NS_Tide improperly weights the residuals and regressors before the Fourier decomposition of the coefficient variance–covariance.

The second objective was to examine whether the considered methods correctly evaluate the uncertainty associated with the tidal parameter estimations. To this end, a Monte Carlo (MC) experiment was used to simulate a large number of water level series with known statistical properties (i.e., known harmonic parameters and noise structure). For each of these simulations, the harmonic regression coefficients were obtained with the UTide code, adapted to treat Eq. (1) and with default IRLS weighting function and tuning parameter. Then, the statistical distributions of the parameters were approximated via the analytical RT_Tide, UTide, robustfit.m, and NS_Tide default formulations and compared with those estimated by two residual resamplings: a moving block bootstrap with fixed block length and a semiparametric bootstrap that simulates noise realizations from the fast Fourier transform of the observed residuals. The analysis was conducted in a stationary framework, as all methods assume that the statistical properties of the observed signals (e.g., the HA tidal constituent parameters, their standard errors, and the residual spectrum characteristics) do not vary significantly with time.

Based on the presented results, robustfit.m and RT_Tide provide well-constrained characterizations of the parameter uncertainty when the error terms of the HA model are generated by a white noise process. The close correspondence between the two methods and the OLS variance–covariance expressions explain this result, while these analytical approximations show poor performance for series with autocorrelated residual. In these case, RT_Tide generally overestimates the parameter variance since it ignores the effect of the regressors and assumes zero covariances between the coefficients related to different constituents.

Equivalently, robustfit.m underestimates the variability for low-frequency constituents and overestimates the uncertainty at the highest frequencies when dealing with signals perturbed by colored noise, as it does not adequately represent the residual autocorrelation structure.

UTide and NS_Tide oversmooth the regression variance–covariance matrix, resulting in underestimated parameter variances. The incorrect use of weighted residuals in both methods and the band averaging and line decimation of the residual spectrum in UTide explain these results. Hence, it was possible to define a corrected version of NS_Tide by eliminating the residual down-weighting from Matte et al.’s (2013) formulation.

The NS_Tide unweighted formulation provided the most realistic estimates of the tidal amplitude variability among the considered analytical methods. This suggests that the NS_Tide unweighted approach could provide reliable error variance–covariance structures for GLS regressions for water level series collected at evenly spaced time steps. However, the method does not currently integrate the error variance–covariance matrix into the regression coefficient estimator and needs some adjustments for unevenly spaced series (e.g., regridding and/or interpolation of residuals).

The two residual bootstraps lead to realistic assessments of the parameter uncertainty by accounting for the colored nature of noise signals typically found in tidal records. Notably, the MBB avoids model selection and conveys a complete and skillful description of the tidal parameter probability distributions. However, the results for SPB are not entirely conclusive. On the one hand, the method reproduces the expected $\sigma_{A_k}$ trend over the frequency spectrum in the presence of autocorrelated residuals. On the other hand, it tends to underestimate the absolute value of the amplitude standard errors for signals perturbed by white noises. In this case, the SPB performance also varied with the sampling conditions (series length, as well as the presence and distribution of missing values). This result might be to some extent related to the use of the FFT to estimate the residual spectrum since Fourier-based methods tend to filter part of the skewness and kurtosis of the analyzed signals. This may result in weakening the randomness of the simulated residual samples (Kim et al. 2013). Other nonparametric estimators may prove better for generating realizations of noise with unknown spectrum (Munteanu et al. 2016).

Such differences among the methods are of concern for the applications that directly involve the specification of tidal parameter probability distributions. For instance, an accurate characterization of tidal parameter standard error is needed to predict tides into future climate and evaluate sea level trends (Ray and Foster 2016; Li et al. 2021; Talke et al. 2018). Similarly, accounting for tidal reconstruction accuracy is essential to support the analysis of short and sparsely sampled gauge records and satellite measurements (Cherniawsky et al. 2001; Stammer and Cazenave 2017).

The last study objective was to determine the adequacy of constituent selection strategies based on the proposed uncertainty assessments. To this end, the differences between the standard errors estimated by the six methods were evaluated by comparing the probability coverage of the corresponding tidal amplitude confidence intervals (CIs). When the CI coverage does not correspond to the nominal confidence level, the tidal constituent significance cannot be adequately assessed. In all cases, standard errors smaller than those estimated via MBB correspond to CIs that do not respect the nominal confidence level, as the range of covered values is too narrow.
Hence, the reduction of the CI width implies the underestimation of the tidal parameter uncertainty. Conversely, standard errors larger than the MBB values lead to slightly better CIs when considering series perturbed by white noise but too wide intervals for the colored noise case. Hence, while the MBB produces better uncertainty assessments than the other methods for series characterized by autocorrelated residuals, it slightly underestimates the overall residual energy and parameter variance. In this context, CI definitions that do not presume the Gaussianity of the tidal amplitude may perform better. Finally, some specific constituents presented low CI coverage for all the considered methods due to a substantial discrepancy between the point $A_k$ values (estimated via IRLS) and the true amplitude parameter. These differences may arise from temporal modulations of the tidal parameters that are not accounted for by the considered nodal corrections (not shown) and convey into the residuals as all other unresolved components. For instance, at the estuarine location (St.2), the bottom friction can also significantly modify the nodal modulations described by the equilibrium satellite constants (Ray and Talke 2019; Godin 1986).

In view of these findings, the MBB demonstrates the ability of calculating a wide range of statistics related to the parameter statistical distributions and provides valuable information on the parameter uncertainty without using stringent parametric hypothesis. The MBB can also be applied to different noise structures with few modifications in the resampling strategy (e.g., changes in the block length). This is particularly interesting in the presence of strong autocorrelation of residuals, a situation in which the statistical inference from the other considered methods is not robust and cannot be safely applied for testing the constituent significance. The effects of possibly erroneous constituent selections on derived statistics (e.g., estimated trends for sea levels, their dependence on tidal constituent changes, and the corresponding changes in coastal flooding risk; Pirazzoli et al. 2006; Ray and Foster 2016; Schindelegger et al. 2018; Li et al. 2021; Enríquez et al. 2022) may be important and should be the focus of future research.

**Limitations and future work**

Correlated regression residuals may result from misspecified regression models which neglect some signal variability components. In these cases, the OLS and IRLS regression coefficient estimates are not unrealistic on average (the $\hat{\beta}$ estimator is still unbiased). However, OLS and IRLS regressions produce biased estimators of the regression coefficient variances. To benefit of the optimal properties of least squares estimators for estimating the $\beta$ values, the autocorrelation of the observed signal should be explicitly accounted for in the regression model. Hence, future studies should develop GLS frameworks for tidal HA. Even more important, future work must explore the possibility of using some externally estimated GLS variance–covariance to complete and support the analysis of sparsely sampled data. Such improved information would help at resolving the centimeter-scale variability of some tidal components from satellite measurements, such as those from the upcoming Surface Water and Ocean Topography (SWOT) mission and forthcoming missions (e.g., Morrow et al. 2019; Pietroniro et al. 2019).

Moreover, the ability of residual bootstraps to fully replicate the temporal covariance structure of sparsely sampled water level records should be further validated (Vogel and Shalsercross 1996). Finally, additional research is needed to improve the spectrum-based residual simulations and better reproducing the heavy-tailed nature of noise typically found in tidal records. Resampling methods suitable for analyzing extreme events or strongly nonstationary residuals could then be defined. At this stage, before applying the resampling proposed in the present study, it is important to verify the specification of the HA model, evaluating the inclusion of long-term trends and/or nonstationary components in Eq. (1). Considering these limitations and our general results, a more comprehensive assessment of the properties of the proposed bootstraps for irregularly sampled series and for nonstationary tidal–fluvial models [e.g., NS_Tide or continuous wavelet transforms (CWTs)] is recommended.

**Acknowledgments.** The presented research widely relied on the use of the existing statistical packages for tidal analysis. Specifically, the study would not have been possible without the UTide, T_Tide, RT_Tide, and NS_Tide codes being made available. Comments by David A. Jay (Portland State University) and Edward D. Zaron (Oregon State University) on earlier drafts of the paper are gratefully appreciated. The authors gratefully thank Dr. Richard D. Ray and the two anonymous reviewers for the constructive comments and recommendations which definitely helped to improve the paper.

**Data availability statement.** Hourly water level records at the two gauge stations considered in the analysis come from the Canadian Tides and Water Levels Data Archive of Fisheries and Oceans Canada (http://www.isdm-gdsi.gc.ca/). The source code used to produce the results from this paper is published as “Innocenti, Silvia 2022: boot_tide: residual bootstrap methods for parameter uncertainty assessment in tidal analysis with temporally correlated noise. https://doi.org/10.5281/zenodo.7085587” and can be accessed from the https://github.com/SilvInn/boot_tide GitHub repository.

**APPENDIX A**

**GLS in NS_Tide**

A GLS regression is used in the presence of heteroscedastic autocorrelated errors. In this case, $\epsilon$ is a zero-mean random variable with (nondiagonal) variance $\Sigma_\epsilon = \sigma^2 \mathbf{V}$, where $\Sigma_\epsilon$ and $\mathbf{V}$ are positive definite squared matrices. As a result, it is always possible to factorize these matrices using $\mathbf{V} = \mathbf{G}\mathbf{G}^T$, where $\mathbf{G}$ is a square matrix of size $T$. Accordingly, the GLS problem can be rewritten as an OLS model.
FIG. B1. HA uncertainty estimation for OLS tidal regressions over 500 annual fixed start \( h^* \) series at St.1: (a), (b) median of the amplitude standard errors estimated for each tidal component (x axis), (c) distribution of the differences between the OLS and IRLS amplitude estimates, (d) probability coverage of the amplitude CIs constructed at the 0.95 confidence level. T_Tide, UTide, and NS_Tide consider the options presented in the left panel of Fig. 1b (excluding the linear methods).
with i.i.d. errors \( \hat{\epsilon} \), by simply premultiplying each term by \( \mathbf{G}^{-T} \), i.e.,

\[
\mathbf{h} = \mathbf{G}^{-T} \mathbf{\hat{\epsilon}}, \quad \text{with } \Sigma_{\hat{\epsilon}} = \mathbf{G}^{-1} \Sigma_{\epsilon} \mathbf{G}^{-T} = \sigma_{\epsilon}^2 \mathbf{I},
\]

(A1)

where \( \hat{\mathbf{h}} = \mathbf{G}^{-T} \mathbf{h}, \mathbf{\hat{\Omega}} = \mathbf{G}^{-1} \mathbf{\Omega}, \) and \( \hat{\epsilon} = \mathbf{G}^{-1} \epsilon \). Using the regular OLS expression, the GLS variance-covariance of \( \hat{\mathbf{\beta}}_{\text{GLS}} \) is written as

\[
\Sigma_{\hat{\mathbf{\beta}}_{\text{GLS}}} = \sigma_{\hat{\epsilon}}^2 \mathbf{\Omega}^{-1} \mathbf{\Omega}^{-T} = \sigma_{\epsilon}^2 \mathbf{\Omega}^{-1} \mathbf{\Omega}^{-T} = \mathbf{\Omega}^{-1} \Sigma_{\epsilon} \mathbf{\Omega}^{-1}.
\]

(A2)

Note that, for \( \mathbf{V} = \mathbf{W}^{-1} \), the error variance-covariance is \( \Sigma_{\epsilon} = \sigma_{\epsilon}^2 \mathbf{W}^{-1} \) and the GLS model in Eq. (A2) is equivalent to WLS. Finally, consider the Fourier transforms \( \mathbf{F}_{\Omega} \) and \( \mathbf{F}_\epsilon \) of the GLS regressors and errors. These transforms can be expressed, respectively, as \( \mathbf{F}_{\Omega} = \mathbf{G}^{-T} \mathbf{\Omega} = \mathbf{\Omega} \) and \( \mathbf{F}_\epsilon = \mathbf{G}^{-1} \epsilon = \hat{\epsilon} \). Hence, based on Eq. (A2), the \( \text{NS}_{\text{Tide}} \) variance-covariance matrix \( \Sigma_{\hat{\mathbf{\beta}}_{\text{NS}_{\text{Tide}}}} \) can be recovered as

\[
\Sigma_{\hat{\mathbf{\beta}}_{\text{NS}_{\text{Tide}}}} = \sigma_{\hat{\epsilon}}^2 \mathbf{F}_{\Omega} \mathbf{F}_{\Omega}^{-1} = \mathbf{F}_{\Omega}^{-1} \Sigma_{\epsilon} \mathbf{F}_{\Omega}^{-1} \mathbf{F}_{\epsilon}^{-1} \mathbf{F}_{\epsilon}^{-1},
\]

(A3)

where \( \mathbf{F}_{\epsilon} = \mathbf{F}_{\epsilon}^{-1} \mathbf{F}_{\epsilon}^{-T} \mathbf{F}_{\epsilon}^{-1} \). Note that, to respect the GLS formulation of \( \Sigma_{\hat{\mathbf{\beta}}_{\text{GLS}}} \) [Eq. (A2)] the unweighted residuals and observations must be used in Eq. (A3).

APPENDIX B

HA Uncertainty Estimates for OLS Regressions

Figure B1 compares the six uncertainty estimation methods for OLS regressions (i.e., \( \hat{\mathbf{\beta}}_{\text{OLS}} \) regression coefficient estimates) applied to the 500 annual fixed start \( \text{h}_t \) series at St.I. In this figure, \( \text{T}_{\text{Tide}} \) is presented as OLS version of RT_Tide.

While the general IRLS conclusions also apply to the OLS case, this example expressly confirms the following results: (i) the considered tidal analysis packages use the IRLS weights to reduce the amplitude standard errors and CI width improperly. While, by definition, the OLS should provide the minimum residuals, \( \text{T}_{\text{Tide}}, \text{UTide}, \text{and NS}_{\text{Tide}} \) present OLS \( \sigma_{A_k} \) values higher than the IRLS values (compare, for instance, the \( \sigma_{A_k} \) value ranges in Figs. 4 and B1a,b). (ii) The absence of the IRLS weights, on the other hand, increases the \( \sigma_{A_k} \) \( \text{T}_{\text{Tide}} \) overestimation since \( \mathbf{W} \) is typically expected to be inversely proportional to the diagonal of the regressor quadratic term \( \mathbf{\Omega}^{-1} \mathbf{\Omega}^{-T} \) needed in \( \Sigma_{\hat{\mathbf{\beta}}} \). On the other hand, the OLS condition \( \mathbf{W} = \mathbf{I} \) improves the UTide and NS_Tide performance, with \( \sigma_{A_k} \) values closer to the MBB standard errors and better CI coverage (Figs. B1a,b,d); in fact, note that for the OLS NS_Tide corresponds to \( \text{NS}_{\text{Tide}} \) unweighted presented in section 7c, (iii) robustfit.m disregards the WLS formulation of the IRLS regression when estimating \( \sigma_{A_k} \). Specifically, the robustfit.m uncertainty estimates are weakly sensitive to the computed weights since \( \Sigma_{\hat{\mathbf{\beta}}_{\text{GLS}}} \) [Eq. (2)] and \( \Sigma_{\hat{\mathbf{\beta}}_{\text{robustfit.m}}} \) [Eq. (4)] differ only on the residual scalar variance estimation. For the same reason, the method shows good performance for both IRLS and OLS regressions when considering the WN series (see Figs. 7b and B1d). (iv) No significant change is observed between the OLS and IRLS MBB performance, which is consistent with the small differences between the OLS and IRLS tidal amplitudes (between \( -1 \) and \( 1.5 \) mm for most of the \( \text{h}_t \) series and tidal constituents; Fig. B1c). (v) Some inconsistent results are found for the SPB, which produces very low \( \sigma_{A_k} \) and CI coverage values for some frequency bands (for constituents with period lower than 1 yr and for diurnal and semi-diurnal constituents when considering the WN series).

REFERENCES


