https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

Robust and Flexible Tidal Reconstruction from Sparse High Water - Low Water Observations

Joris (J.G.W.) Beemster[1], Pascal Matte[2], Silvia Innocenti[2], Bas (D.S.) Van Maren[3,4], and Ton (A.J.F.) Hoitink[1]

[1]Department of Environmental Sciences, Wageningen University & Research, Wageningen, The Netherlands
[2]Numerical Prediction and Modelling Research Division, Environment and Climate Change Canada, Quebec, QC, Canada
[3]Ecosystems and Sediment Dynamics, Deltares, Delft, The Netherlands
[4]Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands

Correspondence: Joris (J.G.W.) Beemster (joris.beemster@wur.nl)

Abstract. Tidal analysis and prediction are traditionally based on the harmonic decomposition of continuous water-level records. This limits the applicability to sparse, historical observations of high and low waters. Here, we adopt a high–low tidal analysis (HLTA) framework that directly models tidal extrema and their temporal modulation using lunar transit timing and astronomical forcing. Two formulations are explored: a long-period harmonic (LPH) approach and an empirical–astronomical (EA) representation. Application to tide-gauge data from the Western Scheldt demonstrates that HLTA predicts tidal extrema with accuracy comparable to harmonic analysis of 10-minute observations for water levels. Performance is also largely improved for the prediction of extrema timing, and bias is reduced. In contrast, harmonic analysis applied directly to high–low data performs poorly, not only due to aliasing, but also because of broad-scale dependencies between constituents introduced by sparse sampling. The HLTA framework is robust to observational errors and can be extended naturally to non-stationary conditions by incorporating, for example, river discharge. Coupled with simple interpolation, HLTA enables accurate reconstruction of the continuous tidal signal, matching or exceeding harmonic analysis on high-resolution data in shallow systems where the tidal wave is strongly distorted. These results demonstrate that accurate tidal reconstruction from high–low observations is feasible even in strongly distorted, shallow systems, with performance comparable to modern high-resolution analyses. This enables improved use of historical datasets for applications such as storm surge analysis, sea-level rise, and the analysis of changing tides, while also suggesting potential for improved modern tidal prediction in shallow and non-linear environments.

1 Introduction

Tides have been observed for millennia, but systematic measurement of coastal water levels developed only in the 18th to 20th centuries with the advent of mechanical and self-registering tide gauges (Cartwright, 2000). Although modern digital gauges provide continuous high-frequency records at minute-scale resolution, much of the world's historical tidal archive consists only of the times and heights of successive high and low waters (Talke and Jay, 2017; Latapy et al., 2023). These high water – low water (HW/LW) records are often incomplete: in some cases, only daytime observations were registered, in others only high waters were recorded, and in some instances only the timing or only the height of each extremum is available. Nevertheless,

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

such records are widespread, often extending back a century or more, and in many regions represent the longest available observations of coastal water levels. Properly analyzed, HW/LW datasets have the potential to provide valuable information on past tidal conditions, long-term changes in tidal regimes, surges and mean sea level, and the interaction of tides with river discharge, surges, and other non-tidal drivers. However, their use has been limited by the lack of robust methods for extracting tidal information from sparse and irregular extrema observations.

Tidal analysis and prediction are traditionally based on the harmonic decomposition of continuous water-level records, where the tide is represented as a sum of sinusoidal constituents at known astronomical frequencies (Pugh et al., 2014). Although classical harmonic analysis (HA) can accommodate irregular sampling, it implicitly treats observations as independent measurements of a deterministic tidal signal perturbed by random noise. HW/LW observations violate the prerequisites of this framework because sampling times are generated by the extrema of the tidal signal itself. This phase-locked, sparse, and endogenous sampling reduces the effective information content and enhances aliasing between constituents (Ji and Guohong, 1987). Consequently, applying standard HA directly to HW/LW data often yields ill-conditioned or physically ambiguous solutions, even when derivative constraints (Foreman and Henry, 1979) are imposed. These limitations motivate the development of approaches that operate directly on the tidal extrema rather than reconstructing the full continuous tidal curve from sparse observations.

A key observation is that, although HW/LW data are sparse in time, they can be reorganized into slowly varying time series when expressed relative to astronomical reference times, such as lunar transits. This idea builds on earlier extrema-based formulations developed by Lubbock (1835) and later formalized in harmonic form by Horn (1948). In this representation, the fast semidiurnal oscillation is removed, and the remaining variability reflects the slow modulation of the tidal envelope by astronomical forcing. This transformation concentrates the information content of HW/LW records into a small number of smoothly varying signals, which are more robust to sampling irregularities and observational uncertainty.

Building on this idea, we revisit and extend a Lunar-Transit High–Low Tidal Analysis (HLTA) approach that models the timing and height of the tidal extrema directly. Our approach combines and modernizes earlier extrema-based approaches by representing high- and low-water quantities as smooth functions of astronomical forcing. Two complementary formulations are presented: (i) a long-period harmonic (LPH) representation, in which extrema are described using a compact set of long-period tidal constituents (adapted from Horn (1948)), and (ii) an empirical–astronomical (EA) representation, in which extrema are related directly to physically meaningful astronomical forcing parameters (adapted from Lubbock (1835)). Both formulations can accommodate incomplete and irregular sampling (e.g., HW-only or daytime-only observations). Relative to earlier implementations, the HLTA approach is further modernized through robust parameter estimation using IRLS, reformulation of the extrema series into mean and deviation components to improve numerical conditioning, and extensions for incorporating non-stationary forcings and reconstructing continuous water-level time series from predicted extrema.

The aim of this study is to evaluate the extent to which tidal water levels can be recovered from HW/LW observations using these extrema-based approaches. Motivated by the need to more quantitatively exploit historical records that are often limited to tidal extrema, we assess the performance of these methods using high-resolution and high–low water-level data from two contrasting sites. We examine how well the approaches reproduce the modulation of tidal extrema and how robust they are

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

to noise and non-stationary conditions. In addition, we evaluate their ability to reconstruct continuous water-level time series from predicted extrema using a range of interpolation schemes, and compare performance against classical harmonic analysis applied to both high-resolution and HW/LW data, thereby also exploring their relevance for modern applications including tidal prediction.

The remainder of the paper is organized as follows. We first review existing approaches to tidal analysis from high–low water data (Section 2) and discuss the limitations of classical harmonic analysis for HW/LW observations (Section 3). We then present the Lunar-Transit High–Low Tidal Analysis approach (Section 4) and describe the datasets and experimental design (Section 5). The results are presented in Section 6, followed by a discussion and conclusions in Sections 7 and 8.

2 Tidal Analysis Methods for HW/LW Data

2.1 Early empirical and astronomical approaches

Systematic attempts to predict tides long predate modern harmonic analysis, and most early methods were built directly on observations of high and low waters. In the 11th century, practical prediction schemes were based on the empirical relationship between the upper moon's transit (its meridian passage above the horizon) and the timing of the next high water (Moule, 1923; Zuosheng et al., 1989; Lubbock, 1837). These approaches typically relied on the "lunar establishment" (or lunitidal interval), defined as the observed time of high water at a new or full moon, combined with an assumed constant daily offset. Although simple, such methods were sufficient to produce rudimentary tide tables and remained in use for centuries. However, by the 18th and 19th centuries, it became clear that the daily offset was neither constant nor universal. This led to more systematic descriptions of tidal variability, culminating in the work of Lubbock (1835) and his contemporaries (e.g. Cassini, 1710), who analyzed the timing and height of high and low waters relative to lunar transit and expressed their variations in terms of slowly varying astronomical parameters such as lunar age, declination, and distance, as well as solar declination (Cassini, 1710; Lubbock, 1835). These formulations represent an early form of empirical–astronomical tidal modeling, in which tidal extrema are described directly as functions of astronomical forcing.

2.2 Harmonic analysis and its extension to HW/LW data

The introduction of harmonic analysis in the late nineteenth century marked a major conceptual shift in tidal analysis, recasting the tidal signal as a sum of harmonic constituents at known astronomical frequencies (Thomson, 1868; Darwin, 1883). Within this framework, tidal prediction became a problem of estimating constituent amplitudes and phases from observations of water level. Following the development of harmonic methods, several researchers sought to adapt them to HW/LW observations. Procedures were developed to extract dominant constituents from extrema records using empirical formulations and stepwise reductions (Ferrel, 1874; Darwin, 1891; Schureman, 1941; Doodson, 1951; Zetler, 1953). These approaches attempted to recover the harmonic structure of the tide from sparse observations, but were limited by the reduced information content of HW/LW data.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

Horn (1948) proposed an intermediate formulation that provides a conceptual bridge between empirical–astronomical and harmonic approaches. In this framework, high- and low-water heights and time lags relative to lunar transit are represented as sums of long-period harmonic constituents. By focusing on tidal extrema rather than the full water-level signal, this approach effectively removes the dominant semidiurnal oscillation and expresses the remaining variability as slowly varying modulations driven by astronomical forcing. This approach has been applied in a range of settings, including long-term analyses of the tidal extrema in the St. Lawrence estuary (Godin et al., 1967), and remains in operational use along the German North Sea coast, where it is commonly referred to as a harmonic representation of inequalities (HRoI, e.g. Boesch and Müller-Navarra, 2019). Despite its continued operational use, this class of methods has received relatively limited attention in the recent scientific literature.

2.3 Transition to modern harmonic analysis

The advent of least-squares harmonic analysis in the mid-twentieth century transformed the field by making it practical to estimate large numbers of tidal constituents from continuous water-level records. As digital tide gauges became more widespread, these methods quickly became the standard approach for tidal analysis and prediction, reducing the central role of HW/LW-specific approaches. Nevertheless, several studies continued to investigate the extraction of tidal information from HW/LW data within the harmonic framework (Foreman and Henry, 1979; Ji and Guohong, 1987; Li et al., 2004). A common strategy in such approaches is the use of derivative constraints, imposing the requirement of zero slope at the observed extrema to augment the information content of the HW/LW time series.

More recent developments in harmonic analysis, including Bayesian regularization (Sarkar et al., 2018) and smoothness constraints based on admittance principles (Munk and Cartwright, 1966; Pan et al., 2023; Monahan et al., 2025), have further improved the stability of constituent estimation when data are limited. However, these methods remain fundamentally tied to the reconstruction of a continuous tidal signal from discrete observations.

2.4 Remaining challenges for HW/LW data

Despite these advances, analyzing HW/LW data remains a challenge. Classical harmonic analysis relies on conditions that are difficult to satisfy for extrema-based observations. In particular, HW/LW records are sparse and phase-locked to the tidal cycle, so that the timing of observations depends on the signal itself. Altering the sampling scheme, this changes the fundamental statistical properties of classical spectrum and harmonic analysis estimators and complicates the identification of individual tidal constituents. As a result, the application of harmonic methods to HW/LW data is subject to aliasing, parameter dependence, and increased sensitivity to observational uncertainty. The limited number of observations per tidal cycle limits the ability to resolve constituents with closely spaced frequencies, while the phase-locked sampling and structured noise characteristics invalidate several assumptions underlying harmonic analysis. Even when augmented with derivative constraints, these approaches attempt to infer the full harmonic structure of the tide from a limited and structurally constrained set of observations.

Early empirical–astronomical formulations and Horn's long-period harmonic representation provide an alternative perspective, in which the information contained in HW/LW records is expressed directly in terms of slowly varying modulations of

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

the heights and timing of tidal extremes. This viewpoint shifts the focus from reconstructing the full tidal curve to modeling the temporal evolution of the tidal extrema themselves. However, these approaches have received relatively limited attention in the broader scientific literature over the past half-century.

In the following section, we examine in more detail why classical harmonic analysis performs poorly when applied to HW/LW data and how the structure of extrema-based sampling imposes fundamental limitations in resolving tidal constituents.

3 Limitations of harmonic analysis

Classical harmonic analysis (HA) represents the tidal signal as a sum of a finite set of deterministic sinusoidal constituents, whose amplitudes and phases can be estimated by a least squares fitting procedure. This formulation relies on several key assumptions: (1) the observed signal can be represented by a finite set of constituents; (2) these constituents are independent and non-stochastic; (3) the water level signal is stationary over the analysis period; (4) the record is sufficiently long to distinguish closely spaced frequencies to satisfy the Rayleigh criterion, and (5) no other significant source of variability act on the water signal so that the HA error term can be represented as a white noise process. When only HW/LW data are available, several of these conditions, particularly (2), (4), and (5), are not met. These limitations are examined in the following subsections, focusing on sampling structure and aliasing (Section 3.1), sampling–signal dependence and parameter coupling (Section 3.2), and sensitivity to noise and non-tidal variability (Section 3.3).

3.1 Sampling structure and aliasing

As only four observations are available per lunar day and these observations are phase-locked to the tidal cycle, rather than uniformly distributed in time, the first challenge in HW/LW analysis is the drastic reduction in effective sampling. This situation is often described in terms of a Nyquist frequency, but the principal difficulty is not the Nyquist limit per se; rather it is that phase-locked sampling causes persistent aliasing and can prevent constituent separation under the Rayleigh criterion. As demonstrated by satellite altimetry, dominant tidal constituents can be resolved from much sparser but regular sampling schemes with repeat periods of 10 to 30 days (Parke et al., 1987), provided that the record length is sufficient to separate aliased frequencies according to the Rayleigh criterion. In contrast, if the sampling interval coincides with sums or differences of tidal frequencies (as in HW/LW records, or sun–synchronous satellite missions), the minimum frequency distance that satisfies the Rayleigh criterion becomes infinite.

For idealized and perfectly regular sampling at six lunar-hour intervals, the classical results show the formation of exact alias pairs whose frequency sums or differences coincide with the reciprocal of the sampling interval (Ji and Guohong, 1987). This behavior is illustrated in Figure 1a, where alias pairs appear as perfect correlations between the constituent amplitudes. Well-known examples include the aliasing of M_2 with M_6, and of M_4 with the mean sea level. More generally, long-period constituents correlate with quarter-diurnal constituents, diurnal constituents with third-diurnal constituents, and semidiurnal constituents with both other semidiurnal and sixth-diurnal species (Ji and Guohong, 1987). Deviations from perfectly regular spacing, caused by tidal asymmetry, diurnal inequality, modulation by other constituents, and non-tidal noise, have been argued

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

to improve resolvability by breaking the strict phase alignment responsible for these exact alias pairs (Ji and Guohong, 1987), with larger deviations generally expected to increase separability.

3.2 Sampling–signal dependence and parameter coupling

This interpretation from existing literature is incomplete. As the deviations from exact four-times-per-lunar-day sampling arise from the same tidal dynamics that generate the analyzed signal, they are not random. Therefore, the timing of successive HW and LW events contains a harmonic structure with frequencies corresponding to the tidal constituents and their combinations. As a result, the sampling scheme is spectrally correlated with the tidal signal itself. This dependence alters the classical aliasing structure and introduces dependencies between constituents whose frequencies overlap with those present in the sampling pattern. This is illustrated in Figure 1b, which shows the corresponding bootstrap correlations for true HW/LW sampling. While the classical alias pairs remain visible, their correlations are reduced, and strong dependencies emerge among a much broader set of constituents, including pairs that do not alias under strictly regular sampling. The net effect is that resolvability may improve slightly for a limited number of canonical alias pairs but degrades across a much wider range of constituent combinations. In practice, this manifests itself not only as ambiguity between specific pairs such as M_2–M_6, but also as a more diffuse pattern of parameter coupling.

3.3 Sensitivity to noise and non-tidal variability

Classical HA assumes that non-tidal contributions act as random noise, uncorrelated with the tidal signal. In HW/LW records, this assumption is necessarily violated by the autocorrelation involved in the sampling. Moreover, as non-tidal residuals typically exhibit temporal correlations (Innocenti et al., 2022) and seasonal patterns (e.g. due to storm climatology), and interact with the tidal signal itself (Horsburgh and Wilson, 2007; Jenkins et al., 2025), this residual autocorrelation structure strongly increases the parameter estimation uncertainty with sparse HW/LW observations. In particular, low-frequency variability associated with meteorological forcing or seasonal effects overlaps with long-period constituents. Through the broad parameter coupling inherent in HW/LW harmonic analysis, this variability can further propagate across the spectrum, affecting higher-frequency constituents and reducing parameter identifiability. Differences in the sensitivity of high and low waters to non-tidal perturbations may introduce additional structure into the sampling, further complicating the separation of tidal and non-tidal signals.

3.4 Implications for harmonic analysis of HW/LW data

Taken together, these limitations imply an inherent mismatch between the assumptions of least-squares HA and the structure of HW/LW observations, which limits the validity of classical HA for reconstructing the full harmonic representation of the tide from extrema alone. In the following section, we therefore adopt an alternative perspective in which HW/LW observations are treated not as sparse samples of a continuous signal, but as slowly varying quantities that directly reflect the modulation of tidal amplitude and timing.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

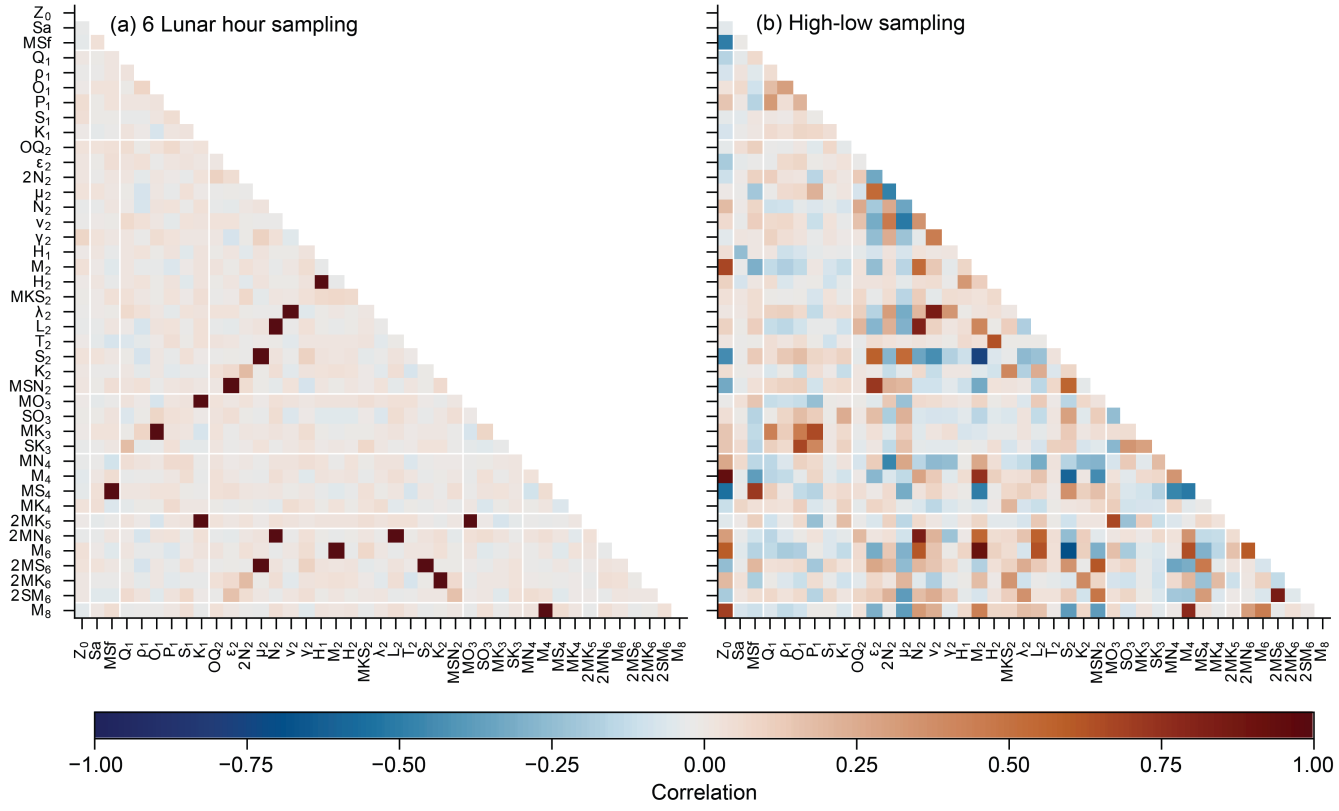

Figure 1. Pairwise correlations between tidal constituents estimated over 1000 bootstrap samples for six-lunar-hour sampling (a) and true high–low sampling (b). Correlations are computed for tidal amplitudes (68-constituent list from Utide) based on 1-month-block bootstrapping of residuals from the Bath tide gauge (The Netherlands) over the period 2015–2019 (Innocenti et al., 2022). Both positive and negative correlations are observed, indicating whether pairs of constituents tend to co-vary (positive) or compensate for one another (negative). Tidal species are delineated by white lines.

4 Lunar-Transit High–Low Tidal Analysis

4.1 Conceptual framework

In semidiurnal tidal regimes, observations of high and low water provide four extrema per lunar day, each characterized by a water level $h_{i,j}$ and a time $t_{i,j}$, where $i \in \{\mathrm{HW}, \mathrm{LW}\}$ distinguishes high and low waters and $j \in \{\mathrm{U}, \mathrm{L}\}$ denotes whether the event follows the upper or lower lunar transit, thereby separating the two semidiurnal tides within each lunar day, which may differ due to diurnal inequality. Following Horn (1948), we replaced absolute event times with time lags between successive lunar transits, which yields sequences $\Delta t_{i,j}(n)$ and $h_{i,j}(n)$ that vary smoothly from one lunar day to the next. These series are primarily governed by the slow astronomical modulation of the tide, rather than the fast semidiurnal oscillation. Because each lunar day contains only two transits, oscillations at diurnal and higher frequencies are absent by construction, and many

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

higher-frequency tidal constituents affect the observed values only via their long-term modulations. In this reparameterization, the HW/LW observations are expressed as a set of slowly varying time series that capture the modulation of tidal amplitude and timing. In this sense, aliasing is not eliminated, but reorganized: constituents that are difficult to distinguish in classical harmonic analysis of sparse data collapse onto shared long-period signals in the extrema representation.

For each transit j, the paired high- and low-water quantities can be decomposed into a mean component and a deviation,

$$\overline{h_j} = \frac{1}{2}(h_{\mathrm{HW},j} + h_{\mathrm{LW},j}), \quad h'_j = \frac{1}{2}(h_{\mathrm{HW},j} - h_{\mathrm{LW},j}), \tag{1}$$

$$\overline{\Delta t_j} = \frac{1}{2}(\Delta t_{\mathrm{HW},j} + \Delta t_{\mathrm{LW},j}), \quad \Delta t'_j = \frac{1}{2}(\Delta t_{\mathrm{HW},j} - \Delta t_{\mathrm{LW},j}). \tag{2}$$

The mean components $\overline{h_j}$ and $\overline{\Delta t_j}$ describe the slowly varying behavior of the water level and the average timing of the tide, including long-period modulation and non-tidal variability. The deviation components h'_j and $\Delta t'_j$ represent the tidal amplitude and the asymmetry of the tidal duration. This formulation is equivalent to the original one based on extrema variables $(h_{i,j},\ \Delta t_{i,j})$ and preserves all information, but reorganizes the estimation problem so that most of the tidal energy is concentrated in a subset of variables (e.g. height deviations). A complete HW/LW record thus provides eight variables per lunar day:

$$\{h_{\mathrm{HW},U}, h_{\mathrm{LW},U}, \Delta t_{\mathrm{HW},U}, \Delta t_{\mathrm{LW},U}, h_{\mathrm{HW},L}, h_{\mathrm{LW},L}, \Delta t_{\mathrm{HW},L}, \Delta t_{\mathrm{LW},L}\},$$

which are reorganized into four means and four deviations:

$$\{\overline{h_U}, h'_U, \overline{\Delta t_U}, \Delta t'_U, \overline{h_L}, h'_L, \overline{\Delta t_L}, \Delta t'_L\},$$

In many settings, long-period modulations act primarily on either the mean or the deviation, allowing the other component to be simplified or omitted, and thereby reducing the number of parameters to estimate. This decomposition also extends easily to non-stationary formulations, in which external forcing, such as river discharge or meteorological surge, can affect the mean water level differently from the tidal amplitude or asymmetry (Hoitink and Jay, 2016).

We consider two complementary approaches for modeling these series of slowly varying extrema:

1. **Long-Period Harmonic High-Low Tidal Analysis (LPH-HLTA)**, which represents the extrema as sums of long-period harmonic constituents, following the structure introduced by Horn (1948).

2. **Empirical–Astronomical High-Low Tidal Analysis (EA-HLTA)**, which relates the extrema directly to slowly varying astronomical variables such as lunar phase, declination and distance, extending the empirical-astronomical formulations of Lubbock (1835).

Both approaches model each of the eight daily extrema series separately and can accommodate incomplete or irregular observations, including HW-only, LW-only, or daytime-only sampling. Additional slowly varying forcings, such as river discharge or storm surge, may be incorporated, where appropriate (Section 4.4). Once the extrema $h_{i,j}$ and $\Delta t_{i,j}$ have been estimated, the corresponding sequence of high and low waters follows directly. The continuous tidal signal can then be reconstructed by interpolating successive extrema, as described in Section 4.5.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

4.2 Long-Period Harmonic High-Low Tidal Analysis (LPH-HLTA)

The LPH-HLTA formulation follows Horn (1948) and represents each extrema as a sum of long-period harmonic terms associated with Doodson numbers 2–6, which define the fundamental astronomical frequencies of the tidal potential. The first Doodson number, corresponding to mean lunar time and associated with the high-frequency tidal oscillation, is omitted. These long-period terms can be interpreted mostly as modulations of tidal elevation and timing, rather than independent components of the full oscillatory tidal signal. To illustrate this, consider the upper-transit deviation term $h'_U(t)$:

$$h'_U(t) = \beta_0 + \sum_{n=1}^{N} \left(\beta_n^c \cos(\sigma_n t) + \beta_n^s \sin(\sigma_n t)\right), \tag{3}$$

where σ_n are the angular frequencies of the selected long-period constituents, N is the number of retained constituents, β_0 is a constant offset, and β_n^c and β_n^s are the cosine and sine coefficients, respectively. The remaining extrema are modeled using equations of the same form, each with its own set of coefficients.

4.3 Empirical Astronomical High-Low Tidal Analysis (EA-HLTA)

In the EA-HLTA formulation, each variable is expressed as a function of slowly varying astronomical quantities. At each lunar transit time t, we define an astronomical state vector

$$S(t) = (\phi_m(t), \phi_s(t), \delta_m(t - \tau_{\delta_m}), R_m(t - \tau_{R_m})), \tag{4}$$

where ϕ_m is the lunar phase, ϕ_s the solar phase, δ_m the lunar declination, and R_m the Earth-Moon distance. In practice, the lags τ_{δ_m} and τ_{R_m} are determined prior to fitting, for instance, by maximizing the correlation with the extrema series. Each extrema variable is subsequently modeled as a function of the state vector. For instance, for the upper-transit deviation term $h'_U(t)$ this can be expressed as:

$$h'_U(t) = \beta_0 + F(S(t)), \tag{5}$$

where $F(\cdot)$ is an empirical response function. In practice, this function is represented using a set of smooth basis functions $\{\psi_k(S)\}$, yielding the regression model

$$h'_U(t) = \beta_0 + \sum_{k=1}^{K} \beta_k \psi_k(S(t)), \tag{6}$$

where β_k are the regression coefficients.

The basis functions consist of linear and quadratic transformations of the S(t) terms. Circular variables (lunar and solar phase) are represented using sine-cosine pairs (e.g., $\sin(2\phi)$, $\cos(2\phi)$, and, optionally, higher harmonics), allowing phase lags to be estimated directly within the regression framework. This yields a compact and physically interpretable relationship between the extrema and the astronomical forcing (section 6.1.3).

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

4.4 Incorporating non-stationarity

Both the LPH-HLTA and EA-HLTA formulations can be extended to account for non-harmonic forcings. Following the philosophy of non-stationary tidal analysis (Matte et al., 2013), the model coefficients are allowed to vary as functions of a time series $H(t)$ representing the temporal evolution of an external forcing, such as river discharge or meteorological surge. For any extrema (consider here $h'_U(t)$), the stationary model can be written as

$$h'_U(t) = \beta_0 + G(t)\boldsymbol{\beta}, \tag{7}$$

where $G(t)$ represents the astronomical predictors (harmonic for LPH-HLTA, empirical for EA-HLTA). In the proposed non-stationary extension, the coefficients are allowed to vary with an external forcing $H(t)$:

$$\beta_k(t) = \alpha_{k0} + \alpha_{k1}H(t) + \alpha_{k2}H(t)^2, \tag{8}$$

where a quadratic dependence on $H(t)$ is used, as in krvavica2025impact. Linear forms can be used by simply omitting the $\alpha_{k2}H(t)^2$ term, with resulting model:

$$h'_U(t) = \beta_0(t) + X(t)\beta(t). \tag{9}$$

where $\beta_0(t)$ represents the externally forced, non-harmonic component of the signal, and $X(t)\beta(t)$ describes the non-stationary tidal response modulated by the external forcing through $H(t)$.

This formulation naturally separates the externally forced, non-harmonic variability from the modulated tidal response, and can be applied consistently within both the LPH-HLTA and EA-HLTA frameworks.

4.5 Interpolation to a continuous tidal signal

The HLTA approach only predicts tidal extrema, i.e. the times and heights of high and low tides, but not a continuous water level series. However, once the heights and timings of the extrema have been estimated, a continuous tidal signal can be obtained by interpolating between consecutive high and low waters. We employ two interpolation methodologies, one based on a cosine function and the other on the characteristic tidal shape.

A simple and commonly used approach is half-cosine interpolation (Zhang et al., 2018). For two neighboring extrema at times t_1 and t_2 with heights h_1 and h_2, the interpolated level at time $t \in [t_1, t_2]$ is given by

$$\hat{h}(t) = \frac{h_1 + h_2}{2} + \frac{h_2 - h_1}{2}\cos\left(\pi\frac{t - t_1}{t_2 - t_1}\right), \tag{10}$$

which preserves the correct derivative sign at each extreme, avoids overshoot, and requires no calibration.

A more flexible alternative is to use a characteristic tidal shape. Here, we fit a constrained polynomial to high-resolution observations from the calibration period, providing an empirical representation of the local tidal curve. The polynomial is defined in normalized time between successive extrema and is constructed to satisfy two conditions: it must pass exactly through the prescribed extrema, and its time derivative must reduce to zero at those extrema. These constraints ensure consistency with the turning points of the tidal cycle and prevent spurious oscillations between successive extrema, while allowing asymmetric and nonlinear tidal shapes to be represented (see Appendix A).

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

5 Materials and Methods

5.1 Study site and data sources

The methodology is evaluated using observations from two stations in the Scheldt estuary (Figure 2): Bath, The Netherlands, and Melle, Belgium. Furthermore, synthetic datasets derived from a harmonic fit of water levels at Bath are used to assess the sensitivity to measurement uncertainty (Section 5.5). Although both stations are located within the same estuarine system, they exhibit markedly different hydrodynamic regimes. Bath is located in the Western Scheldt, where the tidal dynamics are predominantly stationary. In contrast, Melle is situated in the upper tidal river, where tidal propagation is strongly modulated by river discharge. These two sites therefore provide complementary test cases for evaluating the performance of the HLTA framework under stationary and non-stationary conditions.

Bath is characterized by a dominantly semidiurnal tide (form factor 0.06) with a mean tidal range of 4.83 m. The tidal signal exhibits pronounced nonlinear distortion, originating from shallow-water processes in the lower Scheldt and adjacent North Sea (Horsburgh and Wilson, 2007). This results in a strongly asymmetric tidal wave with significant overtides, including M_4 (0.11 m) and M_6 (0.12 m). Non-tidal variability is relatively small (variance ratio 0.03) and is primarily associated with meteorological forcing. Bath therefore provides a representative case for evaluating HLTA under stationary but distorted tidal conditions.

Melle is located in the upper Sea Scheldt, where the tidal signal is strongly influenced by river discharge. The mean tidal range is 2.58 m and the form factor is 0.07 (2004-2023). Variations in discharge significantly affect tidal amplitude, propagation, and asymmetry, leading to pronounced non-stationarity. The non-tidal contribution is substantial (variance ratio 0.30). Over 1971-2022, the mean discharge is 35 $m^3 s^{-1}$, with an average annual maximum of 249 $m^3 s^{-1}$. This site provides a challenging test case for evaluating non-stationary HLTA formulations.

High-resolution (10-minute) water-level data and corresponding high-low records were obtained from Rijkswaterstaat for Bath (https://waterinfo.rws.nl) and from MOW-HIC for Melle (https://waterinfo.be) over the 2015-2021 period. For Melle, the corresponding river discharge records were also obtained from MOW–HIC, where discharge is measured using an acoustic discharge meter (ADM).

5.2 Preprocessing and astronomical reference

All timestamps were converted to Coordinated Universal Time (UTC) and water levels were referenced to a consistent vertical datum (NAP - Dutch Ordnance Datum). High-low times are rounded to the nearest minute. Lunar transit times and associated astronomical parameters were computed using the DE421 planetary and lunar ephemeris (Folkner et al., 2009). The upper and lower lunar transits were evaluated at Greenwich (0° longitude), providing a consistent astronomical reference independent of station location.

Each observed extreme was assigned to the closest preceding lunar transit, and the corresponding time lag relative to that transit was computed. When extrema occur close to a transit, this procedure can lead to incorrect assignments. To correct for this, the distribution of time lags was analyzed and the median lag was used to identify and reassign mismatched extrema,

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

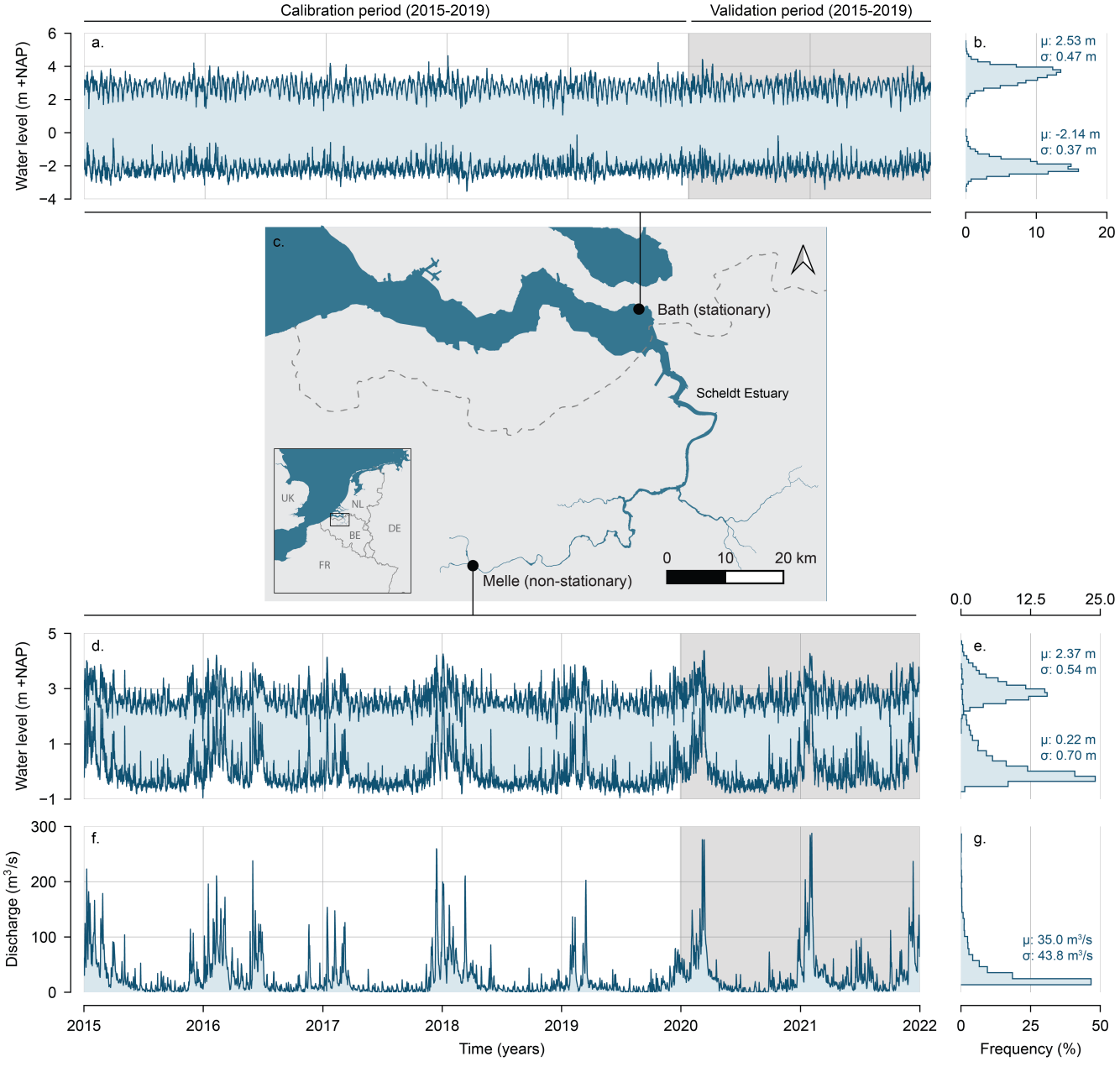

Figure 2. Locations of the Bath and Melle tide gauges within the Scheldt estuary and corresponding records. (a) Time series of water levels at Bath; (b) corresponding distributions of high and low water levels at Bath; (c) map of the Scheldt estuary showing tide gauge locations; (d) time series of water levels at Melle; (e) corresponding distributions of high and low water levels at Melle; (f) time series of river discharge at Melle; and (g) distribution of discharge at Melle. For all distributions, μ and σ indicate the empirical mean and standard deviation estimated on the whole record.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

ensuring a consistent phase relationship. At each transit time, astronomical predictors were evaluated, including lunar phase, solar phase, lunar declination, and Earth-Moon distance.

5.3 Model setup

The HLTA framework was applied using two formulations: LPH-HLTA and EA-HLTA (Section 4). For each station, both models were fitted separately to the extrema-based variables, including the mean and deviation components for the upper and lower transits.

The LPH-HLTA formulation represents each extrema variable as a sum of selected long-period harmonic constituents. Rather than including all possible constituents, only those that are statistically significant (Section 5.6) and supported by the spectral characteristics of the extrema series (Section 6.1.2) were retained.

The EA-HLTA formulation instead expresses extrema as functions of the astronomical state. The predictors include lunar phase, solar phase, lunar declination, and Earth–Moon distance. Lunar and solar phases are represented using sine–cosine pairs, allowing the model to estimate phase shifts directly. For the lunar declination and Earth–Moon distance, an optimal lag relative to lunar transit was determined prior to model fitting. This was achieved by applying time shifts to the predictor time series and selecting the lag that maximized correlation with the corresponding extrema series. This procedure aligns each predictor with the timing of its strongest influence on the extrema and accounts for the lag commonly observed between full moon or new moon and spring tides, often referred to as *the age of the tide*.

The model parameters in both formulations were estimated using iteratively reweighted least squares (IRLS; Holland and Welsch 1977). To account for non-stationary effects at Melle, coefficients were allowed to vary with river discharge using both linear and quadratic relationships (Equation 8).

As a baseline for comparison, we also included a constant–extrema model. This model assumes that extrema occur with fixed mean heights and fixed time offsets relative to lunar transit, independent of astronomical variability. Therefore, it represents the simplest possible HLTA formulation and serves as a benchmark for quantifying the added value of the more advanced models.

5.4 Bootstrap uncertainty analysis

Uncertainty in model parameters was assessed using a residual block bootstrap approach. For each model, bootstrap realizations $n = 1000$ were generated and parameter distributions were estimated on each resample. Uncertainty was quantified using standard deviations and percentile-based confidence intervals.

Residuals were calculated for each tidal transit, including errors in both water level and timing of high and low waters. The bootstrap resampling was then performed at the transit level, preserving the joint structure of height and timing errors. To account for seasonal variability, a seasonal stratification was applied: blocks of approximately one month were sampled in sequence, ensuring that the synthetic series follows the annual cycle while allowing interannual variability. For each bootstrap realization, resampled residuals were added to the model predictions to generate synthetic observation series, after which model parameters were re-estimated. The bootstrap framework was also used to assess the sensitivity to residual noise by scaling the

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

resampled residuals by uniform factors of 0–2. In this sensitivity analysis, $n = 25$ bootstrap realizations (per error level) were used to quantify the precision of the methods.

The use of a block bootstrap is essential because residual errors exhibit significant temporal autocorrelation and cross-dependence between height and timing (see Appendix B). Standard uncertainty estimates from commonly used tidal analysis packages typically assume simplified error structures and tend to underestimate uncertainty. By resampling residuals in blocks, the method preserves the empirical dependence structure of the errors (see Innocenti et al. 2022 for a similar approach).

5.5 Experimental design

The performance of the HLTA framework was evaluated under three complementary settings: (i) stationary tidal conditions (Bath), (ii) controlled perturbation experiments to assess sensitivity to measurement uncertainty, and (iii) non-stationary conditions influenced by river discharge. In addition, predicted extrema were used to reconstruct continuous water-level time series by interpolation between successive high and low waters. For all model comparisons, a common calibration period (2015-2019) and an independent validation period (2020-2021) were used. This split was applied consistently across all methods to allow for direct comparison of predictive performance.

Bath was used as the stationary test case. At this site, the constant-extrema model, LPH-HLTA, and EA-HLTA were compared against harmonic analysis applied to both high-resolution (10-minute) water-level data and high-low observations. Harmonic analysis was performed using a modified version of the UTide package (Codiga, 2011), with trend and nodal corrections disabled and using the robust fitting option. The harmonic analysis of the 10-minute data serves as the reference solution, while the harmonic analysis of the high-low data was evaluated with and without derivative constraints, which are not available in the original UTide package.

Melle was used as the non-stationary test case because river discharge strongly affects tidal amplitude and timing. In this case, the HLTA formulations were extended to include discharge dependence and compared with non-stationary harmonic analysis estimated through the NS_Tide package on hourly water-level data using a similar formulation with linear relationships with discharge (Matte et al., 2013; Krvavica et al., 2025; Innocenti et al., 2022). The application of NS_Tide directly to high–low observations was tested but is not presented here, as results suggest that additional methodological development may be required before the approach can be consistently applied to HW/LW data.

Sensitivity to measurement uncertainty was assessed by perturbing the Bath high- and low-water predictions based on a harmonic fit on 10-minute data perturbed with random errors derived from the residuals. These realizations were generated using seasonally–stratified bootstrap sampling, thereby preserving the covariance structure of errors in both water level and timing (Innocenti et al., 2022). The perturbations were scaled to prescribed noise levels before being added to the predicted extrema, and the resulting experiments were used to quantify the robustness of the different methods to uncertainty in the high-low observations.

Finally, predicted extrema were used to reconstruct continuous tidal signals by interpolating between successive high and low waters. Reconstruction skill was evaluated for both stationary and non-stationary conditions relative to the reference water-level record.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

5.6 Parameter selection

The selection of harmonic constituents and empirical predictors was guided by the spectral characteristics of the extrema series and their statistical significance. A preliminary set of candidate terms was defined for each formulation. For the harmonic approach, this included the standard set of UTide constituents appropriate for a 5-year record length at 10-minute resolution (68 constituents). For the LPH formulation, the candidate terms consisted of long-period constituents used by Horn (1948, 1960). For the empirical formulation, the candidate predictors consisted of quadratic relationships with the four astronomical forcing variables.

Parameter selection was performed using a seasonally-stratified block residual bootstrap to assess the distribution and significance of the fitted coefficients. Details of this approach are provided in Section 5.4 and in Appendix B.

For harmonic constituents, terms were retained when the bootstrap statistics met the following criteria: (i) the lower 2.5% amplitude quantile exceeded 0.5 cm for height and 0.5 min for lag variables, corresponding to half the measurement resolution, (ii) the phase circular standard deviation was less than 30°, and (iii) the median amplitude exceeded half the width of the corresponding 97.5-2.5 quantile bootstrap interval, $(a_{97.5} - a_{2.5})/2$.

For the empirical formulation, predictors were retained when their bootstrap confidence intervals indicated a statistically robust relationship with the extrema variables. In addition to satisfying the amplitude and stability criteria above, terms were required to exhibit a non-zero response over part of the predictor domain, as indicated by the 95% bootstrap confidence interval. The final selection of parameters is presented and discussed in Section 6.1.

5.7 Evaluation metrics

Model performance was evaluated based on the overall error of the predicted tidal extrema and the reconstructed water levels. The Root-mean-square error (RMSE) was used as a measure of total error, while the mean error (ME) was used to quantify systematic bias.

For high-low predictions, RMSE and ME were computed separately for water levels (in meters) and timing (in minutes), and evaluated for all extrema combined, as well as for high and low waters individually. For harmonic-analysis-based approaches, the predicted high and low waters were determined by first reconstructing water-level time series at a 1-min temporal resolution, followed by identification of local extrema; consequently, the predicted extrema have the same temporal resolution (1 min) as the observations. In contrast, for the HLTA approaches, tidal extrema are predicted directly without first reconstructing the continuous water-level signal. For reconstructed water-level time series, RMSE was computed over the full record relative to the observed water levels. All performance metrics were evaluated over the independent validation period (2020-2021).

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

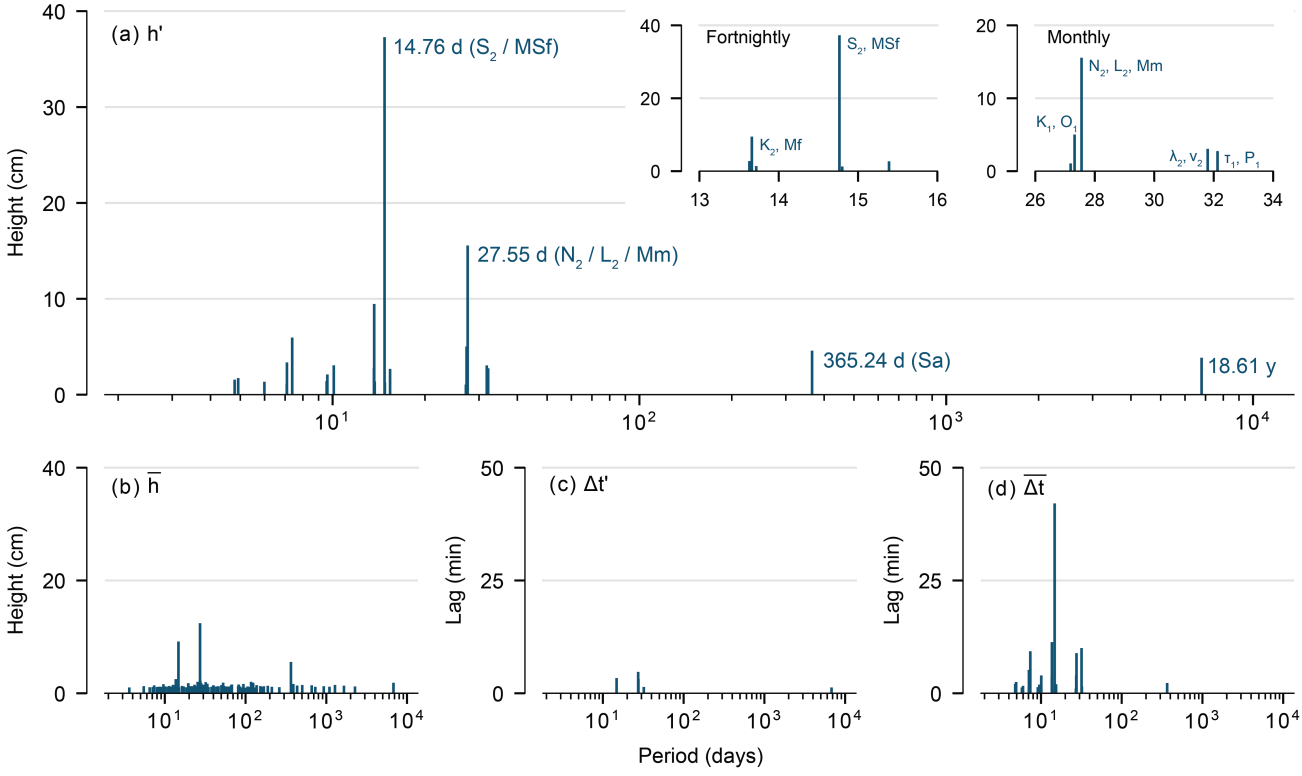

Figure 3. Spectral decomposition of high- and low-water observations for the upper transit at Bath over the period 2000–2021: a) Spectrum of the height deviation (range) term (h'), with dominant periods highlighted (corresponding high-frequency constituents indicated in brackets) and inset panels zooming into the fortnightly and monthly bands. b) Spectrum of the mean height ($\overline{h}$). c) Spectrum of the time-lag deviation ($\Delta t'$). d) Spectrum of the mean time lag ($\overline{\Delta t}$).

6 Results

6.1 Spectral Structure and Parameter selection

6.1.1 Spectral structure of height and time-lag series

Figure 3 shows the spectra of the high- and low-water heights and time lags for the upper transit, together with the corresponding mean and deviation components introduced in Section 4. Expressing the extrema in terms of mean and deviation concentrates most of the energy density in height deviation h' and mean lag $\overline{\Delta t}$, which represent the modulation of tidal amplitude and mean phase of the tide.

Because the extrema are referenced to lunar transits, the semidiurnal oscillation is absent from the spectra, while the diurnal signal is captured through differences between upper and lower transits. As a result, variability is dominated by slowly varying modulations of the tidal envelope. The strongest peak occurs at the fortnightly constituent MSf (14.77 d), associated with the

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

spring-neap interaction of M_2 and S_2. Additional energy density corresponds to the monthly Mm cycle (27.55 d), related to modulation by N_2 and L_2, and to the lunar fortnightly Mf (13.66 d). Smaller peaks show the importance of constituents with a period of one month (27.3 d), associated with the diurnal constituents K_1 and O_1, as well as those at seasonal (Sa) and nodal time scales.

Higher harmonics of the dominant modulation frequencies are also visible, including a pronounced component near 7.38 d corresponding to an overtide of the spring-neap cycle. These arise naturally from the superposition of multiple tidal constituents with slightly different frequencies. When expressed as an amplitude-modulated oscillation, the resulting modulation of the tidal envelope is not a purely sinusoidal function of the difference frequency but contains harmonics of that frequency. Consequently, the spectra of the extremum series exhibit peaks at frequency values corresponding to multiples of the dominant long-period component frequencies, such as the spring-neap frequency. A short derivation illustrating this representation is provided in Appendix C. Overall, the spectra indicate that the variability of the extremum series is governed by a limited set of long-period modulations. This concentration of energy density provides a compact basis for both LPH-HLTA and EA-HLTA representations.

6.1.2 Selection of long-period constituents (LPH-HLTA)

The initial set of candidates for long-period constituents for the LPH-HLTA representation was based on the long-period tidal terms used by Horn (1948, 1960). To account for the limited record length, closely spaced constituents were first screened using a relaxed Rayleigh criterion (factor 0.5), allowing partially resolved frequencies to be retained when supported by the data. Relaxed Rayleigh criteria can be justified when the signal-to-noise ratio is sufficiently large (Munk and Hasselmann, 1964). The remaining candidate constituents were evaluated using the seasonally constrained bootstrap procedure described in Section 5.4.

The resulting constituent sets differ among the high and low extrema (Table 1). Consistent with the spectral structure described above, a substantially larger number of constituents are retained for the height deviation h' and the mean lag $\overline{\Delta t}$ (13 and 14, respectively) than for the mean height $\bar{h}$ and lag deviation $\Delta t'$ (3 and 4, respectively). This reflects the concentration of tidal variability in the modulation of the amplitude and the mean phase components of the extrema series.

The length of the record places a practical limit on the number of long-period constituents that can be resolved. For example, separating the monthly group (Mm, N_2, L_2) from the nearby diurnal group (K_1, O_1) requires a record length of 8.85 years under strict Rayleigh criteria. The relaxed criterion adopted here allows both groups to be retained, and both are found to be robust in bootstrap analysis. Resolving the full 18.6-year nodal modulation requires records approaching the length of the nodal cycle itself (or at least half its length under a 0.5 Rayleigh criterion).

6.1.3 Selection of astronomical predictors (EA-HLTA)

Figure 4 shows the relationships between the extrema variables and the selected astronomical predictors. Candidate predictors included the lunar phase ϕ_m, solar phase ϕ_s, lunar declination δ_m, and lunar distance R_m, for the linear and quadratic representations. Because the tidal response to astronomical forcing is not instantaneous, optimal lags were introduced for the

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

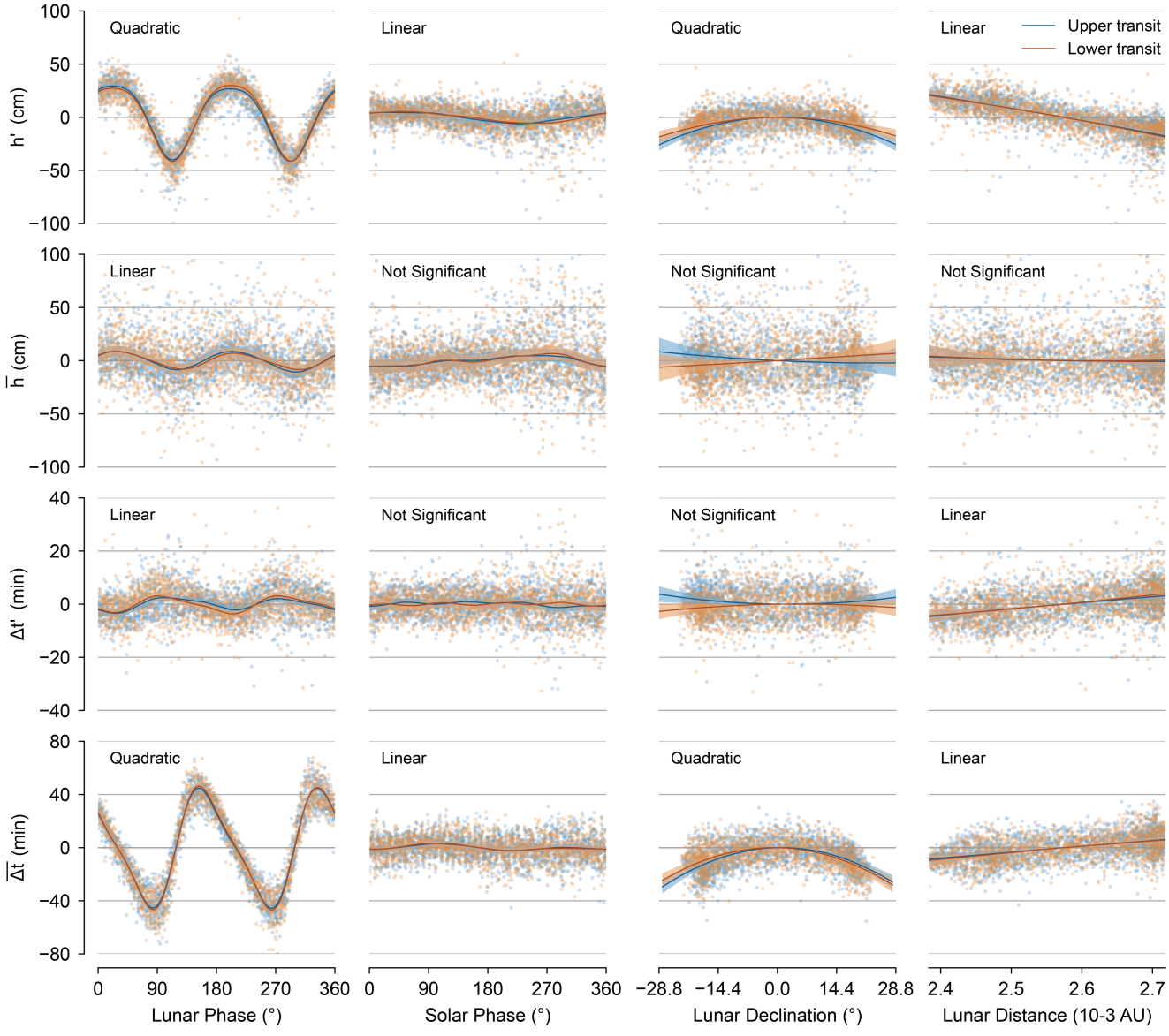

Figure 4. Relationships between extrema variables (rows) and selected astronomical predictors (columns) for upper and lower transits. Candidate predictors include lunar phase (ϕ_m), solar phase (ϕ_s), lunar declination (δ_m), and lunar distance (R_m), represented using low-order (linear and quadratic) fits. Each subplot shows the fitted relationship, with annotations in the upper-left corner indicating whether the fit is statistically significant (linear or quadratic) or not significant. Blue lines correspond to the upper transit and orange lines to the lower transit. Shaded regions indicate the 95% confidence intervals, while dots represent observations after removing the influence of other predictors (i.e., showing the fitted relationship with residual variability).

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

Table 1. Selected set of long-period tidal constituents used in the LPH–HLTA representation. The first column lists the constituent period (in days), followed by the corresponding Doodson numbers (indices 2–5; index 1 is 0, index 6 is irrelevant at the timescales considered). Columns three to six indicate whether each constituent is found to be significant and robust in the spectral decomposition of high–low water height differences (h'), mean heights ($\overline{h}$), timing differences ($\Delta t'$), and mean timings ($\overline{\Delta t}$), respectively. The final column provides the equivalent constituents in a continuous water-level signal representation.

Period (days)	Doodson Numbers	h'	$\overline{h}$	$\Delta t'$	$\overline{\Delta t}$	Equivalent Constituents
365.2422	[0, 1, 0, 0]	✓	✓		✓	Sa
32.1284	[1, -2, 0, 0]	✓		✓		τ_1, P_1
31.8122	[1, -2, 1, 0]	✓			✓	Msm, λ_2, ν_2
27.5546	[1, 0, -1, 0]	✓		✓	✓	Mm, N_2, L_2
27.3217	[1, 0, 0, 0]	✓	✓	✓	✓	K_1, O_1
15.3874	[2, -3, 0, 0]	✓			✓	T_2
14.7654	[2, -2, 0, 0]	✓	✓	✓	✓	MSf, S_2, μ_2
13.6608	[2, 0, 0, 0]	✓			✓	Mf, K_2
10.0847	[3, -4, 1, 0]	✓			✓	
9.6138	[3, -2, -1, 0]				✓	ϵ_2, ζ_2
7.3827	[4, -4, 0, 0]	✓			✓	
7.0958	[4, -2, 0, 0]	✓			✓	
5.9921	[5, -6, 1, 0]				✓	
4.9218	[6, -6, 0, 0]	✓			✓	
4.7926	[6, -4, 0, 0]				✓	

predictor variables. For circular phases, these lags are implicitly accounted for through the sine and cosine representations. For the non-circular predictors, the optimal lags were determined from their correlation with the height deviation h'. The strongest relationships were obtained for lags of approximately 13 hours for lunar declination and 47 hours for lunar distance, and these values were adopted in the regression analysis. Bootstrap resampling was used to identify which predictor relationships can be robustly estimated using the available samples. The terms were retained when they satisfied the stability criteria described above. In part of the predictor domain, they exhibited a statistically significant response (Section 5.6).

The resulting predictor selection differs across the eight daily extrema. The strongest and most consistent relationships occur for the height deviation h' and the mean lag $\overline{\Delta t}$, where the dominant dependence is on the second harmonic of the lunar phase, represented by $\sin(2\phi_m)$ and $\cos(2\phi_m)$, corresponding to the spring-neap cycle. The first harmonic ($\sin(\phi_m)$ and $\cos(\phi_m)$) is not retained for any variable, consistent with the absence of asymmetry between the waxing and waning lunar phases. The solar phase contributes less and is retained only through its first harmonic, representing seasonal modulation. Lunar declination is also a significant predictor of h' and $\overline{\Delta t}$, with a quadratic dependence reflecting the symmetric influence of declination on

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

tidal amplitude and timing. The lunar distance is linearly related to most high and low extrema, and is retained for all except the mean height $\bar{h}$.

In general, the empirical-astronomical representation confirms the spectral interpretation of Section 6.1.1: the variability among the extremum series is dominated by spring-neap modulation, with secondary contributions from declination, lunar distance, and seasonal forcing. Because these effects are expressed directly in terms of astronomical parameters, long-period modulations, such as the nodal cycle, are implicitly captured through their influence on the forcing variables, without requiring explicit harmonic representation (i.e. nodal corrections).

6.1.4 Constituent selection for harmonic-analysis benchmarks

To construct harmonic-analysis benchmarks comparable to HLTA approaches, the selection of constituents was performed using the same seasonally stratified bootstrap approach described in Section 5.4. For the reference 10-minute water-level record, applying the selection criteria to the UTide constituent basis for the available 5-year record length (68 constituents) yielded 41 significant constituents. These constituents form the basis for the reference harmonic analysis.

To assess the impact of high-low sampling, the same procedure was applied to the harmonic analyses performed on the extrema datasets. This yielded 22 significant constituents for the standard harmonic analysis and 24 constituents when a derivative constraint was applied. Lists of retained constituents for each configuration are provided in Table A1.

6.2 Performance under stationary conditions (Bath)

6.2.1 Accuracy of HW/LW prediction

Large differences in predictive skill are observed between the different approaches (Figure 5). Harmonic analysis applied directly to high-low (HL) observations performs poorly, particularly for timing. Using the reduced set of statistically significant constituents yields validation errors of approximately 0.46 m in height and 59 minutes in timing (RMSE), which is substantially worse than the HLTA formulations (Figure 5e). Introducing a derivative constraint improves the results, reducing errors to approximately 0.26 m and 23 minutes. Although this brings height predictions close to the accuracy of the HLTA approaches, timing errors remain significantly larger.

The constant-extrema baseline model yields validation errors of approximately 0.41 m in height and 36 minutes in timing. While this representation captures the mean tidal structure, it does not reproduce the temporal modulation of the tidal cycle.

Both HLTA formulations substantially improve predictive accuracy. The long-period harmonic HLTA (LPH-HLTA) achieves a validation RMSE of approximately 0.25 m in height and 9.5 minutes in timing, while the empirical-astronomical formulation (EA-HLTA) yields errors of approximately 0.27 m and 12.4 minutes, respectively. These results are comparable to, and for timing even better than, the performance of harmonic analysis applied to the full 10-minute water-level record (0.25 m and 15.4 minutes). The LPH-HLTA formulation reduces timing errors by approximately 38% relative to the reference solution (10-minute HA), while EA-HLTA achieves an improvement of approximately 20%. The HLTA formulations also exhibit reduced timing bias compared to harmonic analysis of the 10-minute record. The reference harmonic analysis shows systematic offsets

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

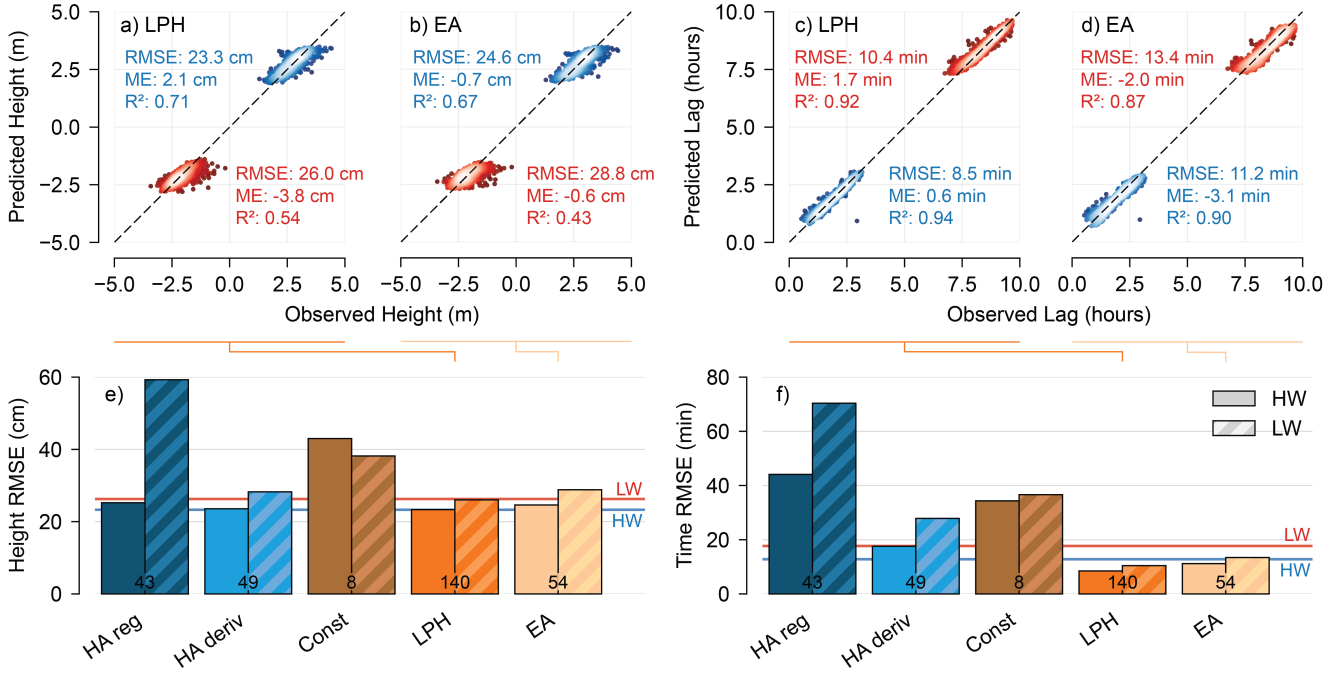

Figure 5. Validation results at Bath (2020–2021) for predicting high- and low-water extrema. The upper panels show comparisons between predicted and observed extrema heights (a–b) and time lags (c–d) for the LPH method (a, c) and the EA method (b, d). In these scatter plots, lighter point colors indicate regions of higher local point density. The lower panels present RMSE values for extrema heights (e) and time lags (f) for harmonic analysis applied to high–low data, the constant-extrema baseline, and the two HLTA approaches. Horizontal lines indicate the performance of harmonic analysis using 10-minute water-level data (HW in blue; LW in red). Black numbers at the base of each bar denote the number of fitted parameters for the corresponding method.

of approximately +4.9 minutes for high water and -10.2 minutes for low water. In contrast, LPH-HLTA yields biases of +0.6 minutes (HW) and +1.7 minutes (LW), while EA-HLTA produces similarly small biases of -3.1 minutes and -2.0 minutes, respectively (Figure 5c-d).

The different approaches also vary substantially in model complexity. Harmonic analysis applied to HL data uses 43-49 parameters (21-24 constituents), but produces the largest prediction errors. The empirical-astronomical HLTA formulation uses approximately 54 parameters and achieves near-optimal performance, whereas the long-period harmonic HLTA uses a larger parameter set (140 parameters) and yields the most accurate predictions overall.

In addition to overall performance differences, a systematic asymmetry is observed between high water (HW) and low water (LW) predictions (Figure 5). Across most methods (LPH-HLTA, EA-HLTA, and harmonic analysis), LW predictions exhibit larger errors in both timing and height. This is notable given the greater variability in HW heights (Figure 2a), suggesting a higher prediction difficulty for HW. The constant-extrema model further illustrates this contrast: because it effectively represents the mean state, its RMSE largely reflects the variability of the underlying data, showing larger RMSE in HW heights

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

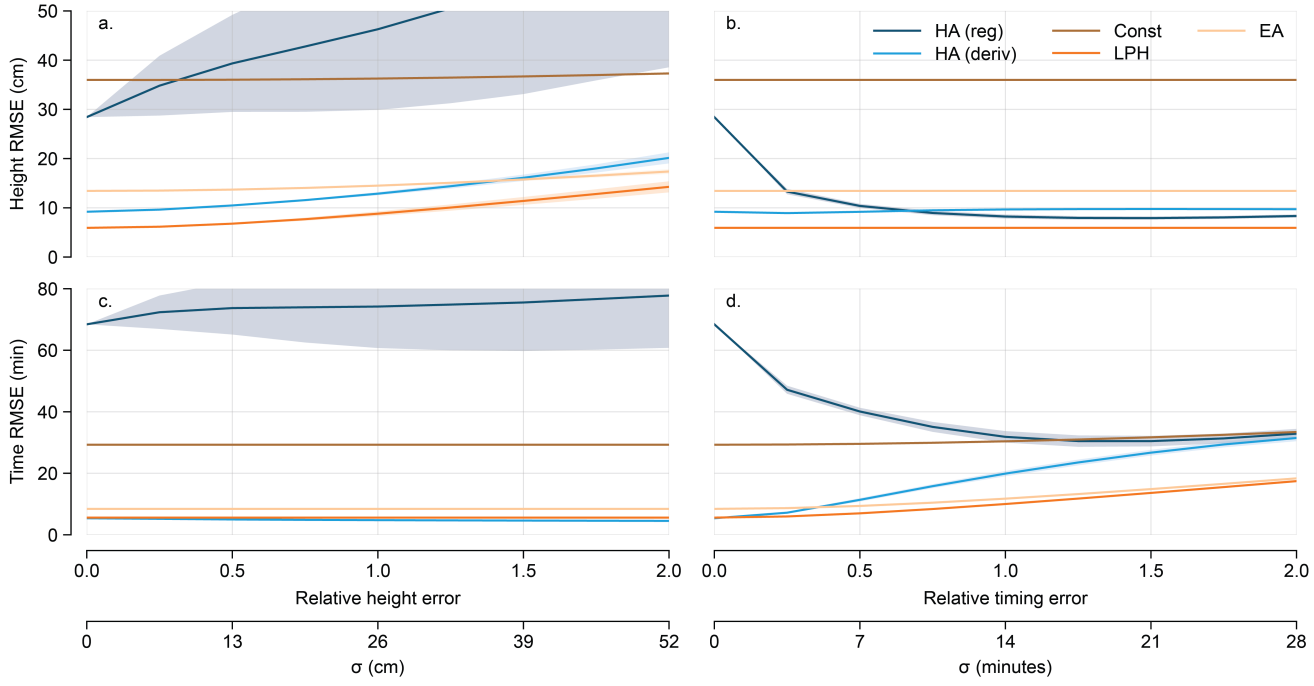

Figure 6. Sensitivity of the tested approaches to perturbations in high–low water observations. Errors are introduced using a bootstrap-based perturbation framework, and model performance is evaluated in terms of RMSE in height (a-b) and timing (c-d). The left column (a, c) shows sensitivity to height perturbations, while the right column (b, d) shows sensitivity to timing perturbations. Noise levels are scaled by the standard deviation, with the maximum values corresponding to twice the standard deviation at Bath. Shading represents the spread (standard deviation) of the parameter estimates over the bootstrap replicates (N=25).

510 but greater temporal variability for LW (Figure 5e–f). A likely explanation is that low waters are more strongly influenced by shallow-water processes and non-tidal disturbances. As water levels approach minimum depth, nonlinear effects such as bottom friction are expected to become more pronounced, and in combination with non-tidal contributions (e.g., surges), which may have a relatively larger impact during low water conditions, this leads to increased variability, particularly in shallow coastal systems such as the Scheldt estuary. These processes introduce additional variability in both the timing and magnitude of LW
515 that is not fully captured by the models, likely explaining the systematically reduced predictive skill compared to HW.

Overall, these results demonstrate that HLTA methods can recover tidal extrema with high accuracy using only high-low observations. In particular, the LPH-HLTA formulation achieves prediction accuracy comparable to harmonic analysis of continuous water-level data, while both HLTA approaches substantially outperform harmonic analysis applied directly to HL observations.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

6.2.2 Sensitivity to noise and measurement uncertainty

The different approaches exhibit markedly different sensitivity to observational uncertainty, both in magnitude and in how errors propagate between height and timing (Figure 6). HLTA formulations (constant, LPH-HLTA, and EA-HLTA) exhibit complete decoupling between height and timing: height perturbations affect only height predictions, while timing perturbations influence only timing predictions. This reflects the structure of the HLTA models, in which height and timing are treated as independent response variables. In addition, the overall sensitivity to noise is low: even for substantial perturbations (up to twice the standard deviation), the RMSE increases only moderately, and variability across the bootstrap realizations is small, indicating stable and robust parameter estimation. Among the HLTA approaches, LPH-HLTA is slightly more sensitive to noise than EA-HLTA, whereas the constant-extrema model is least sensitive due to its limited flexibility.

In the absence of added noise, HLTA models exhibit a finite baseline error (approximately 5.9 cm and 5.6 minutes for LPH-HLTA, and 13.4 cm and 8.4 minutes for EA-HLTA), reflecting that they do not capture the full harmonic signal. Because harmonic analysis is applied using a reduced set of significant constituents, it also exhibits a nonzero baseline error. Using the full set of constituents would yield a near-zero baseline error (not shown). With the reduced set, regular harmonic analysis shows a relatively large baseline error (29.0 cm and 1.15 hours), which decreases substantially with increasing timing perturbations, likely due to aliasing and a degree of equifinality.

Height perturbations (Figure 6a,c) lead to only moderate increases in height RMSE for HLTA models, while leaving timing errors unaffected. In contrast, HA of HL data shows strong and coupled responses: height errors increase rapidly with noise level, and timing errors increase simultaneously, indicating that amplitude errors propagate into phase estimates. This sensitivity is particularly pronounced for regular harmonic analysis, which exhibits large errors and substantial variability across realizations. Derivative-constrained harmonic analysis is more robust, but still shows stronger sensitivity than the HLTA approaches.

A similar pattern is observed for timing perturbations (Figure 6b,d). HLTA models again show isolated and limited responses, with timing RMSE increasing gradually while height predictions remain unaffected. In contrast, harmonic-analysis-based approaches exhibit coupled error propagation, with timing perturbations affecting both timing and height predictions. The increase in timing RMSE is substantial for derivative-constrained harmonic analysis, while height errors also increase, although less drastically, further highlighting the sensitivity of harmonic representations to phase uncertainty in sparse observations.

Overall, these results demonstrate that HLTA methods provide stable and robust estimates of tidal extrema under realistic measurement uncertainty, with errors that remain effectively decoupled. In contrast, harmonic analysis applied directly to high–low data is highly sensitive to observational errors, with strong dependence between height and timing and substantially larger prediction uncertainty.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

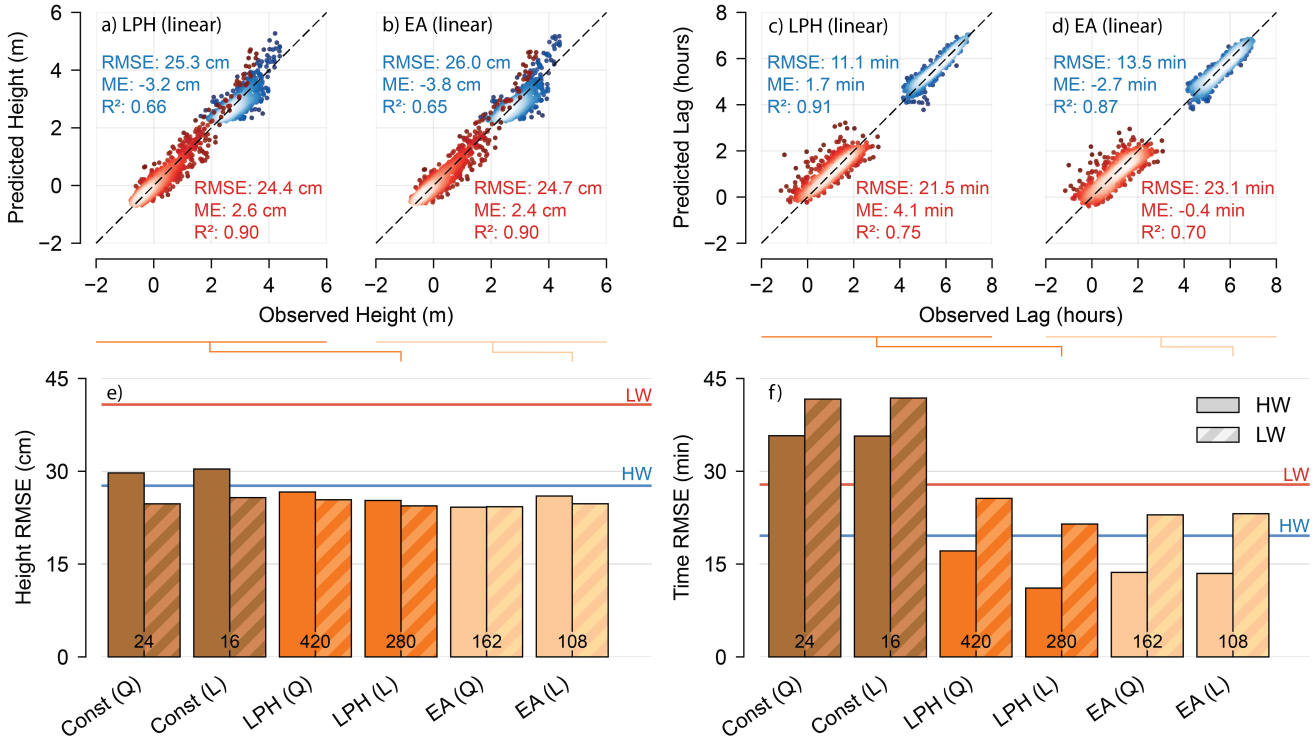

Figure 7. Validation results at Melle (2020–2021) predicting high- and low-water extrema. The upper panels show comparisons between predicted and observed extrema heights (a–b) and time lags (c–d) for the LPH method (a, c) and the EA method (b, d), both with a linear discharge relationship. In these scatter plots, lighter point colors indicate regions of higher local point density. The lower panels present RMSE values for extrema heights (e) and time lags (f) for the constant-extrema baseline, and the two HLTA approaches applied to non-stationary high–low data. Horizontal lines indicate the performance of non-stationary harmonic analysis (NS_Tide) using hourly water-level data (HW in blue; LW in red). Black numbers at the base of each bar denote the number of fitted parameters for the corresponding method. The Q/L labels on the bars indicate whether a Quadratic (Q) or Linear (L) discharge–tide relationship was fitted.

6.3 Performance under non-stationary conditions (Melle)

6.3.1 Model extension and reference solution

To evaluate performance under non-stationary conditions at the Melle tide gauge, the HLTA formulations were extended to include river discharge as an external covariate. The underlying tidal structure, including the selected long-period constituents and astronomical predictors (Section 6.1), was retained. As a reference, non-stationary harmonic analysis using NS_Tide was applied to hourly water-level data (Matte et al., 2013). The selection of constituents within this framework resulted in a model with approximately 150 parameters (37 constituents), based on statistical significance testing (Innocenti and Matte, 2025). The direct application of NS_Tide to high-low observations yielded poor performance (e.g. RMSE $>$ 1 m in height), and adapting

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

the method to the high-low setting would require additional methodological developments, beyond the scope of the present study.

Under non-stationary conditions, the HLTA approaches show a strong predictive skill (Figure 7). For water-level prediction, all HLTA formulations substantially outperform NS_Tide for low waters, reducing the validation RMSE from approximately 0.41 m to approximately 0.25 m. For high waters, performance is similar across the methods, with RMSE values ranging from 0.24 to 0.28 m. The discharge-dependent constant-extrema model achieves comparable accuracy for water levels, although it performs slightly worse for high waters. The differences between the methods are more pronounced for timing prediction. Both EA-HLTA and LPH-HLTA yield consistently lower RMSE values than NS_Tide, with improvements in the order of 10–29% for both high and low waters. In contrast, the constant-extrema model fails to capture discharge-induced shifts in the tidal phase and shows substantially larger timing errors.

Within the HLTA framework, the differences between the EA and LPH results are small. Both approaches yield comparable prediction errors for height and timing. The choice between linear and quadratic discharge dependence has limited impact on performance; for the LPH formulation, linear dependence performs slightly better, suggesting that higher-order terms may introduce mild overfitting. The EA formulation achieves similar performance with fewer parameters by directly including the astronomical predictors, indicating a more compact representation of the non-stationary tidal response.

The relatively good performance of the discharge-dependent constant-extrema model for water level heights suggests that further upstream, tidal modulation is damped, and discharge becomes a dominant control on the heights of the extrema. However, the larger timing errors for this model indicate that phase modulation depends on both tidal forcing and river discharge, requiring a more explicit representation as provided by the HLTA formulations.

6.4 Reconstruction of continuous tidal signals from high-low data

The HLTA formulations provide predictions of high and low water heights and timings, but do not directly reconstruct the continuous tidal signal. To obtain a continuous water-level time series, interpolation between successive extrema is required. The successive extrema were interpolated using either (i) cosine interpolation (Eq.10) or (ii) a fitted polynomial (Appendix A) representation of a characteristic tidal hydrograph derived from the calibration period 2015-2019 (Figure A1). The resulting time series were compared against the harmonic analysis of the high-resolution data (Figure 8) and the original observed data.

For the stationary case at Bath (Figure 8a), the harmonic analysis applied directly to high-low data performs poorly, with RMSE values exceeding 0.65 m. Although extrema are reproduced reasonably well, the intra-tidal structure is not captured. This is evident in the example of reconstructed time series (Figure 8c), where regular harmonic analysis produces spurious oscillations (including double highs and lows), while derivative-constrained harmonic analysis yields unrealistically flat high- and low-water stands.

In contrast, HLTA-based reconstructions perform substantially better. Even the constant-extrema model, combined with simple cosine interpolation, outperforms harmonic analysis applied to high-low data. The LPH-HLTA and EA-HLTA approaches yield reconstruction errors close to those obtained from harmonic analysis of the 10-minute record when cosine interpolation is used (LPH: 38.7 cm, EA: 41.2 cm, compared to 29.9 cm for HA_{10min}).

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

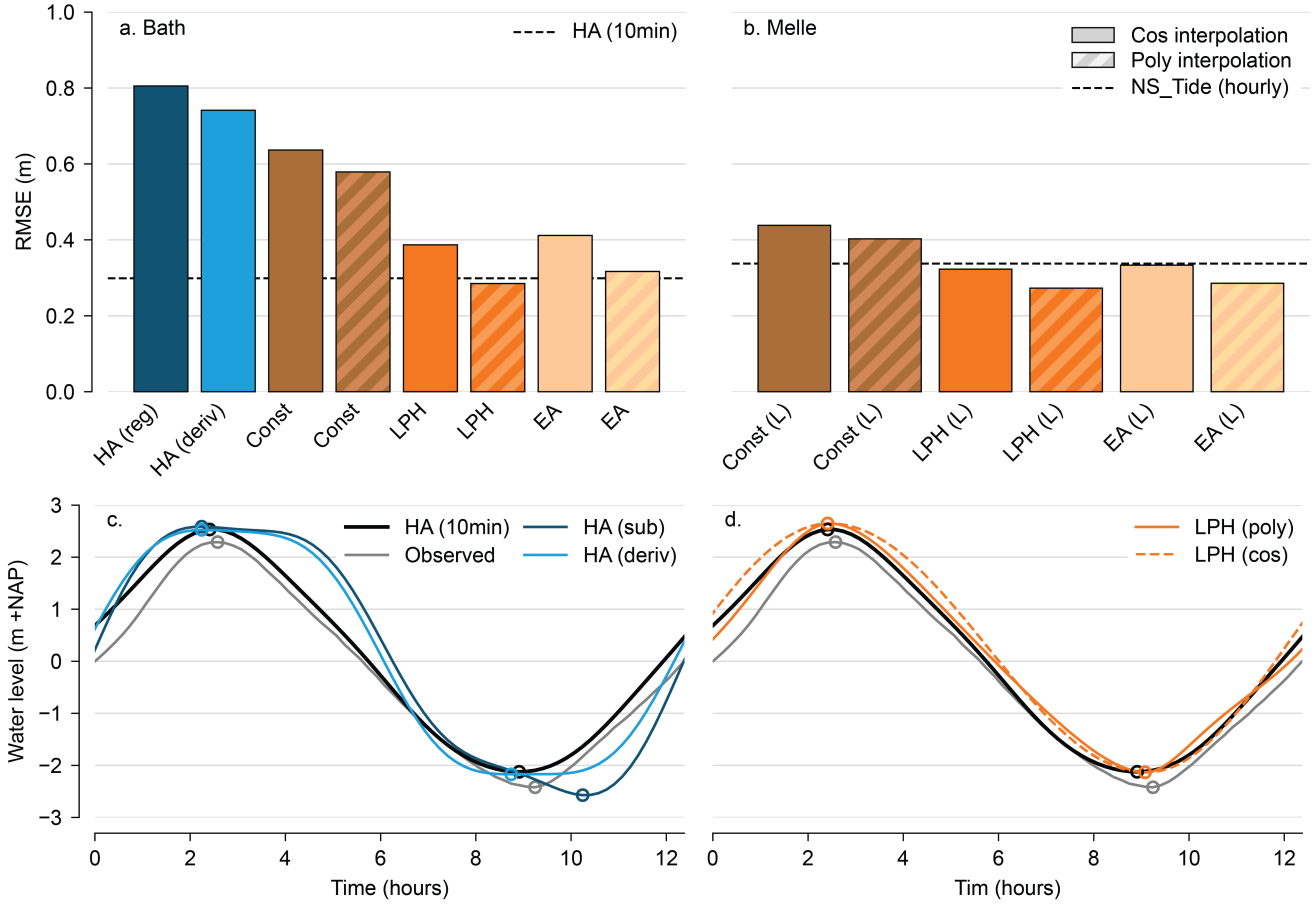

Figure 8. Reconstruction performance of different methods for generating continuous tidal time series at (a) Bath and (b) Melle. Bars show RMSE values for each approach, with solid bars representing polynomial interpolation (fitted over the 2015–2019 calibration period) and dashed bars representing cosine interpolation applied to the HLTA methods. (c) Example time series segment comparing observed and reconstructed water levels for selected methods; dots indicate the observed and predicted high- and low-water extrema.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

Using a fitted polynomial interpolation further improves performance. The EA-HLTA reconstruction achieves an RMSE of approximately 31.7 cm, while LPH-HLTA slightly outperforms the harmonic-analysis reference (28.5 cm versus 29.9 cm). These results indicate that accurate prediction of extrema, combined with a representative tidal hydrograph, is sufficient to reconstruct the continuous signal with an accuracy comparable to, or exceeding, the harmonic analysis of high-resolution data.

A similar pattern is observed under non-stationary conditions at Melle (Figure 8b). The constant-extrema model already provides a reasonable reconstruction when combined with interpolation, approaching the performance of NS_Tide applied to hourly data. The EA-HLTA and LPH-HLTA approaches match NS_Tide when cosine interpolation is used and outperform it when polynomial interpolation is applied. The differences between the EA and LPH formulations, as well as between the linear and quadratic discharge dependence, are small, consistent with results presented in 6.3.

Overall, these results show that the primary limitation in reconstructing continuous tidal signals from high-low data lies in the accuracy of the predicted extrema rather than in the interpolation itself. Once extrema are predicted accurately, even simple interpolation yields high-quality reconstructions, while more representative tidal shapes further improve performance. In contrast, harmonic analysis applied directly to high-low data fails to reproduce the shape of the tidal hydrograph, resulting in large reconstruction errors.

7 Discussion

7.1 Accurate and robust prediction of tidal extrema

This study demonstrates that high- and low-water observations can be used to achieve accurate and robust predictions of tidal extrema. The HLTA formulations reproduce both heights and timings with skill comparable to or better than harmonic analysis of high-resolution data. In particular, substantial improvements are obtained for the timing of extrema, with both reduced RMSE and strongly reduced bias relative to harmonic analysis of 10-minute observations. These results are consistent with the findings of Boesch and Müller-Navarra (2019) and reinforce the long-standing practical observation that HLTA-based approaches have remained in operational use well beyond the advent of harmonic analysis. In particular, the British Admiralty continued to rely on the Lubbock method for HW/LW prediction well into the twentieth century (Cartwright, 2000), while variants of the Horn-based German method are still operationally employed for extrema prediction today (Boesch and Müller-Navarra, 2019). Together, these results suggest that HLTA methods are not merely historical artifacts, but remain well suited for the prediction of tidal extrema, especially when formulated within a modern statistical framework.

The improved performance for timing is expected to be especially notable in shallow-water environments, where non-linear processes and tidal deformation play a dominant role. While harmonic analysis reconstructs the full signal through a spectral representation, HLTA directly models the height and time modulations of the extrema. This limitation of harmonic analysis is particularly evident when considering tidal asymmetry: for example, a representation based solely on an M_4 overtide (without allowing double maxima or minima) can produce a maximum rise or fall duration of approximately 4.73 h, whereas in many shallow systems tides are substantially more asymmetric, with tidal bores representing an extreme end-member (Hoitink and Jay, 2016). Consistent with this, George and Simon (1984) showed that in the upper Loire and Gironde estuaries, constituents

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

with frequencies exceeding 20 cycles per lunar day are required to accurately represent tidal deformation, whereas conventional tidal analyses typically include only lower-order constituents (UTide: 8th diurnal). This highlights that for many applications, direct modeling of the extrema may provide a more effective approach than reconstruction of the full tidal signal.

7.2 Suitability of HLTA for high-low observations

Harmonic analysis performs poorly when applied directly to high–low observations, as it is not designed for sparse sampling of the tidal cycle. In this study, harmonic analysis of extrema data yielded ill-conditioned solutions, strong parameter dependencies, and large prediction errors. In particular, overtides often lead to unrealistic double high- or low waters, while derivative-constrained formulations produce artificially flat high- and low water stands. In practice, this suggests that including higher-frequency constituents, particularly overtides exceeding the effective Nyquist sampling frequency, should be avoided when analyzing high-low data using harmonic analysis. Although studies based on satellite altimetry suggest that strict Nyquist limits do not necessarily apply (Parke et al., 1987), our results show that, for high–low sampling, inclusion of such constituents tends to degrade rather than improve the solution.

These issues arise not only from classical aliasing effects (Ji and Guohong, 1987), but also from the sampling itself. Specifically, they stem from the high–low sampling schemes, which give rise to complex aliasing patterns and data covariance structures. In high–low observations, the sampling is not independent but determined by the tidal signal itself, as observations occur at tidal extrema. As a result, the sampling scheme is intrinsically linked to the constituents being resolved, leading to broad-scale dependencies between harmonic constituents and parameter identifiability issues. As a result, multiple combinations of constituents can produce similar fits to the extrema, even when they represent physically distinct tidal signals. This introduces ill-conditioning in the parameter estimation and undermines the robustness and validity of harmonic solutions. In effect, the inverse problem becomes underdetermined at the level of individual constituents.

In contrast, HLTA directly targets the observed quantities, extrema heights and timings, and their modulation. This makes the method inherently suited to high–low data, particularly in shallow and strongly non-linear tidal systems. Previous applications of harmonic analysis to extrema observations have generally focused on less distorted tidal regimes (Foreman and Henry, 1979; Ji and Guohong, 1987; Li et al., 2004); the present results demonstrate that such approaches are especially problematic in environments with strong tidal deformation.

7.3 Choice of spectral versus astronomical representation

The similar performance of the long-period harmonic (LPH) and empirical–astronomical (EA) formulations demonstrates that tidal modulation can be described effectively using either a spectral or a forcing-based approach. The LPH formulation, based on Horn (1948), represents modulation through a set of long-period constituents, analogous to conventional harmonic analysis, while the EA formulation, based on Lubbock (1835), relates extrema directly to astronomical predictors, in line with classical response methods (Munk and Cartwright, 1966; Monahan et al., 2025). Despite these conceptual differences, both approaches yield comparable accuracy. The EA formulation achieves similar performance with fewer parameters, suggesting a more compact representation of the system, while the LPH formulation remains closer to conventional harmonic analysis and provides

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

more transparent insight into the dominant frequencies. For long stationary records (e.g. >= 2 years), the LPH formulation is expected to be advantageous because it can more directly and efficiently resolve the dominant tidal frequencies. In contrast, the EA formulation may offer greater stability for short, noisy, or non-stationary records, where its more compact predictor-based structure may reduce sensitivity to overfitting or ill-conditioning. In this study, relatively simple relationships between predictors and extrema were used. Further improvements may be possible by incorporating interactions among predictors or using more flexible functional forms and machine learning approaches.

7.4 Robustness to observational uncertainty

The HLTA formulations exhibit robustness to simulated measurement errors in high–low observations. Prediction errors increase only weakly with increasing noise levels, and the variability across the bootstrap replicates remains small. In addition, height and timing errors are fully decoupled, reflecting the decoupled formulation of the HLTA models. In contrast, harmonic analysis of high-low data degrades rapidly with increasing noise, particularly for timing. This sensitivity limits its applicability for sparse datasets. The robustness of HLTA makes it particularly suitable for historical observations and for shallow-water systems where tidal distortion may amplify the observational uncertainty.

7.5 Extension to non-stationary conditions

The HLTA framework extends naturally to non-stationary tidal systems by allowing model parameters to depend on external forcing, such as river discharge. This extension leads to improved predictions of both water levels and timing compared to non-stationary harmonic analysis (NS_Tide), particularly for low waters and for extrema timing. The relatively good performance of the constant-extrema model for height prediction indicates that friction-induced damping reduces tidal range modulations upstream, while timing remains strongly modulated. Beyond its predictive capability, the two HLTA approaches provide a useful tool to analyze the interaction between tides and river discharge and gain insight into estuarine dynamics.

7.6 Reconstruction of tidal signals from extrema

Accurate prediction of tidal extrema, combined with interpolation, allows reconstruction of the continuous tidal signal with high fidelity. Even simple cosine interpolation yields good results, while using a representative tidal hydrograph further improves performance, achieving accuracy comparable to or exceeding that of harmonic analysis based on high-resolution data. This indicates that reconstruction errors are primarily controlled by extrema prediction rather than the choice of interpolation.

In shallow and estuarine systems, where tidal waveforms are strongly deformed and asymmetric, harmonic analysis struggles to reproduce the signal without a large number of higher-order constituents (George and Simon, 1984). In contrast, HLTA combined with interpolation captures asymmetric rise and fall behavior through accurate extrema timing and a representative shape, without explicitly resolving high-frequency components.

For historical high-low datasets, the availability of a representative tidal shape may be limited. In such cases, several approaches can be adopted. A baseline reconstruction can be obtained using cosine interpolation when only high–low observa-

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

tions are available. Where modern high-resolution data exist, the shape of the tidal hydrograph can be derived from contemporary observations at the same or a nearby station. In cases where data originate from marigrams, which are often laborious to digitize in full, it may be sufficient to digitize only a limited number of tidal cycles to extract a representative tidal shape, and combine this with longer high–low records to reconstruct a continuous time series.

Further improvements may be expected by allowing the interpolation shape to vary with tidal conditions, such as tidal range, rise and fall duration, lunar phase, or river discharge. The extent to which the dimensionless shape of the tidal hydrograph varies between tidal cycles or over longer time scales remains an open question, particularly in the context of secular changes in estuarine morphology and dynamics. Notably, traditional tide tables, such as those of the Admiralty, already employ a related concept by combining high–low predictions with representative tidal shapes (typically one for neap and one for spring tides), motivated by practical space constraints (United Kingdom Hydrographic Office, 2025).

Several limitations should be noted. The HLTA framework does not directly produce continuous water levels and, therefore, relies on interpolation, which was shown to introduce only minor errors. However, interpolation choice affects higher-frequency constituents: cosine interpolation degrades them, whereas polynomial interpolation better preserves and resolves them. The approach requires sufficient data to resolve tidal modulation, typically at least one year, with longer records needed to capture longer-period components and ensure adequate coverage of astronomical predictors, although this was not explicitly analyzed in this study. In addition, accurate estimation of mean sea level remains challenging, as high–low observations primarily capture the mean tide level (Woodworth, 2017). Accurate interpolation of the tidal shape may help distinguish between the mean tide level and the mean sea level.

The results demonstrate that HLTA provides a powerful framework for analyzing tidal extrema from sparse observations. This opens new possibilities for the use of historical tide-gauge records, including datasets containing only high waters, and for applications such as detiding, e.g., for analyzing storm surges, flood timing prediction, and the study of tide–surge and tide–discharge interactions. More broadly, the results show that accurate and robust tidal analysis does not require continuous observations but can be achieved using well-structured representations of the tidal extrema and their modulation.

8 Conclusions

This study introduces a high–low tidal analysis (HLTA) framework to predict tidal extrema and reconstruct tidal signals from sparse high-low tidal observations. The main findings are:

1. High- and low water observations contain sufficient information to accurately predict the tidal extrema. Both HLTA formulations achieve prediction skill comparable to harmonic analysis of high-resolution data for water levels, and substantially improved performance for extrema timing, with reduced RMSE and bias.

2. Harmonic analysis applied directly to high–low data performs poorly, not only due to aliasing but also because constituent-driven sampling introduces broad-scale dependencies between constituents, leading to ill-conditioned parameter estima-

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

tion and unstable solutions. In contrast, HLTA directly models extrema and their modulation, making it inherently better suited to such data.

3. Two conceptually different HLTA formulations, a long-period harmonic (LPH) representation and an empirical–astronomical (EA) formulation, yield similar performance. This demonstrates that tidal modulation can be described effectively either through a spectral representation or through direct relationships with astronomical forcing.

4. The HLTA approaches are robust to observational errors and remain stable as noise level increases, in contrast to harmonic analysis, which degrades rapidly, particularly for timing predictions.

5. Either of the two HLTA formulations can readily be extended naturally to non-stationary conditions by incorporating external forcing, such as river discharge. In this setting, HLTA approaches outperform non-stationary harmonic analysis for low-water levels and extrema timing, demonstrating their ability to capture both amplitude damping and phase modulation.

6. Accurate prediction of the tidal extrema, combined with simple interpolation, allows reconstruction of the continuous tidal signal with high fidelity. In shallow and strongly distorted systems, HLTA-based reconstructions can match or exceed the performance of harmonic analysis applied to high-resolution data.

Overall, the results demonstrate that HLTA approaches effectively capture the compact representation of tidal variability contained in high–low observations. They both provide a robust and efficient alternative to harmonic analysis for both stationary and non-stationary tidal systems, with particular advantages for historical, sparse, or highly distorted tidal records.

Code and data availability. The code used for the High–Low Tide Analysis (HLTA) tools developed and applied in this study will be made publicly available via a GitHub repository (https://github.com/JorisBeemster/HighLow_TA). The data supporting the conclusions of this study are obtained from publicly available sources, which are cited in the main text and references.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

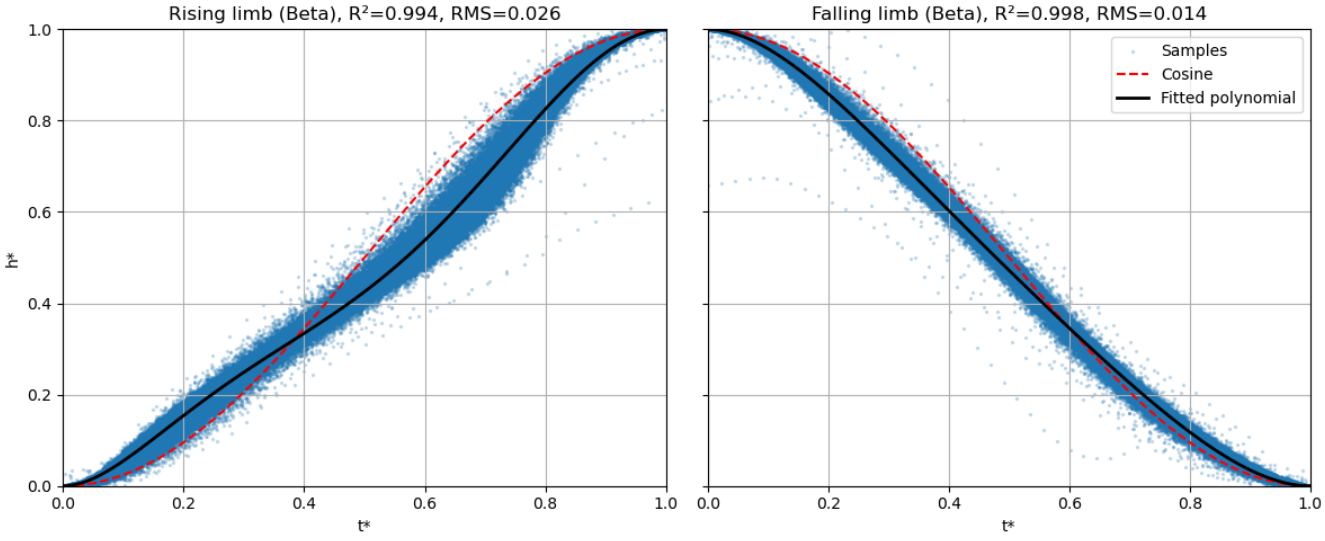

Figure A1. Observed and fitted dimensionless tidal shapes at Bath. Scatter points show normalized water level observations (h^*) as a function of normalized time (t^*) over the calibration period 2015-2019. The red dashed line indicates the reference cosine shape, while the black solid line shows the fitted polynomial shape. Results are shown separately for the rising limb (left) and falling limb (right).

Appendix A: Polynomial interpolation of tidal shape

To reconstruct a continuous tidal signal from predicted high and low waters, we employ a constrained polynomial interpolation between successive extrema. This approach provides a flexible representation of tidal asymmetry while ensuring consistency with the prescribed extrema.

A1 Dimensionless formulation

Consider two successive extrema at times t_1 and t_2 with corresponding water levels h_1 and h_2. We define a dimensionless time variable

$$\tau = \frac{t - t_1}{t_2 - t_1}, \quad \tau \in [0, 1], \tag{A1}$$

and a dimensionless water level

$$h^*(\tau) = \frac{h(t) - h_1}{h_2 - h_1}, \tag{A2}$$

such that $h^*(0) = 0$ and $h^*(1) = 1$ for a rising limb (and the reverse for a falling limb).

The interpolation problem is then reduced to constructing a smooth function $S(\tau)$ satisfying

$$S(0) = 0, \quad S(1) = 1, \quad S'(0) = 0, \quad S'(1) = 0. \tag{A3}$$

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

A2 Base polynomial shape

A minimal function satisfying these constraints is the cubic Hermite polynomial

$$S_0(\tau) = \tau^2(3 - 2\tau), \tag{A4}$$

which provides a symmetric, monotonic transition between the extrema.

A3 Flexible polynomial representation

To allow for asymmetric and nonlinear tidal shapes, we extend this representation using additional basis functions that vanish at the endpoints together with their first derivatives. The interpolating function is written as

$$S(\tau) = S_0(\tau) + \sum_{k=1}^{K} a_k \, \psi_k(\tau), \tag{A5}$$

where the basis functions $\psi_k(\tau)$ are constructed such that

$$\psi_k(0) = \psi_k(1) = \psi_k'(0) = \psi_k'(1) = 0. \tag{A6}$$

Examples of such basis functions include

$$\psi_1(\tau) = \tau^2(1-\tau)^2, \quad \psi_2(\tau) = \tau^2(1-\tau)^3, \quad \psi_3(\tau) = \tau^3(1-\tau)^2, \tag{A7}$$

with higher-order terms allowing additional flexibility in the interior of the interval.

The coefficients a_k control the shape of the tidal curve while preserving the boundary and smoothness constraints.

A4 Fitting procedure

The coefficients a_k are determined by fitting the interpolating function to observed or reference tidal curves, expressed in nondimensional form. Given a set of observations $\{(\tau_i, h_i^*)\}$, the parameters are estimated by minimizing the mean squared error

$$\min_{a_k} \sum_i \left[h_i^* - S(\tau_i)\right]^2. \tag{A8}$$

Rising and falling limbs are treated independently, with separate fits applied to each segment between successive extrema. This allows the interpolation to represent asymmetries between flood and ebb phases without imposing any symmetry constraints.The fitted shapes for Bath over the calibration period (2015–2019) are shown in Figure A1.

A5 Properties

The resulting interpolation has the following properties:

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

- Exact reproduction of the prescribed extrema ($h(t_1) = h_1$, $h(t_2) = h_2$),
- Zero slope at the extrema, consistent with turning points of the tidal cycle,
- Flexibility to represent asymmetric and nonlinear tidal shapes,
- Stability and absence of spurious oscillations due to the constrained basis.

Appendix B: Residual structure and justification for block bootstrap resampling

To quantify uncertainty in model parameters derived from high–low water observations, we employ a block bootstrap approach based on the residual errors between observed and predicted tidal extrema. This section provides additional detail on the structure of these residuals and motivates the use of block-based resampling.

B1 Definition of residual errors

Residual errors were computed by comparing observed and predicted high and low waters for each tidal transit. For each matched transit, errors were calculated in both water level and timing, yielding four residual variables:

- high-water height error ($e_{high,h}$)
- low-water height error ($e_{low,h}$)
- high-water timing error ($e_{high,t}$)
- low-water timing error ($e_{low,t}$)

These residuals represent the combined effects of observational uncertainty, unresolved physical processes, and model imperfections.

B2 Temporal autocorrelation and cross-dependence

The residual errors exhibit significant temporal autocorrelation, reflecting the persistence of meteorological forcing, surge conditions, and other low-frequency processes that are not fully captured by the tidal models. Figure B1 shows an example time series of residual errors together with their autocorrelation function, illustrating that correlations persist over timescales of several days to weeks.

In addition to temporal dependence, the residuals exhibit cross-dependence between variables. Height and timing errors are not independent, and errors in high and low water within the same tidal cycle are often correlated. This reflects the fact that both quantities are influenced by the same underlying physical processes, including tidal distortion, surge, and measurement uncertainty.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

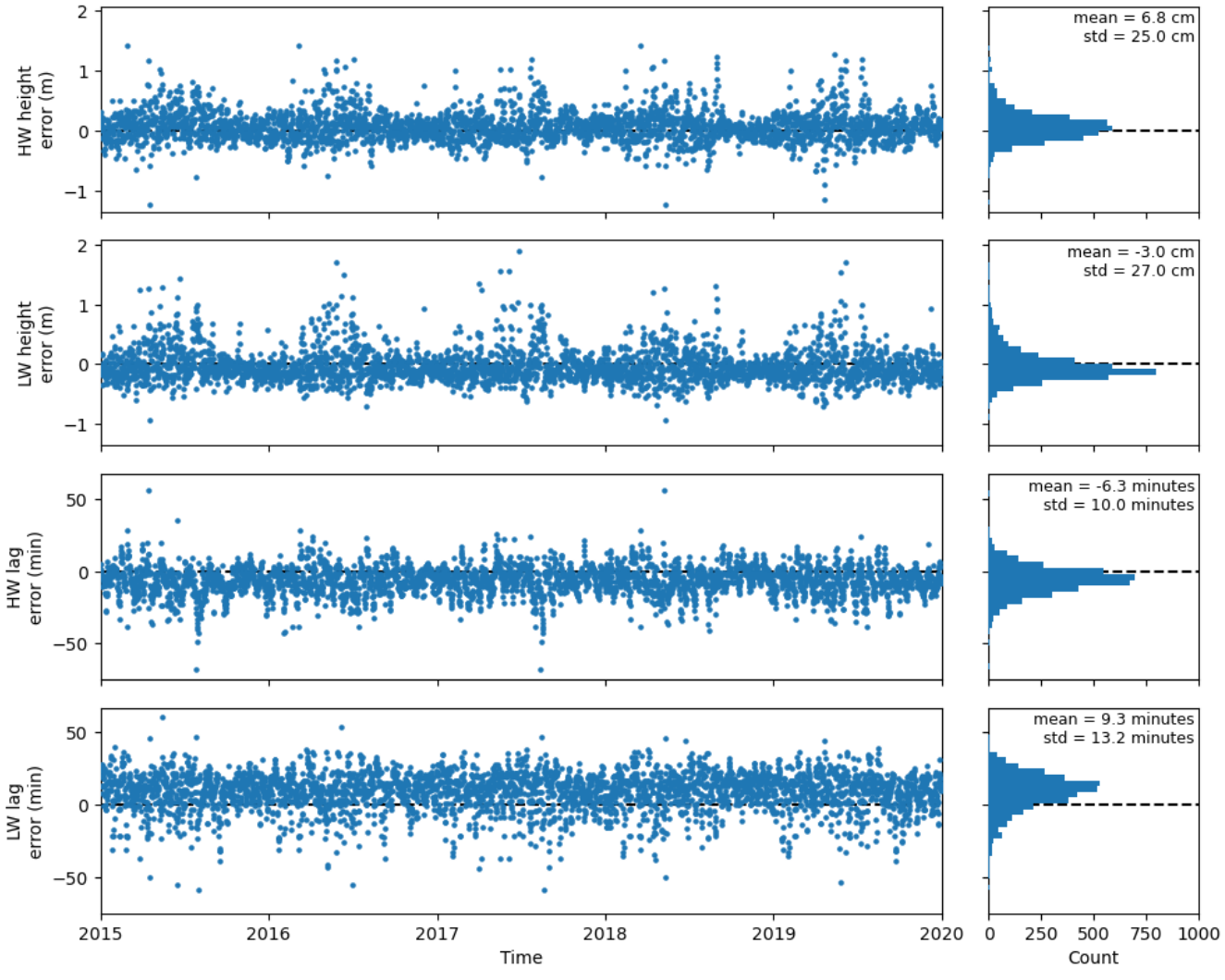

Figure B1. High- and low-water height and timing errors used for bootstrap resampling, shown as time series (left) and corresponding distributions (right). The distributions summarize the error structure, with the mean and standard deviation reported in the upper-right corner of each panel.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

B3 Implications for uncertainty estimation

The presence of temporal autocorrelation and cross-variable dependence violates the assumptions underlying standard (independent) bootstrap methods, which assume that residuals are identically and independently distributed. Applying such methods would lead to an underestimation of uncertainty, as correlated errors would be treated as independent realizations.

To address this, bootstrap resampling was performed at the level of tidal transits, preserving the joint structure of the residual vector. Each sampled element therefore consists of the complete set of height and timing errors associated with a single transit.

B4 Seasonally constrained moving-block bootstrap

To account for both temporal dependence and seasonal variability in residual statistics, a seasonally constrained moving-block bootstrap was used. Synthetic residual series were constructed using blocks of approximately one month (30.5 days), chosen to capture short-term temporal dependence while remaining short relative to the seasonal cycle.

The first block begins at a randomly selected transit and spans one month. Subsequent blocks are drawn from the corresponding calendar month, but may originate from a different year. This ensures that the synthetic residual series progresses through the annual cycle in the correct order while allowing inter-annual variability. The procedure is repeated until the synthetic series reaches the same length as the original data set.

B5 Advantages of the block bootstrap approach

This approach preserves:

- temporal autocorrelation in residual errors
- cross-dependence between height and timing errors
- seasonal variability in error statistics

and therefore provides a more realistic representation of uncertainty than analytical or independent-resampling methods. In particular, it avoids the assumption of independent or normally distributed errors and instead retains the empirical dependence structure of the residuals, consistent with the framework proposed by Innocenti et al. (2022).

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

Appendix C: Amplitude and phase modulation of two tidal constituents

Consider the superposition of two tidal constituents, for example M_2 and S_2,

$$\eta(t) = H_{M2}\cos(\omega_{M2}t + \phi_{M2}) + H_{S2}\cos(\omega_{S2}t + \phi_{S2}). \tag{C1}$$

This sum can be written as a single cosine with time-varying amplitude and phase,

$$\eta(t) = A(t)\cos\big(\omega_{M2}t + \phi_{M2} + \psi(t)\big), \tag{C2}$$

where the modulation depends on the frequency and phase differences

$$\Delta\omega = \omega_{S2} - \omega_{M2}, \qquad \Delta\phi = \phi_{S2} - \phi_{M2}. \tag{C3}$$

The instantaneous amplitude is

$$A(t) = \sqrt{H_{M2}^2 + H_{S2}^2 + 2H_{M2}H_{S2}\cos(\Delta\omega t + \Delta\phi)}. \tag{C4}$$

Thus, the amplitude varies on the beat timescale set by $\Delta\omega$. For the pair M_2–S_2, this is the spring–neap modulation. However, $A(t)$ is not a pure cosine in $\Delta\omega t$, because of the square root. Therefore, the modulation generally contains not only the fundamental difference frequency $\Delta\omega$, but also higher harmonics.

C1 Fourier-series representation of the amplitude modulation

Define

$$\theta = \Delta\omega t + \Delta\phi. \tag{C5}$$

Then

$$A(\theta) = \sqrt{H_{M2}^2 + H_{S2}^2 + 2H_{M2}H_{S2}\cos\theta}. \tag{C6}$$

This is a 2π-periodic function of θ and may therefore be written as a Fourier cosine series,

$$A(\theta) = a_0 + \sum_{n=1}^{\infty} a_n\cos(n\theta). \tag{C7}$$

Hence the amplitude modulation contains frequencies

$$n\Delta\omega, \qquad n = 1, 2, 3, \ldots \tag{C8}$$

in addition to the mean term.

A convenient parameterization is obtained by defining

$$a = H_{M2}^2 + H_{S2}^2, \qquad \varepsilon = \frac{2H_{M2}H_{S2}}{H_{M2}^2 + H_{S2}^2}, \tag{C9}$$

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

so that

$$A(\theta) = \sqrt{a}\,\sqrt{1+\varepsilon\cos\theta}. \tag{C10}$$

The first terms of the Fourier expansion are

$$A(\theta) = \sqrt{a}\left[1 - \frac{1}{16}\varepsilon^2 - \frac{15}{1024}\varepsilon^4 + \left(\frac{1}{2}\varepsilon + \frac{3}{64}\varepsilon^3\right)\cos\theta\right. \tag{C11}$$

$$\left. + \left(-\frac{1}{16}\varepsilon^2 - \frac{5}{256}\varepsilon^4\right)\cos 2\theta + \frac{1}{64}\varepsilon^3\cos 3\theta - \frac{5}{1024}\varepsilon^4\cos 4\theta + \cdots\right]. \tag{C12}$$

This shows explicitly that the amplitude envelope contains not only the fundamental beat frequency $\Delta\omega$, but also higher harmonics $2\Delta\omega$, $3\Delta\omega$, $4\Delta\omega$, etc. In practice, the $\Delta\omega$ term is usually dominant, while the higher harmonics become more important as the two constituent amplitudes become more similar.

C2 Phase modulation

The phase correction $\psi(t)$ is also periodic in θ and is therefore likewise modulated on the difference-frequency timescale. Writing

$$\eta(t) = A(t)\cos\big(\omega_{M2}t + \phi_{M2} + \psi(t)\big), \tag{C13}$$

the phase modulation is given by

$$\psi(t) = \tan^{-1}\left(\frac{H_{S2}\sin(\Delta\omega t + \Delta\phi)}{H_{M2} + H_{S2}\cos(\Delta\omega t + \Delta\phi)}\right). \tag{C14}$$

Since $\psi(t)$ is also a nonlinear periodic function of θ, it can be expanded in a Fourier series of the form

$$\psi(\theta) = \sum_{n=1}^{\infty} b_n \sin(n\theta), \tag{C15}$$

so the phase modulation also introduces components at

$$n\Delta\omega, \qquad n = 1, 2, 3, \ldots \tag{C16}$$

Therefore, the interaction of two nearby tidal constituents such as M_2 and S_2 produces modulation not only at the primary beat frequency, but also at its higher harmonics in both amplitude and phase.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

Table A1. Tidal constituents used in the analysis, including periods and aliasing pairs. Constituents are selected based on bootstrap-derived significance and robustness from 10-minute harmonic analysis of Bath (2015–2019). Flags indicate whether constituents are also significant/robust when estimated from high- and low-water data using regular (Reg HA) and derivative-constrained (Deriv HA) harmonic analysis.

Constituent	Period (hours)	Reg HA	Deriv HA	Aliasing Pair(s)
z0	-	✓	✓	M4, M8
SA	8766.231		✓	
MSf	354.367	✓	✓	MS4
Q1	26.868	✓	✓	
RHO1	26.723			
O1	25.819	✓	✓	MK3
P1	24.066	✓		
S1	24.000			
K1	23.934	✓	✓	2MK5, MO3
OQ2	13.162			
EPS2	13.127	✓		MSN2
2N2	12.905			
MU2	12.872	✓	✓	S2, 2MS6
N2	12.658	✓	✓	L2, 2MN6
NU2	12.626			LDA2
GAM2	12.452			
H1	12.438		✓	H2
M2	12.421	✓	✓	M6
H2	12.403	✓	✓	H1
MKS2	12.386			EPS2
LDA2	12.222	✓	✓	NU2
L2	12.192	✓	✓	N2, 2MN6
T2	12.016			
S2	12.000	✓	✓	MU2, 2MS6
K2	11.967	✓	✓	2MK6
MSN2	11.786	✓		
MO3	8.386		✓	K1, 2MK5
SO3	8.192	✓		
MK3	8.177		✓	O1
SK3	7.993			
MN4	6.269	✓	✓	
M4	6.210	✓	✓	Z0, M8
MS4	6.103	✓	✓	MSf
MK4	6.095	✓		
2MK5	4.931		✓	K1, MO3
2MN6	4.166		✓	N2, L2
M6	4.140	✓	✓	M2
2MS6	4.092		✓	S2, MU2
2MK6	4.089		✓	
2SM6	4.046			
M8	3.105			Z0, M4

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

Author contributions. Conceptualization (all); Methodology development (JB, PM, SI); Data collection (JB); Data curation and quality control (JB); Formal analysis (JB, PM); Software and implementation (JB, PM, SI); Interpretation of results (all); Visualization (JB); Writing—original draft (JB); Writing—review and editing (all); Supervision (BM, TH); Project administration (TH); Funding acquisition (TH, JB)

Competing interests. The authors declare that they have no competing interests.

Acknowledgements. This research was funded by the Netherlands Organization for Scientific Research (NWO) through the Vici project "Deltas out of Shape: Regime Changes of Sediment Dynamics in Tide-Influenced Deltas" (NWO-TTW Grant 17062) (JB, DM & AH). JB also acknowledges support from the WIMEK PhD Grant 2021. The authors thank David Jay, Stefan Talke, and Paul Torfs for valuable discussions over the years on high- and low-water tidal analysis.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

References

Boesch, A. and Müller-Navarra, S.: Reassessment of long-period constituents for tidal predictions along the German North Sea coast and its tidally influenced rivers, Ocean Science, 15, 1363–1379, 2019.

Cartwright, D. E.: Tides: a scientific history, Cambridge University Press, 2000.

Cassini, J.: Sur le flux et le reflux, Mémoires de mathématique et de physique, tirés des registres de l'Académie royale des sciences, series of memoirs on the tides at French ports, read in the Académie royale des sciences in 1710, 1712, 1713, 1714, and 1720; often cited collectively as: Cassini, J. Sur le flux et le reflux, Mém. Acad. Roy. Sci. (1710, 1712, 1713, 1714, 1720)., 1710.

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

Darwin, G. H.: The harmonic analysis of tidal observations. (Report of a Committee for the Harmonic Analysis of Tidal Observations), Report of the British Association for the Advancement of Science, pp. 49–118, 1883.

Darwin, G. H.: On the harmonic analysis of tidal observations of high and low water, Proceedings of the Royal Society of London, 48, 278–340, publisher: The Royal Society London, 1891.

Doodson, A. T.: The analysis of high and low waters, The International Hydrographic Review, 1951.

Ferrel, W.: Tidal researches, US Government Printing Office, 1874.

Folkner, W. M., Williams, J. G., and Boggs, D. H.: The Planetary and Lunar Ephemeris DE 421, Interplanetary Network Progress Report 42-178, Jet Propulsion Laboratory, California Institute of Technology, https://ipnpr.jpl.nasa.gov/progress_report/42-178/178C.pdf, 2009.

Foreman, M. G. G. and Henry, R.: Tidal analysis based on high and low water observations, Institute of Ocean Sciences, Pacific Marine Science Report, pp. 79–15, 1979.

George, K. and Simon, B.: The species concordance method of tide prediction in estuaries, The International Hydrographic Review, 1984.

Godin, G., Eldring, S., and Taylor, J.: The analysis of nineteen years of observations on the high and low water with the aid of the German method, 1967.

Hoitink, A. F. and Jay, D. A.: Tidal river dynamics: Implications for deltas, Reviews of Geophysics, 54, 240–272, 2016.

Holland, P. W. and Welsch, R. E.: Robust regression using iteratively reweighted least-squares, Communications in Statistics-theory and Methods, 6, 813–827, 1977.

Horn, W.: Über die Darstellung der Gezeiten als Funktion der Zeit, Deutsche Hydrographische Zeitschrift, 1, 124–140, in German, 1948.

Horn, W.: Some recent approaches to tidal problems, The International Hydrographic Review, 1960.

Horsburgh, K. and Wilson, C.: Tide-surge interaction and its role in the distribution of surge residuals in the North Sea, Journal of Geophysical Research: Oceans, 112, 2007.

Innocenti, S. and Matte, P.: Tide Estimator - python, https://doi.org/10.5281/zenodo.17131694, 2025.

Innocenti, S., Matte, P., Fortin, V., and Bernier, N.: Analytical and residual bootstrap methods for parameter uncertainty assessment in tidal analysis with temporally correlated noise, Journal of Atmospheric and Oceanic Technology, 39, 1457–1481, 2022.

Jenkins, L. J., Haigh, I. D., Sifnioti, D. E., Rascon, J. A. P., Inayatillah, A., and Kassem, H.: Non-linear tide-surge interactions around the coast of the UK through the lens of tidal level, phase, and skew surge, Estuarine, Coastal and Shelf Science, 321, 109 323, 2025.

Ji, W. and Guohong, F.: The extraction of harmonic tidal constants from high and low waters, Chinese Journal of Oceanology and Limnology, 5, 228–240, https://doi.org/10.1007/BF02843987, 1987.

Krvavica, N., Gržić, M. M., Innocenti, S., and Matte, P.: Impact of storm surge and power peaking on tidal-fluvial processes in microtidal Neretva River estuary, Estuarine, coastal and shelf science, 318, 109 227, 2025.

https://doi.org/10.5194/egusphere-2026-2651
Preprint. Discussion started: 29 May 2026

Latapy, A., Ferret, Y., Testut, L., Talke, S., Aarup, T., Pons, F., Jan, G., Bradshaw, E., and Pouvreau, N.: Data rescue process in the context of sea level reconstructions: An overview of the methodology, lessons learned, up-to-date best practices and recommendations, Geoscience Data Journal, 10, 396–425, 2023.

Li, P., Li, L., Zuo, J., Zhao, W., and Chen, Z.: Tidal analysis of high and low water data, Journal of Ocean University of China, 3, 10–16, https://doi.org/10.1007/s11802-004-0002-2, 2004.

Lubbock, J. W.: XV. Discussion of tide observations made at Liverpool, Philosophical Transactions of the Royal Society of London, pp. 275–299, 1835.

Lubbock, J. W.: IX. On the tides, Philosophical Transactions of the Royal Society of London, pp. 97–140, 1837.

Matte, P., Jay, D. A., and Zaron, E. D.: Adaptation of Classical Tidal Harmonic Analysis to Nonstationary Tides, with Application to River Tides, Journal of Atmospheric and Oceanic Technology, 30, 569–589, https://doi.org/10.1175/JTECH-D-12-00016.1, publisher: American Meteorological Society Section: Journal of Atmospheric and Oceanic Technology, 2013.

Monahan, T., Tang, T., Roberts, S., and Adcock, T. A.: RTide: Automating the tidal response method, Journal of Geophysical Research: Machine Learning and Computation, 2, e2024JH000 525, 2025.

Moule, A. C.: The Bore on the Ch'ien-T'ang River in China, T'oung Pao, 22, 135–188, 1923.

Munk, W. and Cartwright, D.: Tidal spectroscopy and prediction, Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 259, 533–581, https://doi.org/10.1098/rsta.1966.0024, 1966.

Munk, W. and Hasselmann, K.: Super-resolution of tides, Studies on oceanography, pp. 339–344, 1964.

Pan, H., Xu, T., and Wei, Z.: A modified tidal harmonic analysis model for short-term water level observations, Ocean Modelling, 186, 102 251, https://doi.org/10.1016/j.ocemod.2023.102251, 2023.

Parke, M. E., Stewart, R. H., Farless, D. L., and Cartwright, D. E.: On the choice of orbits for an altimetric satellite to study ocean circulation and tides, Journal of Geophysical Research: Oceans, 92, 11 693–11 707, 1987.

Pugh, D., Woodworth, P. L., and Woodworth, P.: Sea-level science: understanding tides, surges, tsunamis and mean sea-level changes, Cambridge university press, 2014.

Sarkar, D., Osborne, M., and Adcock, T.: Prediction of tidal currents using Bayesian machine learning, Ocean Engineering, 158, 221–231, https://doi.org/10.1016/j.oceaneng.2018.03.007, publisher: Elsevier Ltd, 2018.

Schureman, P.: Manual of harmonic analysis and prediction of tides, 98, US Government Printing Office, 1941.

Talke, S. A. and Jay, D. A.: Archival water-level measurements: Recovering historical data to help design for the future, 2017.

Thomson, W.: Report to Committee for the purpose of promoting the extension, improvement, and harmonic analysis of tidal observations, Report of the British Association for the Advancement of Science, pp. 489–510, committee appointed 1867; first major formalization of tidal harmonic analysis., 1868.

United Kingdom Hydrographic Office: Admiralty Tide Tables, UK Hydrographic Office, Taunton, UK, annual publication, NP201–NP208 series, 2025.

Woodworth, P. L.: Differences between mean tide level and mean sea level, Journal of Geodesy, 91, 69–90, 2017.

Zetler, B. D.: A method of inferring the amplitudes of the tidal constituents M2 and (K1+ O1), The International Hydrographic Review, 1953.

Zhang, W., Cao, Y., Zhu, Y., Zheng, J., Ji, X., Xu, Y., Wu, Y., and Hoitink, A.: Unravelling the causes of tidal asymmetry in deltas, Journal of Hydrology, 564, 588–604, 2018.

Zuosheng, Y., Emery, K., and Yuix, X.: Historical development and use of thousand-year-old tide-prediction tables, Limnology and Oceanography, 34, 953–957, 1989.