

Reply to RC1 (Stephen Taylor)

Overall Assessment:

This is a strong and potentially impactful paper. The authors visit a long standing question of how to provide tidal predictions given just high and low tidal data, i.e. height and times of turning points, rather than regular time series of tidal heights. They develop and describe a modernized High–Low Tidal Analysis (HLTA) framework that operates directly on the tidal extrema (i.e. HW/LW measurements) rather than continuous water-level time series records. The paper goes on to provide a means of predicting the time series tidal values given the HW/LW predictions. I recommend acceptance with minor revisions.

We thank the referee for the positive assessment of our work and for recognizing the potential value of the proposed HLTA framework. We appreciate the constructive comments and have revised the manuscript accordingly.

Summary:

The paper is well motivated by the need to exploit such historical HW/LW datasets. The paper develops two formulations of HLTA - the long-period harmonic (LPH) and empirical–astronomical (EA) formulations and compares the formulations with results from Harmonic Analysis methods.

General Comments:

1. The manuscript repeatedly argues that HLTA can outperform harmonic analysis and is particularly advantageous in shallow, nonlinear systems. The claims of general applicability of the HLTA framework and its two formulations, LPH and EA, are perhaps somewhat overstated since the study was carried out at just two tidal stations (Bath and Melle) in the Scheldt Estuary. It is worth examining whether its predictor selection and lag structure generalise to regimes where diurnal inequality is large and the dominant modulation is not the spring-neap cycle. The paper mentions diurnal tidal regimes only in passing (section 4.1) but does not explicitly evaluate the framework for mixed or diurnal regimes. The HLTA framework as presented assumes four extrema per lunar day. An explicit statement about how the framework should be adapted, or fails, for diurnal systems would prevent misapplication. The authors should either test the framework on more contrasting tidal locations/sites or they should consider qualifying the scope in their conclusions.

We thank the reviewer for raising this important limitation. We agree that evaluating the framework only at Bath and Melle is insufficient to substantiate broad claims regarding general applicability, particularly to mixed and diurnal tidal regimes. Our intention in selecting these two locations had been to test the framework under particularly challenging conditions characterized by strong shallow-water distortion. Nevertheless, we agree that including additional sites with contrasting tidal characteristics improves the transparency and generality of the analysis.

We have therefore extended the evaluation in the revised manuscript to a broader set of 15 stations representing predominantly semidiurnal, mixed, and diurnal regimes, as well as systems with different degrees of tidal distortion and non-stationarity. We also clarify how the extrema representation is adapted when fewer than four extrema occur per lunar day.

2. The claim of 38% improved timings and equivalent comparable heights for LPH relative to HA of 10 minute records at Bath, is certainly noteworthy. However it too would benefit from further comparative analysis at more locations. Regarding (line 517) the comparisons with Harmonic Analysis as directly applied to HW/LW values, it is well known that HA works poorly on HW/LW records because, to put it crudely, the harmonic method does not "know" that the data points are turning points, losing a most valuable piece of information. The statement that "both HLTA approaches substantially outperform harmonic analysis applied directly to HL observations" (line 518) is therefore not surprising. The authors repeatedly labour this point in varying detail throughout the paper (see lines 7-9, 34-37, 84-88, 112-119, Section 3 128-184, 518 referring to the inaccuracy of HA to analyse HW/LW records. Perhaps this point is too frequently made. Fortunately, the study does indeed make the more reasonable comparison between HLTA and with HA using a ten minute time series in Figure 5 and Line 495.

We agree that the limitations of applying harmonic analysis directly to HW/LW observations may have been emphasized too frequently, and we reduced the repetition throughout the manuscript.

We nevertheless retain this comparison because, as also noted by RC2, harmonic analysis applied to HW/LW observations—particularly when derivative constraints are imposed—can provide a meaningful and reasonably competitive benchmark for some gauges. The inadequacy of this approach is therefore not entirely self-evident and merits quantitative evaluation.

Accordingly, we retained both the unconstrained and derivative-constrained harmonic analysis applied to HW/LW observations, but present these results more concisely. Greater emphasis is placed on the comparison with harmonic analysis based on continuous records, for which the improved timing predictions obtained with LPH-HLTA are among the most noteworthy findings of the study.

Specific Comments:

1. Perhaps the paper understates the extent to which the study relies on the work of Horn 1948 and Lubbock 1837. It states "This idea builds on" . A better statement would be that the method is directly derived or a direct descended from the work of those earlier authors. Line 39.

We agree that the original phrase "builds on" understated the extent to which the proposed framework derives from the earlier work of Lubbock and Horn. We revised the text to state explicitly that the framework is directly descended from the extrema-based formulations developed by Lubbock (1835) and formalized in harmonic form by Horn (1948).

2. The choice of 5-year calibration (line 360, Sec 5.5) (2015–2019) and 2-year validation (2020–2021) is reasonable, but it would be worth noting whether the validation period is anomalous in any

way — for instance, the COVID-era may affect shipping and fluvial flow and in principle influence surge residuals.

We do not expect the COVID-19 period to have directly affected the astronomical tidal signal, although associated changes in river discharge or meteorological forcing may have influenced the non-tidal variability. With the addition of 13 gauges, the calibration and validation periods have shifted and now differ among stations. Based on visual inspection of the records, we found no indication that the selected calibration or validation periods are anomalous. Furthermore, in response to RC2, we have included an additional sensitivity test to assess the effects of using calibration periods that differ in both timing and length.

3. It would be useful to know what is the likelihood that the results of the comparisons on Figure 5, are or are not statistically significant. What is the margin of error on RMSE and ME values. Would these benefit from wider testing.

We agree that quantifying the uncertainty associated with the reported RMSE and ME values could provide additional information when differences between methods are small. However, we have chosen not to include confidence intervals or margins of error for all RMSE and ME values in the revised manuscript. Following the extension of the evaluation from 2 to 15 stations, reporting uncertainty bounds for every station–method combination would substantially increase the width and complexity of the tables and make the corresponding figures considerably more cluttered. In addition, obtaining robust uncertainty estimates through, for example, 1000 seasonally constrained monthly-block bootstrap realizations for each combination of station and method would add substantial computational cost. We have therefore not implemented this analysis for the complete set of RMSE and ME results.

Instead, the revised evaluation provides a substantially broader assessment of the robustness of the comparisons by including 15 stations spanning different tidal regimes and environmental settings. This allows us to assess whether differences between methods are consistent across stations rather than being specific to the two stations considered previously. In addition, the sensitivity analyses (to calibration period length and noise) explicitly quantify uncertainty using bootstrap-derived confidence intervals. Together, the expanded multi-station evaluation and these uncertainty-aware sensitivity analyses provide a broader indication of the robustness of the reported differences, although we refrain from interpreting the RMSE and ME differences themselves as formal tests of statistical significance.

4. Despite the length of the paper, I feel it lacks sufficient detail for a reader to readily duplicate the LPH and EA methods, and to evaluate its performance on their data. This is after all a basic requirement of the scientific method - the ability to duplicate results. It leaves the reader somewhat confused over the details of method used. For example in section 4.3, it is not clear whether these empirical factors are station-specific or whether they are stable across calibration sub periods. Therefore, Section 4 would benefit from more detailed description of the method and role of the algebraic terms.

We agree that the methodological description, particularly for the EA-HLTA formulation, should be more explicit to allow independent reproduction and application at other sites. In the revised manuscript, we have therefore expanded the description of both the LPH- and EA-HLTA formulations, including a clearer step-by-step explanation of the workflow, the construction of the predictor sets, the role of the main algebraic terms, and the distinction between common model structure and station-specific fitted parameters. In particular, we now state explicitly that the time offset used in the EA formulation is station-specific and estimated from the calibration data. We also clarify how model complexity is adapted to record length, with details of the predictor-selection procedure provided in the Methods and Supplement.

In addition, following the suggestions of both reviewers, we have added a more explicit discussion of the practical applicability of the two approaches. The revised manuscript now discusses the relative suitability of LPH-HLTA and EA-HLTA in light of the record-length sensitivity analysis, the distinction between stationary and non-stationary applications, and the availability of different types of observations. To further improve clarity, we have also added a schematic figure summarizing the conceptual HLTA workflow, from extrema referenced to lunar transit through prediction and optional continuous reconstruction.

5. The Github repository download furthermore does not contain adequate instructions on how to run the code and the various input data requirements. It would be helpful if the Github download readme file contained an overview and sufficient instructions to enable it to be readily run. See https://github.com/JorisBeemster/HighLow_TA.

We agree that the current GitHub repository could provide clearer guidance for reproducing and applying the analysis. In the revised version, we improved the README and repository documentation by adding a clearer overview of the workflow, the main input-data requirements, and basic instructions for running the LPH- and EA-HLTA analyses. We also include example material where useful. Together with the additional methodological clarification in the manuscript, this should make the framework easier to follow and apply to new datasets.

Technical Corrections

1. Line 2 is incorrect in implication. Is it meant to mean "This method suffers from poor accuracy when applied to ...

We thank the referee for pointing out this ambiguity. We agree that the original wording may be misleading and have revised it to state that harmonic analysis may suffer reduced accuracy when applied to sparse HW/LW observations.

2. Line 51 IRLS not yet defined

We have now provided the full name, iteratively reweighted least squares (IRLS), and the associated reference.

3. Line 145 Define what is meant by "minimum frequency distance"

We can see that “minimum frequency distance” may be confusing and perhaps was insufficiently defined. We revised the text to refer explicitly to the apparent frequency separation between aliased constituents after sampling, which becomes zero for exact alias pairs:

"In contrast, if the sampling interval coincides with sums or differences of tidal frequencies (as in HW/LW records, or sun-synchronous satellite missions), two constituents can alias onto the same apparent frequency. Their frequency separation after sampling is then zero, implying an infinite record length for separation under the Rayleigh criterion."

4. The manuscript alternates between: "high-low observations", "HW/LW observations", "extrema observations". Consider standardizing terminology throughout.

After initially defining “high water–low water (HW/LW)” we now use “HW/LW” consistently throughout the manuscript.

Conclusion:

This paper makes a valuable methodological contribution. The primary weaknesses of the study are not methodological flaws but rather limitations in the breadth of validation. Thus the authors should consider including more contrasting tidal locations/sites or they should consider qualifying the scope of the method in the conclusion.

Recommendation:

I would therefore suggest the paper is accepted subject to a number of minor revisions which address the above issues.