



Residual bootstrap for tidal model uncertainty analysis

S. Innocenti, P. Matte, V. Fortin, N. Bernier

vEGU 2021

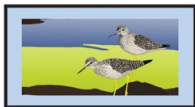


Reconstruction of tidal signals

Water level predictions
and long term analysis

**flood forecasting
risk mapping / mitigation**

**wetland/estuary
conservation**



ecosystem functions



**calibrating/verifying
altimetry products**



freshwater fisheries



river navigation

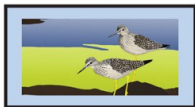
Image adapted from <https://swot.nasa...>

Reconstruction of tidal signals

Water level predictions
and long term analysis

**flood forecasting
risk mapping / mitigation**

**wetland/estuary
conservation**



ecosystem functions



**calibrating/verifying
altimetry products**



freshwater fisheries



river navigation

**Harmonic Analysis (HA)
of water levels**

Image adapted from <https://swot.nasa...>

HA model

[Foreman and Henry, 1989]

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

h_t : surface elevation

τ : time evaluated at time steps $t = 1, 2, \dots, T$

ω_k : harmonic frequencies of constituents $k = 1, 2, \dots, K$

⇒ **tidal constituent selection**

and estimation of $(2 \times K + 1)$
unknown parameters

HA model

[Foreman and Henry, 1989]

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

h_t : surface elevation

τ : time evaluated at time steps $t = 1, 2, \dots, T$

ω_k : harmonic frequencies of constituents $k = 1, 2, \dots, K$

ϵ_t : error $\Rightarrow$ **unresolved component of the signal**

$\Rightarrow$ **tidal constituent selection**

and estimation of $(2 \times K + 1)$
unknown parameters

HA model

[Foreman and Henry, 1989]

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

h_t : surface elevation

τ : time evaluated at time steps $t = 1, 2, \dots, T$

ω_k : harmonic frequencies of constituents $k = 1, 2, \dots, K$

ϵ_t : error $\Rightarrow$ **unresolved component of the signal** $\Rightarrow$ **autocorrelated**

$\Rightarrow$ **tidal constituent selection**

and estimation of $(2 \times K + 1)$
unknown parameters

Interaction between tides,
weather, hydrology, and
other climatological processes

$\Rightarrow$ **substantial residual energy**

HA regression uncertainty

Conventional methods rely on simplifying assumptions to estimate the HA regression variance-covariance matrix

[e.g., normal, independent, and/or homoskedastic residuals]

Analytical definition of the variance-covariance estimator for:

- ◇ UTide [Codiga, 2011]
- ◇ IRLS [Leffler and Jay, 2009; Huber and Ronchetti, 2011]
- ◇ NS_tide [Matte et al., 2013]

⇒ **Obj. 1.** Review each method statistical properties and evaluate their implications in HA uncertainty assessments:

<https://youtu.be/Nqd7qMKJrw0> [4:34 min]

Residual bootstrap for estimating the HA regression uncertainty

No or weak assumptions about the error term distribution to produce random replicates of the observed residuals

[account for the coloured nature of noise signals]

Definition of two resampling algorithms:

- ◇ Moving Block Bootstrap (MBB) with fixed length blocks
- ◇ Semi-Parametric Bootstrap (SPB) based on the residual spectrum

⇒ **Theoretical comparison of conventional and bootstrap methods:** <https://youtu.be/voEg6NEgT8M> [3:23 min]

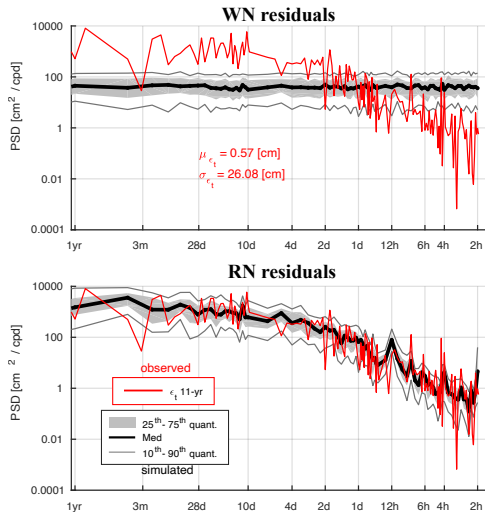
Monte-Carlo Experiment

Simulate a large number of water level series with known statistical properties

- Two sets of simulations corresponding to **two spatial locations** in the St. Lawrence Estuary, Canada.
- 68 tidal constituents with **known amplitudes and phases** (from 11-yr records) to generate the tidal signal.
- Various series lengths at **hourly** resolution and trimmed to contain **high-low water levels**.
- Two **residual structures**:
 - ◇ White Noise [WN], homoskedastic uncorrelated,
 - ◇ Resampled noise [RN], heteroskedastic autocorrelated.

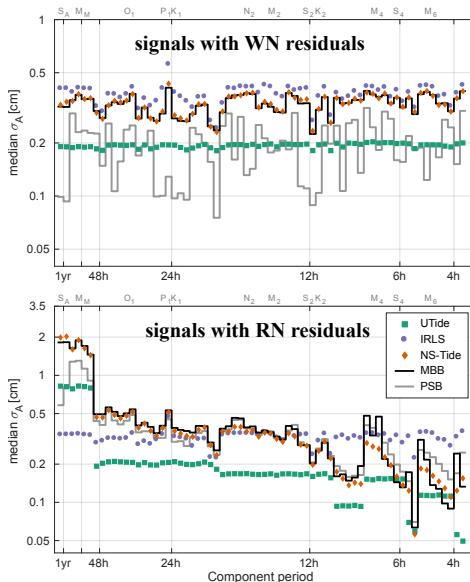
Monte-Carlo Experiment

Power Spectral Density (PSD) of simulated residual series



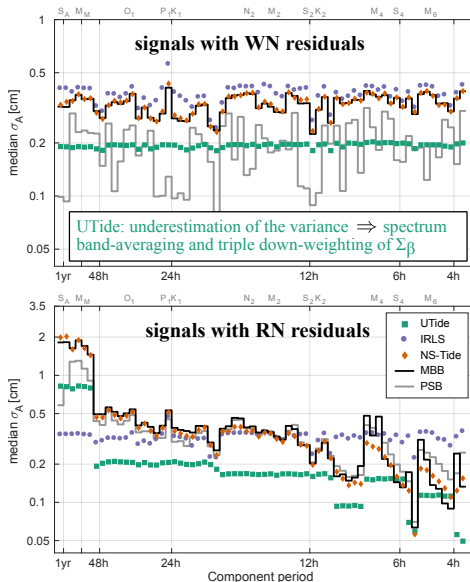
Results

Standard errors of tidal amplitudes



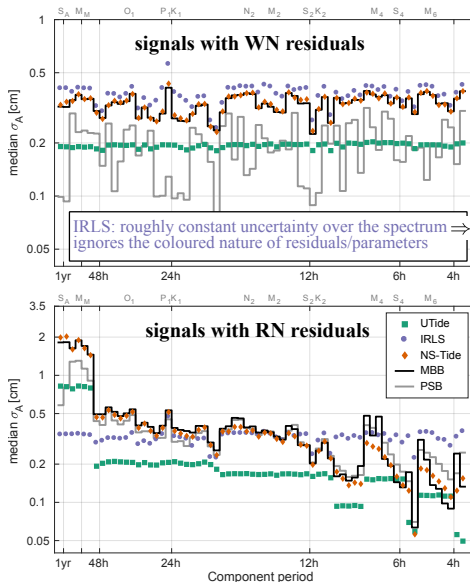
The robust IRLS were used to estimate the regression coefficients for all the uncertainty assessment methods (same β s for all methods)

Standard errors of tidal amplitudes



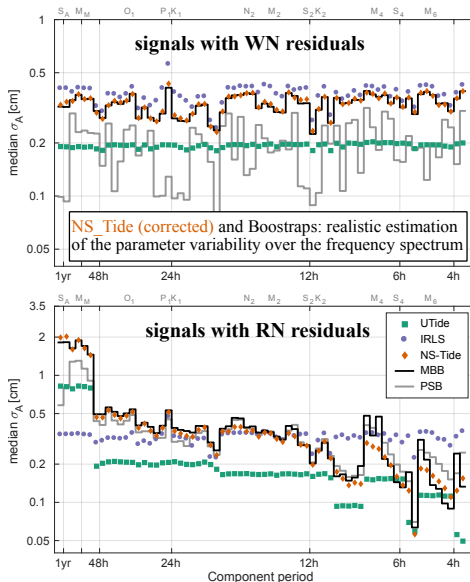
The robust IRLS were used to estimate the regression coefficients for all the uncertainty assessment methods (same β s for all methods)

Standard errors of tidal amplitudes



The robust IRLS were used to estimate the regression coefficients for all the uncertainty assessment methods (same β s for all methods)

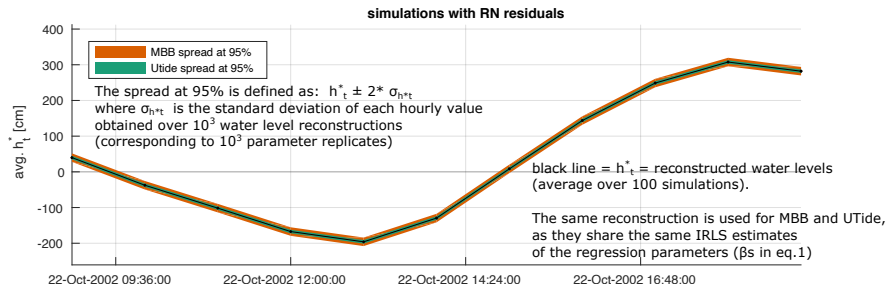
Standard errors of tidal amplitudes



The robust IRLS were used to estimate the regression coefficients for all the uncertainty assessment methods (same β s for all methods)

Uncertainty of reconstructed series

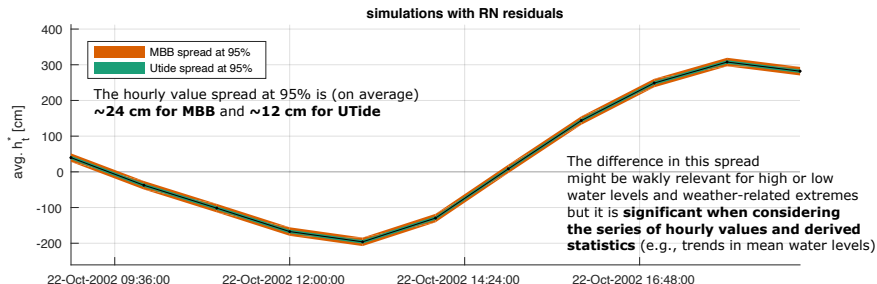
Conventional methods underestimate the variability of water level predictions based on HA



⇒ Calibration and validation of satellite altimetry products,
analysis of sea-level rise, etc.

Uncertainty of reconstructed series

Conventional methods underestimate the variability of water level predictions based on HA



⇒ Calibration and validation of satellite altimetry products, analysis of sea-level rise, etc.

Constituent selection in High-Low water series analysis

Using a large number of simulations with known constant statistical properties (i.e., same true parameters and residual structure), it is possible to evaluate the performance of the Signal-to-Noise Ratio (SNR) criterion for selecting the "significant" tidal constituents

Constituent selection in High-Low water series analysis

SNR selection

It is intended to
select the constituents based on
their relative uncertainty
(~ CI width vs parameter value)

**True if
good
coverage**



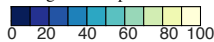
% of simulations excluding the
constituent due to $SNR < 2$



Confidence Interval (CI) coverage

By definition, the CI confidence
level (e.g., $1 - \alpha = 95\%$) must
correspond to the proportion of intervals
containing the true parameter value
over a large number of the interval replicates

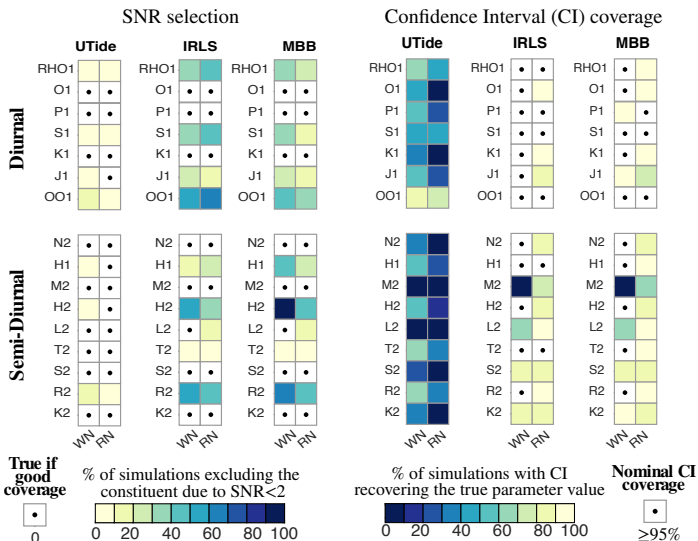
% of simulations with CI
recovering the true parameter value



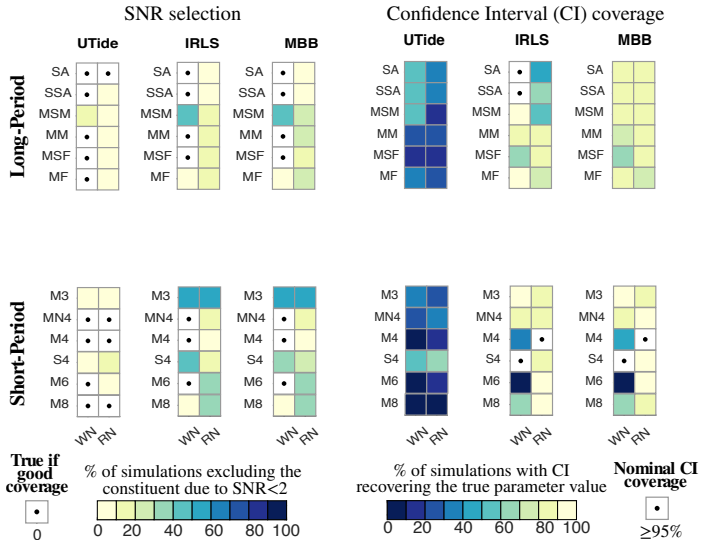
**Nominal CI
coverage**



Constituent selection in High-Low water series analysis

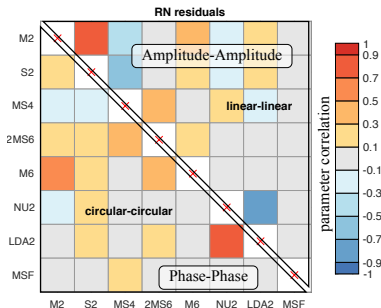


Constituent selection in High-Low water series analysis



Correlation between model constituents

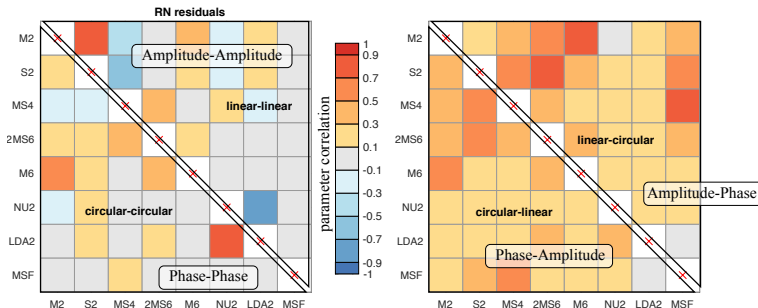
One symptom of inaccurate constituent estimation or selection is the association between the HA model parameters



- bootstrap methods easily provide information on the inter-dependencies among constituents **without** stringent parametric hypotheses
- **circular statistics** to assess the association between phases ... and between phases and amplitudes

Correlation between model constituents

One symptom of inaccurate constituent estimation or selection is the association between the HA model parameters



Spectral leakage due to sparse sampling
 ⇒ physical relevance of the constituent interactions ?

Summary and discussion

Summary

**Bootstrap methods for the inference of tidal HA regressions:
integrate the residual autocorrelation in the model evaluation.**

1. Evaluate the theoretical differences between conventional variance-covariance estimations

2. Assess the impacts of ignoring the residual autocorrelation: parameter variability and CI

3. Examine the constituent selection strategies relying on these uncertainty assessments

- ◇ Multiple use of the weighted residuals: UTide and NS_Tide original.
- ◇ Inconsistent residual PSD estimation.
- ◇ Uncorrelated Gaussian residuals: IRLS.
- ◇ UTide underestimates the parameter variances for all residual structures
- ◇ The IRLS provide adequate estimations for WN residuals and/or long records.
- ◇ NS_Tide, MBB, and SPB provide realistic estimates and adequate CI for most constituents.
- ◇ The SNR provides limited information.
- ◇ Diagnostics based on resampling can describe the inter-dependencies among the parameters (e.g., correlations).

Summary

**Bootstrap methods for the inference of tidal HA regressions:
integrate the residual autocorrelation in the model evaluation.**

1. Evaluate the theoretical differences between conventional variance-covariance estimations

2. Assess the impacts of ignoring the residual autocorrelation: parameter variability and CI

3. Examine the constituent selection strategies relying on these uncertainty assessments

- ◇ Multiple use of the weighted residuals: UTide and NS_Tide original.
- ◇ Inconsistent residual PSD estimation.
- ◇ Uncorrelated Gaussian residuals: IRLS.
- ◇ UTide underestimates the parameter variances for all residual structures
- ◇ The IRLS provide adequate estimations for WN residuals and/or long records.
- ◇ NS_Tide, MBB, and SPB provide realistic estimates and adequate CI for most constituents.
- ◇ The SNR provides limited information.
- ◇ Diagnostics based on resampling can describe the inter-dependencies among the parameters (e.g., correlations).

Summary

**Bootstrap methods for the inference of tidal HA regressions:
integrate the residual autocorrelation in the model evaluation.**

1. Evaluate the theoretical differences between conventional variance-covariance estimations

2. Assess the impacts of ignoring the residual autocorrelation: parameter variability and CI

3. Examine the constituent selection strategies relying on these uncertainty assessments

- ◇ Multiple use of the weighted residuals: UTide and NS_Tide original.
- ◇ Inconsistent residual PSD estimation.
- ◇ Uncorrelated Gaussian residuals: IRLS.
- ◇ UTide underestimates the parameter variances for all residual structures
- ◇ The IRLS provide adequate estimations for WN residuals and/or long records.
- ◇ NS_Tide, MBB, and SPB provide realistic estimates and adequate CI for most constituents.
- ◇ The SNR provides limited information.
- ◇ Diagnostics based on resampling can describe the inter-dependencies among the parameters (e.g., correlations).

Summary

**Bootstrap methods for the inference of tidal HA regressions:
integrate the residual autocorrelation in the model evaluation.**

1. Evaluate the theoretical differences between conventional variance-covariance estimations

2. Assess the impacts of ignoring the residual autocorrelation: parameter variability and CI

- ◇ Multiple use of the weighted residuals: UTide and NS_Tide original.
- ◇ Inconsistent residual PSD estimation.
- ◇ Uncorrelated Gaussian residuals: IRLS.
- ◇ UTide underestimates the parameter variances for all residual structures
- ◇ The IRLS provide adequate estimations for WN residuals and/or long records.
- ◇ NS_Tide, MBB, and SPB provide realistic estimates and adequate CI for most constituents.

**Differences up to 50% in the spread of
the distributions of predicted water levels**

Summary

**Bootstrap methods for the inference of tidal HA regressions:
integrate the residual autocorrelation in the model evaluation.**

1. Evaluate the theoretical differences between conventional variance-covariance estimations

2. Assess the impacts of ignoring the residual autocorrelation: parameter variability and CI

3. Examine the constituent selection strategies relying on these uncertainty assessments

- ◇ Multiple use of the weighted residuals: UTide and NS_Tide original.
- ◇ Inconsistent residual PSD estimation.
- ◇ Uncorrelated Gaussian residuals: IRLS.
- ◇ UTide underestimates the parameter variances for all residual structures
- ◇ The IRLS provide adequate estimations for WN residuals and/or long records.
- ◇ NS_Tide, MBB, and SPB provide realistic estimates and adequate CI for most constituents.
- ◇ The SNR provides limited information.
- ◇ Diagnostics based on resampling can describe the inter-dependencies among the parameters (e.g., correlations).

Conclusion

Resampling methods can be applied to different noise structures and account for several aspects of the error distribution

Provide a more accurate estimation of the HA uncertainty for applications that need cm or higher accuracy water level reconstructions:

- Tides into future climate and sea-level trends ↘
- Calibrate and validate satellite measurements → future research
- Support the analysis of sparsely sampled records ↗

Conclusion

Resampling methods can be applied to different noise structures and account for several aspects of the error distribution

Provide a more accurate estimation of the HA uncertainty for applications that need cm or higher accuracy water level reconstructions:

- Tides into future climate and sea-level trends ↘
- Calibrate and validate satellite measurements → future research
- Support the analysis of sparsely sampled records ↗

Limitations and Future Work:

- ◇ Correlated residuals may result in biased estimators of the point regression coefficients [i.e., the β 's]
- ◇ Resampling methods are not suitable for the analysis of extreme events and/or with strong non-stationarity of the residuals



Environnement et
Changement climatique Canada

Environment and
Climate Change Canada

Thank you
silvia.innocenti@canada.ca



Canada

Bibliography

- Codiga [2011], *Unified tidal analysis and prediction using the UTide Matlab functions*, *Graduate School of Oceanography*
- Foreman and Henry [1989], "The harmonic analysis of tidal model time series", *Advances in Water Resources*
- Huber and Ronchetti [2011], *Robust Statistics*, *Wiley*
- Leffler and Jay [2009], "Enhancing tidal harmonic analysis", *Continental Shelf Research*
- Matte et al. [2013], "Adaptation of Classical Tidal Harmonic Analysis to Nonstationary Tides, with Application to River Tides", *Journal of Atmospheric and Oceanic Technology*
- Pawlowicz et al. [2002], "Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE", *Computers & Geosciences*

PICO slide:

Residual bootstrap for tidal model uncertainty analysis

Innocenti, Matte,
Fortin, and Bernier

Unresolved variability $\Rightarrow$ Autocorrelated HA residuals

1. Review the statistical properties of commonly used HA methods, e.g.,

- ◇ UTide^[1] and T_Tide^[2]
- ◇ IRLS
- ◇ NS_Tide^[3]

Unified notation

[1] Codiga, 2011; [2] Pawlowicz et al. 2002
[3] Matte et al., 2013

Band-averaged and "line-decimated" residual spectrum

Triple down-weighting of the variance-covariance matrix

 Environnement et
Changement climatique Canada

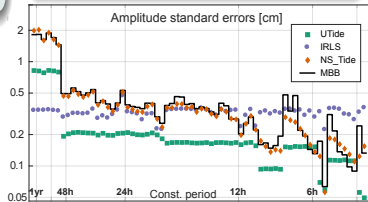
2. Compare the methods with two resampling strategies using simulated series [known properties]

- ◇ Moving Block Bootstrap
 - ◇ Semi-Parametric Bootstrap
- Parameter variances
Confidence Interval coverage

3. Performance with sparsely sampled series

- ◇ High-Low Water Levels

Constituent selection,
variance of water levels



The bootstraps make better estimation and use of:

- ◇ Parameter correlations
- ◇ Derived statistics (e.g., variance of predicted values)

<https://youtu.be/Ye4f7x9QjtA> [2:00 min]