18 November 2018

Journal: ASCMO

Title: An Integration and Assessment of Multiple Covariates of Nonstationary Storm Surge Statistical Behavior by Bayesian Model Averaging

Author(s): Tony Wong

MS No.: ascmo-2018-16

Dear Editors and Reviewers,

Thank you for the insightful and constructive reviews. Your suggestions have greatly helped to improve the quality and clarity of the manuscript. Below I address the open issues. I have formatted the reviewer comments in ​black​ font ​and my responses in ​blue​ text and blockquote indented, in order to better distinguish them from the comments. I have also included a tracked changes version of the manuscript, appended to this letter.

Again, thank you all for your helpful feedback, and for your service to our community.

Sincerely,

Tony Wong

**Reviewer 1**

This article provides a non stationnary storm surge statistical model, which allows to evaluate the effect of choosing different cvariate variables such as sea level rise of teleconnection patterns such as NAO. I think that the method is wel applied and that the topic offers a good example where such approaches are appropriate. I also think that the methods are well explained. This paper could become therefore a strong contribution to the field.

I think that the paper could be improved by extending the discussions on the implications of this work. I could suggest, for example, to better discuss how the proposed approach can support coastal risks assessments. This could be done for example by providing examples of sequences of storms that affected Norfolk and were probably not to be expected with stationnary storm surge analysis. Another option could be to extend the discussion, just briefly included in section 3.4, about the impacts of choosing a particular model structure for coastal zones management. Finally, I think that a discussion on the reasons why different covariate variable perform well in Norfolk could be included. For example, a physical argument explaining why the Temperature covariate performs well would be useful. Is this because Temperature is correlated with NAO?

I hope it is useful.

Very useful indeed! I have added some text to the discussion:

*“For Norfolk and the surrounding area, a difference of about 23 cm in the median estimated return level can lead to millions of dollars in potential damages (Fugro Consultants, 2016). Thus, the present work also serves to demonstrate the potential risks associated with selection of a single model structure.”*

Fugro Consultants, Inc . Lafayette River Tidal Protection Alternatives Evaluation. Work Order No. 7. Fugro Project No. 04.8113009. City of Norfolk, City-wide Coastal Flooding Contract, 2016, 56 pp. https://www.norfolk.gov/DocumentCenter/View/25170 Google Scholar

The reviewer is also quite correct that the potential covariate time series considered here are likely correlated (for example, sea level and temperature are certainly highly correlated). The effects of these correlations are an interesting avenue for future study, but beyond the scope of this work. I have added this note to the Discussion:

*“The accounting and propagation of uncertainty* ***and correlation*** *in the covariate time series would be an interesting avenue for future study, but is beyond the scope of this work.* ***For example, temperature may drive changes in both sea levels and NAO index, so future work might consider disentangling the effects of the multiple covariate time series.****”*

Typo

- page 6 line 3: a word is missing

Quite right, thank you for pointing this out! It has been fixed:

*“For the NAO index covariate time series, I* ***use*** *the historical monthly NAO index data from Jones et al. (1997) …”* (bold-face denotes the missing word)

- consider how useful table A1 is (only 1 model is reported)

I can certainly appreciate that it seems a bit silly to include only a single model in a table! But, it is a requirement of using the CMIP models that the one(s) used are reported in a table such as this, giving proper credit to the modeling group.

- consider reducing the number of significant digits in table A2

I thank the reviewer for the nice suggestion – I have modified the table to be rounded to the nearest centimeter.

**Reviewer 2**

This manuscript presents an approach to integrate multiple non-stationary statistical storm surge models to examine the potential structural uncertainties in coastal hazard projections. In the author’s words, “The main contribution of this work is to demonstrate the ability of the BMA approach to incorporate multiple covariate time series into flood hazard projections, and to examine the impacts of different choices of covariate time series.” Overall I find the paper to be well-written, with a clear motivation and novel methodology for addressing the deep uncertainties underlying extreme sea-level projections. The use of multiple covariates to modulate future sea-level distributions helps to advance our understanding of sea level model uncertainties.

I find this manuscript suitable for publication, with a few minor notes for the author to consider. Many of the notes I have provided are stylistic points with the intent of improving clarity, conciseness, and readability. I have noted the page and line numbers, where applicable:

1. Section 1 - Introduction

The author establishes a good background about why storm surge levels are important for flood risk management, and the role statistical models (in general) play in constraining them. It would be good to note specifically why extreme events are of such concern and how extreme value statistics can derive insights from limited amounts of data.

I thank the reviewer for the many helpful comments and feedback. Here, I have added to the text:

*“In particular, environmental extremes can often carry high risks in terms of widespread damages and economic losses (e.g., Oddo et al., 2017), but extremes are by definition rare, imposing strict limitations on the available data. Extreme value statistical models offer an avenue to estimate extremes with relatively fewer parameters to constrain, as compared to processed-based models.”*

2. Page 2, Line 21

The inclusion of a sentence describing “process-based modeling” feels somewhat out of place in the extended discussion of statistical modeling.

This is a fair and good point. I only mention process-based modeling to give mention to the alternative approaches that are also commonly used to estimate storm surge hazard. I feel it is important to keep this note in the manuscript so it does not send the message that this is the only, or the “correct”, approach to tackling this problem.

3. Page 3, Line 1-2

“In the limit of large sample sizes, the GEV distribution is the limiting distribution…”. The phrasing here wasn’t immediately clear. Is the author saying that GEV distributions are often used in the absence of large amounts of data?

My apologies for the poor wording here. I mean that a GEV distribution is the distribution one will obtain if one samples sets of block maxima from some population, in a similar sense that the Central Limit Theorem gives a normal distribution as the limit distribution for a sequence of block means. I have reworded this sentence and the surrounding discussion to be more clear:

*“The GEV distribution is the limiting distribution of a convergent sequence of independent and identically-distributed sample maxima (Coles, 2001). In extreme sea level analyses, data are frequently binned into sample blocks and a GEV distribution is assumed as the distribution of the appropriately detrended and processed sample block maxima (where this processing serves to achieve independent and identically-distributed sample block maxima). Depending on block sizes (typically annual or monthly), this approach places strict limitations on the available data. By contrast, the PP/GPD modeling approach can yield a richer set of data by making use of all extreme sea level events above a specified threshold (e.g., Arns et al., 2013; Knighton et al., 2017). Additionally, previous studies have demonstrated the difficulties in making robust modelling and processing choices using a GEV/block maxima approach (Ceres et al., 2017; Lee et al., 2017).”*

4. Page 4, Line 7

Change “Sect. 5” to Section 5

It is my understanding that sections should be referred to as “Section 5” (e.g.) when appearing at the beginning of a sentence, and as “Sect. 5” when appearing in the middle of a sentence. From the style guide [linked here](https://www.advances-statistical-climatology-meteorology-oceanography.net/for_authors/manuscript_preparation.html).

5. Section 2 – Methods

The author may want to consider re-arranging the methods sections slightly so that the data source and description (Section 2.2 Data) comes before the discussion of the Extreme value model (Section 2.1). At the very least, a brief overview of the location and record length would be helpful context for the reader, so that they better understand what data is being de-trended and manipulated in Section 2.1.

I thank the review for the helpful suggestion. I have rearranged the two sections and added the distinction between the covariate time series data and the tide gauge data at the beginning of the new Section 2.2:

*“First, to detrend the raw hourly* ***tide gauge*** *sea level time series…”*

6. Page 5, Line 18

“1:100 [year] storm surge return level” may be missing a word. Elsewhere the terminology used is 100-year – should keep it consistent.

Thank you for the important suggestion! This has been revised as suggested:

*“… 1:100-year storm surge return level…”*

7. Page 6, Line 3

Missing word: “For the NA index covariate time series, I [ ] the historical…”

Quite right, thank you for pointing this out! It has been fixed:

*“For the NAO index covariate time series, I* ***use*** *the historical monthly NAO index data from Jones et al. (1997) …”*  (bold-face denotes the missing word)

8. Page 6, Line 7

Is the temperature covariate time series anomalies? If so please state the relative date range.

I thank the review – this is an excellent point to make clear. The data are provided as anomalies relative to the 20th century mean. I have added the following text to make this distinction:

*“For the temperature time series, I use historical annual global mean surface temperature data from the National Centers for Environmental Information data portal (NOAA, 2017a), and as projections I use the CNRM-CM5 simulation (member 1) under Representative Concentration Pathway 8.5 (RCP8.5) as part of the CMIP5 multi-model ensemble (http://cmip-pcmdi.llnl.gov/cmip5/).* ***The time series are provided as anomalies relative to their 20th century mean.****”* (Bold-faced text is added)

9. Page 9, Lines 7-10

The inclusion of global mean temperature appears to widen the uncertainties, which the author potentially attributes to the comparatively larger signal produced by this covariate. Would a local temperature anomaly produce a more functional covariate than the global mean sea level used in this study?

This is a great question! The results and argument set forth by Grinsted et al. (2013; doi: 10.1073/pnas.1209980110) indicate that local mean temperature (or main development region temperatures) are potentially more powerful covariates for the storm surge statistical model (their Table 2). Those authors point out that a key strength of using global climate variables as covariates is that they are less sensitive to subtle regional patterns that process models may have difficulty accurately representing.

I have added the following comment to address this issue in the Discussion:

*“As demonstrated by Grinsted et al. (2013), the use of local temperature or other covariate information may also lead to better constraint on storm surge return levels, but also presents challenges for process models to reproduce potentially complex spatial patterns.”*

10. Page 10, Figure 2 caption

It was not immediately obvious the difference between the green and gray shaded bars below the density functions. It would be helpful to make that more explicit in the caption (as appears to be done in the Figure 4 caption).

Thank you for the helpful suggestion. I have added the following text to the caption to make this clear:

*“… for both the stationary model (green) and Bayesian model averaging ensemble (gray).”*

11. Page 13, Line 11-13

The introduction sentence to this paragraph is attempting to highlight one of the key insights yet the meaning is somewhat unclear, at least to this reviewer.

Thank you for another helpful suggestion. I have revised the beginning of this paragraph as follows:

*“These results are in agreement with the work of Lee et al. (2017) and highlight the importance of carefully considering the balance of model complexity against data availability. Including more complex physical mechanisms into model structures (i.e., nonstationary storm surges) is often important for decision-making, but additional model processes and parameters require more data to constrain them (Wong et al., 2018).”*

12. Page 13, Lines 20-22

The author discusses the increases in the center (2.13 -> 2.36) and 95th percentile upper tail (2.58 -> 3.07) of flood hazard projections when comparing the BMA to the stationary model. In the previous section (Section 3.4), however, the ranges given are the 90% credible range which appears to produce similar numbers (2.36 median and 3.07 upper range). It is somewhat confusing why the switch was made to discussing the 95th percentile here.

This is a great suggestion to help clear up some confusion. The 90% credible range is the same as the 5-95% credible range, so the 95th percentile (3.07 m) is exactly the upper bound of the 90% credible range. To clear up this confusion, I have revised the text throughout the manuscript to refer to the 90% credible interval as instead the 5-95% credible interval.

Is this comparison (the full multi-modal, multi-covariate BMA model vs. the fully stationary model) depicted directly in the figures? To my understanding it appears to be visually comparing the green shaded bar from Figure 2 with the gray shaded bar from Figure 4.

Yes, the reviewer’s interpretation of this comparison is correct, and suggests a good place to make the discussion a bit more clear:

*“Using the stationary model leads to a distribution of 100-year flood level with a median of 2.13 m and upper tail (95th percentile) of 2.58 m. Using the full multi-model, multi-covariate BMA model, however, substantially raises both the projected center (2.36 m) and upper tail (3.07 m) of the distribution of 100-year flood hazard in 2065, relative to using a stationary model.”*

13. Section 4 Discussion

The author goes a good job outlining the main insights, broader impacts, and the caveats and limitations of this work.

Thank you for the positive and constructive feedback!