Review response (#2): Enabling dynamic modelling of global coastal flooding by defining storm tide hydrographs

Job C. M. Dullaart, Sanne Muis, Hans de Moel, Philip J. Ward, Dirk Eilander, and Jeroen C. J. H. Aerts

Reviewer #2 (Summary):

The manuscript present a methodology to generate storm tide hydrographs at global scale by combining tidal cycles (average and spring) with an average storm surge hydrograph. The average storm surge hydrograph is developed by selecting first extreme storm surges, normalizing the time series of these events, calculating their average value at each time step and the duration before and after the storm peak. On the other hand, two average tidal cycles are calculated, namely the average and spring tidal cycles. The effect of the tide-surge interaction can also be included when combining the storm surge and tidal hydrographs by including the mean time offset between the two peaks.

The manuscript is well written and structured; the topic is relevant and the results and findings are interesting and relevant. However, I think that there are a couple of points that should be addressed and I also have some minor comments.

Authors' response

We would like to thank the reviewer for the thorough review and useful comments, which we believe have helped us greatly to improve our manuscript. We are pleased to read that the reviewer considers the manuscript to be well-written and structured, and the presented work to be interesting and relevant. Following the reviewer's suggestions, we have revised our manuscript. In the following sections we respond to each of the reviewer's remarks or questions. Our response is in italic.

Reviewer #2 (Main comments):

My main concern with this methodology is that it only focus on generating average hydrographs and thus the variability at other time steps rather than the peak is neglected/lost. Based on the examples that the authors show, it can be seen that historic storm surge hydrographs and tidal cycles show a large variability at times before and after the peaks. The variability of the water levels at times before and after the peak can largely affect the resulting flooding (see e.g. Quinn et al., 2014; Santamaria-Aguilar et al., 2017). The authors briefly discuss this issue regarding TCs, ETCs and mixed semidiurnal tidal regimes in the discussion, but I miss a more detailed discussion or quantification of the variability of the hydrographs related to the mean. For example, how much can vary a hydrograph in a region with a mixed semidiurnal regime (large tidal variability) and affected by TCs and ETCs? Are there regions where the variability of the hydrograph can largely exceed the one of the peak?

As correctly stated by the reviewer, the focus of the methodology is indeed on generating average hydrographs. We focus on average hydrographs because the goal of our study is to enable large-scale flood modelling using dynamic models. The most common approach for flood risk assessments is to compute flood maps corresponding to different return periods (1 in 10-year, 1 in 100-year, etc.). Having only one hydrograph per return period makes the storm tide hydrographs dataset easy to implement in large scale flood hazard assessments, which are computationally expensive. Another reason for not focusing on the variability of the surge levels is that we only have 38 years of data (resulting in a set of ±38 events). Studying the variability that is caused by specific storm characteristics such as storm length, intensity, and wind direction, would require much longer time. To get some idea of the

variability of the hydrographs, we computed the 25th and 75th percentiles surge hydrographs (Fig. 4a & 4b), and we show all tidal segments that are used to calculate the average and spring tide signal (Fig. 5a & 5b). The variability of the water level at times before and after the peak can indeed largely affect the resulting flooding. Therefore, we have added the two references mentioned by the reviewer in the introduction that now reads as follows:

> *L62: "Hydrograph characteristics that determine the flood severity are, among others, the maximum storm tide level, base duration, and overall shape. For example, when the water level is elevated for a longer period of time, particularly close to the time of high water when defence exceedance is most likely, the water will propagate further inland (Santamaria-Aguilar et al., 2017; Quinn et al., 2014)."*

The authors state that one of the main objectives of this study is to enable dynamic flood modeling at global or large scales by providing storm tide hydrographs. I differ on this as I think that the main limitation for dynamic flood modeling at global scales is still the computational effort required, even by simplified hydrodynamic models like LISFLOOD and SFINCS. However, this methodology contributes to dynamic flood modeling at any other scale as hydrographs are commonly neglected in EVA. Therefore, I think this work would benefit by changing the focus to the limitations of EVA models that neglect the hydrograph information, regardless of the scale of the flood assessment.

We believe we are on the same line as the reviewer when it comes to the motivation behind this study: we would like to move away from the bathtub approach that is based on an EVA analysis of static water levels and subsequently computing flood maps for various return water levels (e.g. RP100). In this study our aim is to develop a relatively simple method to generate storm tide hydrographs such that it can be applied at large scales. It is true that currently simplified hydrodynamic models are not yet applied at the global scale due to the high computational costs, but there are some examples at the continental scale (Vousdoukas et al., 2016; Paprotny et al., 2018)*. Moreover, globally applicable approaches are being developed (e.g. Eilander et al., 2022), which allow for the rapid set-up of inundation models anywhere on the globe, forced by global datasets (or more local data where available) - also for this approach, data on hydrographs are required to move past the bathtub approach. When modelling inundation at the regional to local scale, it would make sense to further advance our methodology such that it is able to incorporate more local characteristics of extreme sea levels, such as was done by Macpherson et al. (2019) for the German Baltic Sea coast. Therefore, we stick to the research gap as described in the original manuscript. However, we adjusted the scale from 'global-scale' to 'continental- to global-scale' throughout the revised manuscript.*

Reviewer #2 (Minor comments):

L11 - Coastal flooding can also arise from other drivers such as waves, precipitation, and river discharge in estuarine regions or a combination of all these drivers.

We changed the first sentence of the introduction as follows:

> *L11: "Coastal flooding is driven by the combination of (high) tide and storm surge, the latter being caused by strong winds and low pressure in tropical and extratropical cyclones."*

L12- Missing "cyclones" at the end of the sentence.

Amended

L16 & L24- At global scale, the main constrain regarding the flood model used (static vs dynamic) is the computational effort. Even simplified hydrodynamic flood models such as LISFLOOD or SFINCS still require huge computational resources compared to bathtub models. In my opinion, the main focus should be placed on losing the hydrograph information when doing the EVA, as most EVA models focus only on storm surges peaks. Therefore, information about the hydrograph or temporal evolution of the event is lost although it is required as boundary conditions for dynamic flood models, even at local scales.

See previous author comment under 'main comments'

L27- Exposure increases because of the increase in population, not the opposite way.

We agree with the reviewer. This line now reads as follows:

> *L27: "In addition, the number of people living in coastal areas below 10 m elevation worldwide is projected to increase from over 600 million people today to more than 1 billion people by 2050 under all Shared Socioeconomic Pathways scenarios (Merkens et al., 2016), which means that the exposure will increase."*

L30- Global coastal flood risk assessments can help identifying…

Amended

L32-Coastal flooding is generally driven by "storm surges" generated from strong winds…

Amended

L47- Authors can cite here Ramirez et al., 2016; Vousdoukas et al., 2016

These studies are now cited in the revised manuscript.

Some lines of the introduction do not connect very well (are not smooth) e.g. L30 to L35. I'll suggest revising it trying to better connect the points between sentences.

To improve the flow of the text we made some changes to the first paragraph of the introduction:

> *L26: "Over the course of the 21st century, coastal populations will be increasingly at risk of flooding due to sea level rise (SLR) (Oppenheimer et al., 2019). In addition, the number of people living in coastal areas below 10 m elevation worldwide is projected to increase from over 600 million people today to more than 1 billion people by 2050 under all Shared Socioeconomic Pathways scenarios (Merkens et al., 2016). Global coastal flood risk*

assessments can help identifying areas that are potentially exposed to flooding under both current and future climate conditions (Ward et al., 2015). To setup these flood risk assessments, it is important to understand the dynamics of storm surges generated from strong winds from both low pressure in tropical- (TCs) and extratropical cyclones (ETCs) and how these generate coastal flooding (Resio and Westerink, 2008). Flood models can be used to model these coastal inundation dynamics resulting from extreme storm tides, where the storm tide is defined as the combination of storm surge and the tide (Colle et al., 2010)."

L88- I think authors should cite all different methods available in the literature (or most of them) even if they only describe in more detail a subset of them.

In the revised manuscript, we now include a much larger number of relevant studies, including (Sebastian et al., 2014; Chbab, 2015; Environment Agency, 2018; MacPherson et al., 2019; Vousdoukas et al., 2016a; Xu and Huang, 2014; Salisbury and Hagen, 2007).

L100-Why that threshold and duration are selected? Was a sensitivity analysis performed in Chbab (2015)? I think this point has to be mentioned and the sensitivity of the approach to these parameters (i.e. threshold and duration of the event) discuss. In addition, several normalized hydrographs are generated from this approach, which one is used for estimating any desired RP hydrograph? The mean normalized hydrograph? How the variability of these extreme hydrographs can affect the resulting design hydrograph? See e.g. Santamaria-Aguilar et al., (2017) and Quinn et al., (2014)

We thank the reviewer for pointing out that the description of the method from Chbab (2015) was lacking some details. We have included any missing relevant information (i.e. the selected threshold and duration). Unfortunately, in doing so we found a typo in the text. The threshold for a surge event to be selected is not 0.5 m but 1.5 m. The reason for selecting 1.5 m as threshold is that it results in a set of surge events that is large enough to calculate a representative average, but not too large because then too many small surge events would be selected that have a different physical behaviour than higher surge levels and influence the resulting standardized shape. The reason for extracting 48 hours of the surge time series (24 hours before until 24 hours after the surge maximum) is that the duration of a typical storm in the area is also 48 hours. The revised manuscript reads as follows:

> *L102: "To test the sensitivity of the surge hydrograph to the chosen parameters, a sensitivity analysis performed. They conclude that the upper 50% of the normalized surge height (normalized surge height > 0.5) is not affected when either the threshold or time window length is increased or decreased. This is an important finding because it indicates that the surge hydrograph is most robust close to the time of high water when defence exceedance is most likely (Santamaria-Aguilar et al., 2017; Quinn et al., 2014). However, a longer time window (of e.g. 72 or 96 hours) results in a longer base duration. The argument given for using a 48-hour time window is that 48 hours is the typical duration of a storm along the Dutch coastline.*

L121- Although this is true for most of the places, a dependency between skew surges and high tidal levels is observed at some locations, se Santamaria-Aguilar and Vafeidis (2018).

To make readers aware that at some locations there can be a dependency between skew surges and high tidal levels, we include the following sentence:

> *L134: "apart from some locations where a dependency between skew surge and high tidal levels is observed (Santamaria-Aguilar and Vafeidis, 2018)"*

L122- How sensitive is this method (i.e. the 15 events selected) to the series length? Is there any limitation of this approach to a minimum length of data?

Information about the sensitivity of this method to the series length and the number of events selected (15) is not provided in the report of the Environment Agency. We added the following sentence to the manuscript to inform readers about this:

> *L137: "An argument for selecting this number of events is not given."*

L157. A similar method than MacPherson et al., (2019) was developed by Wahl et al., (2012) for the German North Sea coast (macrotidal)

The authors are aware that the study by Macpherson et al. (2019) makes use of the methods developed in Wahl et al. (2012). To make readers aware of this as well the text now reads as follows:

> *L142: "The third method by MacPherson et al. (2019), that further developed the method from Wahl et al. (2011, 2012), starts by identifying storm tide events."*

L161. The method of Vousdoukas et al., (2016) overestimates the WLs of the hydrograph assuming the maximum high tidal level along the entire duration of the event (i.e. neglecting the tidal variations/cycle), and this issue can significantly overestimate the WLs in those places of large tidal variability (e.g. places with macrotidal range). I think this point needs to be mentioned here or when the approach is described.

We agree with the reviewer that this limitation of the method from Vousdoukas et al. (2016) could be clarified in the manuscript. Therefore, we included the following sentence:

> *L154: "The assumption that the maximum high tidal level occurs along the entire duration of the event, thereby neglecting tidal variations, can significantly overestimate the water level in places with large tidal variability, such as north-western Australia."*

L167. I'm missing a bit of a discussion about the variability of hydrographs and why this is important.

To emphasize the importance of the variability of the hydrograph, particularly close to the maximum water level, we added the following sentence to section 2 of the revised manuscript:

> *"It is especially important that the hydrograph represents the water level correctly close to high water when defence exceedance is most likely, and because the water will propagate*

further inland if the water level is elevated for a longer period of time (Santamaria-Aguilar et al., 2017; Quinn et al., 2014)."

L183-Is this comparison/validation based only on the peaks? How well these modeled data represent the hydrographs?

This validation study focused on the peaks, or maximum surge levels. To make this clear to the reader we have put the word 'maximum' in front of 'surge heights'.

L207- i.e. dividing each surge level by the peak

Amended

Average tide signal and spring tide signal. Although it is briefly discuss later, I think authors should mention the issues of this approach in places with different tidal regimes, e.g. mixed semidiurnal regimes in which the variability of high tidal levels is large and the spring cycle cannot really be defined?

In areas with different tidal regimes, such as a mixed semidiurnal regime, and where the variability of high tidal levels large our methodology might not be able to correctly represent the average (spring) tide signal. Indeed it is important to make the reader aware of this. Therefore, this is extensively discussed in the fourth paragraph of the discussion.

L256. Storms

Amended

L262. Differences in the time evolution of storms can also contribute to the differences/variability observed at the different time steps of the hydrographs, not only the effects of the tide-surge interaction. This can be particularly true in those places that are affected by TCs and ECs, as the characteristics of the two types of storms differ, but the authors analyze together the storm surges hydrographs that arise from TCs and ECs. (It is discussed later in section 5, but I think it should be at least mentioned here too).

The results section now also mentions differences in the time evolution of storms as a potential contributor to the variability observed at the different time steps of the hydrographs. The following line is now included in the results section 4.1 of the manuscript:

> *L284: "Differences in the evolution of storms over time can also contribute to the variability observed at the different time steps of the normalized surge levels, particularly in areas that are affected by TCs and ETCs as the characteristics of these two types of storms differ considerably* (Domingues et al., 2019)*."*

L283. I would like to see how much the finding are affected by this.

The large variability between tidal cycles makes it more difficult to extract an average (spring) tide signal that represents the tidal characteristics at this location (Montevideo). In the manuscript we argue that the effect on the combined hydrograph is small because this large variability between tidal cycles is something we observe especially in areas with a mixed semidiurnal tide regime and small tidal range. The high tide value (maximum of the average (spring) tide signal) is however still correctly represented by the hydrograph. Therefore we cannot exactly quantify the effect of this finding on the final hydrograph. However, one can look at the spread between tidal cycles in Figure A1 to get some idea.

L294. What are the differences at other time steps and their potential effects?

Generally, when the difference between the average and spring tide signal maximum (high tide) exceeds 0.5 m or 1.0 m, the difference between the average and spring tide signal minima (low tide) also exceeds 0.5 m or 1.0 m. It is mostly the magnitude that differs between the tidal cycles, due to the M2 phase, and not the overall shape (at what time the tide goes up or down). As a result, differences are small between tidal cycles around a water level of 0.0 m, and this difference increases towards the time of high or low tide. Because we are especially interested in maximum water levels, when the water can potentially flood the land, we report the difference between maximum tide levels in the manuscript and not at other time steps. For the two locations that are discussed throughout the paper, the evolution of the average and spring tide signal can be seen in Figure 5a and 5b.

L330- The previous study of Wahl et al., (2011) showed that the peak of the hydrograph can show a dependency to the intensity of the hydrograph. I think this issue should be mentioned here as e.g. a dependency of the shape of the hydrograph shows a dependency to the threshold in the case of Marco Island.

It is now mentioned in the manuscript in section 4.4. that Wahl et al. (2011) found some dependency of the peak of the hydrograph to the intensity of the underlying surge events. The study from Wahl et al. (2011) is cited.

L334- The dependency of the storm surge peak and tidal levels not only can change the duration of the event due to the time offset, but also the magnitude of the event. Not accounting for this time offset overestimates the peak water levels as the NTR peak magnitude might not be the same at high tidal levels (i.e. for the same atmospheric forcing, the NTR magnitude is larger at lower tidal levels than at higher tidal levels). This issue should be mentioned here.

We agree with the reviewer that in theory we could overestimate the peak water levels as the NTR peak magnitude might not be the same at high tidal levels compared to low tide. However, before combining the tide signal and surge hydrograph, we scale the surge hydrograph up to a certain water level such that the combined water level equals a specific return period (e.g. RP100 based on the COAST-RP dataset). This means that we will never overestimate (or underestimate) the peak value. However, what could be wrong is the contribution of the surge and tide levels to the combined water level. One component might generally be larger (or smaller) in reality, and is one of the limitations of our methods. This is now mentioned at the end of the section 4.3:

"Last, the surge hydrographs are based on the surge residual including tide-surge interaction effects. These interaction effects tend to be positive at low tide and negative at high tide (Horsburgh and Wilson, 2007). As a result, we might overestimate the contribution of the surge to the combined hydrograph at high tide."