This paper presents an interesting strategy to address surge and inundation in the bay of Bengal, based on publicly available model and forcings. The quality of the results is much larger than the previous efforts and a detailed analysis on process dominance is presented. The proposed strategy, although apparently customized to this site, is interesting and worth publishing.

Below I have outline major and minor issues to be solved before approval, some of them requiring new runs. Furthermore, the paper is extremely long (sometimes it looks more a report than a scientific paper) and lacks now and then a clear presentation direction. The paper should thus be reviewer for conciseness and easy of reading.

I consider the paper to be accepted with major reviews.

Abstract

- "Despite recent advancements, the complex morphology and hydrodynamics of this large delta and the associated modelling computational costs impede the storm surge forecasting in this highly vulnerable region." – Nowadays, super computers are available to perform forecasts of much larger and more complex domains, integrated with atmospheric models. The author should review this sentence maybe focusing on the quality of the forecasts and the necessary grid refinements and process-knowledge for high quality results.
- "This article shows the proof of the maturity of our framework for operational implementation, which can particularly improve the quality of localized forecast for effective decision-making" – Is the framework generic or only applicable to the bay of Bengal? The authors should clarify this issue at the abstract.

Line 36

- "global weather and forecasting system" – a word is missing of the "the" word needs to be removed

Lines 44-45

- "Nowadays, operational surge forecasting systems typically run on high-performance computing systems, either on a scheduled basis or triggered on-demand during an event (Khalid and Ferreira, 2020)" – The authors should include other references of such system, either applicable at a specific site or of generic application.

Lines 50-52

- "Storm surge forecasts have shown their potential to better target the evacuation decision, to optimize early-engineering preparations, and to improve the efficiency of the allocation of the resources (Glahn et al., 2009; Lazo and Waldman, 2011)." – again an updated and more comprehensive review is missing, along with the identification of what are the major challenges in developing and keeping in operational mode this type of systems.

Lines 58-60

- "In the past decade, unstructured-grid modelling systems are getting more and more popular due to their efficiency in resolving the topographic features and their reduced computational cost compared to structured-grid equivalents (Ji et al., 2009; Lane et al., 60 2009; Melton et al., 2009)." – all these references are not from the last decade. Part of them do not address operational forecast systems. There are several examples of unstructured grid forecast system in operation, some recent some in operation for over a decade. The authors should review carefully the state-of-the-art and improve the current text.

Line 100

- "Due to this interaction, the highest surge is obtained for a storm making landfall around 2 hours before the high tide." – a detailed explanation (or references explaining it) is needed. Is it associated with the specifics of the geometric/bathymetric characteristics of the bay or generic? Is it tidal amplitude dependent?

Line 165-169

- "Our bathymetric dataset is a blend of two digitized sounding datasets in the nearshore zone – one from navigational charts produced by Bangladesh Navy, and another being a bathymetry of the Hooghly estuary provided by IWAI (Inland Waterways Authority of India)"- are there any common areas between the two sources ? if yes, what was the combination procedure? If not, substitute "blend" by "combination". How old is the data?
- "The river bathymetry is composed of a set of cross-sections obtained from the Bangladesh Water Development Board (BWDB)." – what is the spacing between profiles? How old is the data?

Lines 194-195

- "At each of the upstream 195 river open boundaries of Ganges, Brahmaputra, Hooghly, and Karnaphuli, we implemented a discharge boundary condition" – what is the source of the discharge values?

Fig 4 – the map is unreadable. Place it at a larger scale.

Lines 200 to 209

this paper aims at evaluating a procedure (framework is not adequate in this context) to forecast storm surges and evaluates the procedure using a past event. However when running an operational model, reanalysis are not available. Therefore the quality of the model should be evaluated with a past event (so data is available) but for a run under operational forecast conditions. The analysis of fig 4 should therefore by re-done under these conditions.

Lines 213 and following: why is the comparison limited to the storm path? The data is available in the whole domain.

Lines 274 and following: this text belongs in the introduction. Remove.

Lines 282-284: "We communicated the results to Bangladesh local government authority through personal communications, as well as to the scientific community through social media." – this sentence is off context and has no scientific link with the remaining of the chapter. It should be moved to other parts of the paper (introduction?)

Fig 6 only deals with time. It should be improved with a plot on space definition of forcings. The use of the "blend" should be reviewed.

Page 13 – the proposed methodology seems too linked with the specific physics of the site and of this particular event. Small variations of the coupling should be tested and compared with data.

Lines 315 and following: errors are necessary for the forecast runs. The discussion is very weakly supported without them.

Lines 327 until the end of page 15 belong in the introduction as motivation for this study.

Lines 344: "The best way to avoid the error from the analytical wind field might be not using these formulations and rely on the full-fledged atmospheric forecasts". In spite of the limitations of existing atmospheric forecasts, this sentence and the next ones would be better supported with a simulation just based on the available forecasts. I suggest the author to repeat the simulation without the analytical model and evaluate the differences.

Fig 8 – why are the inundation patches for the hindcast experiment and not for the forecast runs? I suggest that those are included too, with another figure.

Lines 400-401 – refer to table A2. In table A2, correct the 2nd link as it is not a public link.