Summary

Thank you for the opportunity to conduct a review of this manuscript. Here, the authors couple an existing storm surge and erosion model to estimate annual rates of coastal erosion for two study areas in the Arctic (Drew Point, AK and Manmontovy Khayata, Siberia). The authors conclude that they can predict multi-annual cumulative erosion on the same order of magnitude of what has been observed at these sites and that their methodology is an important first-step toward an approach for estimating erosion for pan-Arctic scales.

Recommendation

I commend the authors on their writing styles, as evidenced by the small number typographical errors throughout the manuscript. However, this work hosts a multitude of technical issues, most notably the study's methodology and conclusion based therefrom. My feeling is that it does not warrant publication. For this reason, I have limited my review to two major comments, as opposed to more detailed in-line comments.

Major Comments

The authors highlight that "the most important root causes of Arctic shoreline change can only be gained through careful evaluation of the physical processes involved" and yet make no such effort for their own study. For example, one of the two sites where the authors apply their model is Drew Point, AK. Here, it is well known that permafrost blocks bound by ice wedges topple onto the beach due to an undercutting process that is facilitated by storm surge (i.e., "thermo-abrasion"). This reality is in stark contrast with the incremental style of bluff retreat associated with the model of coastal erosion employed by the authors (Figure 1). I understand that the authors ultimately wish to exercise their modeling framework elsewhere, but what is the scientific value of applying such a model to a place like Drew Point? My feeling is that Drew Point is not an appropriate location to apply or test the erosion model the authors use in this study.

My biggest concern regarding the validity of this study is the lack of an error analysis of the model outputs (i.e., annual rates of erosion). The model predictions are higher and lower than the observations and in a somewhat chaotic fashion (Figure 4a-b). In many cases, the model predictions are several factors (approaching an order of magnitude) off. Given that the calibration of the model includes an input of historical retreat rates, is this level of error acceptable? What explains the seemingly non-systematic trends in model error?

I calculated a negative value for the Nash-Sutcliffe Model Efficiency (EF) statistic using the measured vs. modeled erosion rates reported for Drew Point in this study, which indicates that the mean of the Drew Point observations is a better predictor of annual erosion than the author's model. This back of the envelope calculation with a widely used error analysis metric underscores a potentially major issue regarding the predictive power of the author's model.

The EF is given by:

$$EF = \left[\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2 - \sum_{i=1}^{n}\left(P_i - O_i\right)^2\right] \Big/ \sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2,$$

where P_i are the predicted values, O_i are the observed values, n is the number of samples, and $\overline{O}$ is the mean of the observed data. The EF statistic ranges from 1.0 to $-\infty$, with 1.0 indicating a perfect match between P_i and O_i and EF less than zero indicating that $\overline{O}$ is a better model than P_i for simulating O_i.

Without a formal error analysis or comparison to another erosion model, it difficult to argue that this study has advanced our understanding of Arctic coastal erosion processes or produced meaningful insights for the communities that are vulnerable to this environmental problem.