PREDICTING DUNE EROSION WITH COMBINED PROCESS-BASED AND MULTIVARIATE STATISTICAL MODELS

Thomas Wahl, University of Central Florida, t.wahl@ucf.edu
Nathaniel G. Plant, U.S. Geological Survey, nplant@usgs.gov
Joseph W. Long, U.S. Geological Survey, jwlong@usgs.gov
Victor Malagon Santos, University of Central Florida, vmalagon@knights.ucf.edu

INTRODUCTION

Oceanographic variables such as mean sea level, tides, storm surges, and waves are drivers of erosion, and they act on different time scales ranging from hours (associated with weather) to seasonal and decadal variations and trends (associated with climate). Storm erosion of dunes, which often protect coastal communities and built infrastructure from flooding, poses a major risk. Dune erosion is also a challenge to analyze and predict, due to the complex processes involved and the computational costs that come with it. Here, we introduce a novel framework that combines advanced statistical techniques and a process-based numerical model to efficiently simulate storm-induced dune erosion.

DATA AND METHODS

For our case study site, Dauphin Island in the northern Gulf of Mexico, we first pre-process in-situ water level and wave observations (Figure 1, top) and develop and apply a multivariate copula-based model to simulate a large number of sea-storm events (Figure 1, bottom). The Multivariate Sea Storm Model (MSSM) provides information on the tide, storm surge, wave height, wave period, wave direction, and storm duration. It was explicitly developed to preserve the univariate distributions of the different variables and the joint correlations between them, while generating extreme event combinations that have not been observed yet, but are physically plausible (see Wahl et al. (2016) for more details on the model development).

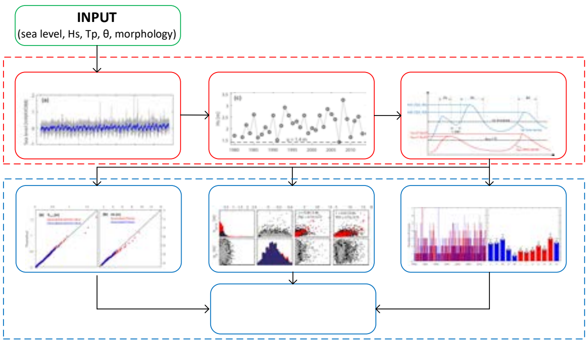

Figure 1 - Steps involved in the data pre-processing (red boxes) and MSSM model development (blue boxes).

Next, we apply a maximum dissimilarity algorithm (MDA) to select 100 sea-storm events (out of 100s of thousands that were simulated with the MSSM) which optimally cover the six-dimensional parameter space. Hourly time series for the selected events are derived with a simple triangular approach for small and moderate extra-tropical events and normalized reference storm surge curves for extreme tropical events. The hourly time series are then used as boundary conditions to force a calibrated and validated XBeach model that covers most of Dauphin Island, to simulate the dune response. The latter is recorded as a set of relevant parameters, including change in the height and cross-shore location of the dune toe and dune crest, dune width, and eroded area. We use the input sea-storm data (from MSSM and MDA sampling) and output dune response data (from XBeach) and test a range of different meta-models for their applicability to replace the computational expensive XBeach model in predicting dune erosion. These include multiple linear regression models (MLRM), artificial neural networks (ANN), multivariate adaptive regression splines (MARS), and random forest trees (RFT).

We use a k-fold cross validation where 5 events are held back for the model setup and then predicted with the statistical models, then the next 5 events are held back, and so on, until all events have been predicted without being used in the model development. Statistical model results are then compared against XBeach output, which is used as reference truth in the absence of observational data.

RESULTS

Preliminary results in Figure 2 show the performance of the different meta-models in predicting changes in the dune crest height, one of the most relevant dune response parameters in terms of flooding risk for the hinterland. Performance metrics used here are correlation, standard deviation, and root mean squared error compared to the XBeach results. These metrics are averaged for all events and across all transects of Dauphin Island (273 in total).

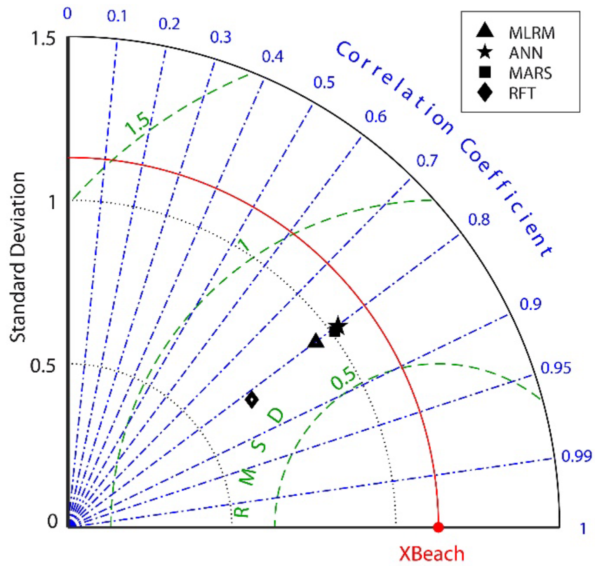

Figure 2 - Taylor diagram outlining statistical model performance relative to XBeach results when evaluating changes in dune crest height.

The results indicate excellent performance of the statistical models for changes in dune crest height, with correlation coefficients consistently larger than 0.8. In this example RFT yield the highest correlation but ANN perform better when taking the other metrics into account. Similar results were achieved for other relevant dune erosion parameters such as changes in dune width or eroded area. Other variables, especially concerned with changes in the dune toe and cross-shore locations are harder to predict and the models tested here resulted in lower correlation coefficients in the order of 0.5 to 0.6.

SUMMARY

The new modelling framework presented here uses advanced statistical models to simulate plausible sea-storm events covering data outside the observational range and machine-learning algorithms to develop meta-models capable of mimicking a state-of-the-art numerical model to simulate dune response during storm events in a computationally highly efficient way. Our results show great potential of the framework be used for both probabilistic long-term coastal change hazard simulations and short-term (ensemble) forecasting (using wave and water level forecast information). The framework itself is transferrable to other coastline stretches where water level and wave observations as well as bathymetric information are available.

REFERENCES

Wahl, T., Plant, N.G., Long, J.W. (2016). Probabilistic assessment of erosion and flooding risk in the northern Gulf of Mexico, Journal of Geophysical Research Oceans, 121, 3029-3043.