Dynamic adjustments to short-term NOAA tide predictions using a multiple linear regression model

UNIVERSITY OF MIAMI
ROSENSTIEL SCHOOL of MARINE & ATMOSPHERIC SCIENCE

Nathan Taminger, Brian McNoldy, & Brian Soden

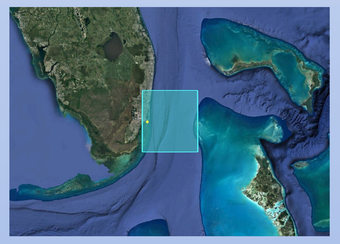
University of Miami | Climate Resilience Academy | 25 April 2022

Background

- NOAA's Virginia Key tide gauge has measured water levels since 1994
- NOAA builds predictions for this tide gauge through a Fourier decomposition of astronomical factors with a reference period from 1983-2001 to capture the 18.6-year lunar nodal cycle
- Predictions currently have a low bias of ~0.5 feet due to sea level rise (SLR)
- Meteorological and oceanographic variables can affect water levels in an area but are not captured in NOAA's predictions.
- Thus, tide-induced flooding events are rarely predicted by NOAA for Virginia Key despite occurring multiple times a year
- Purpose of this study was to see if a model could be built upon NOAA's tide predictions that could forecast maximum daily tide heights out to ten days by accounting for sea level rise and utilizing forecasts for environmental variables off the coast of Virginia Key

Data

- Observed and predicted water level data from Virginia Key was downloaded for every 6 minutes from 01/01/1996 to 12/31/2020
- Maximum daily values were found for observed and predicted tides
- Outliers and the linear trend from sea level rise were removed from the observations and the tide "departure" was found using highest observed water level – highest predicted water level for each day in 25-year period



- SLP, U10, V10, SST, and SWH data was downloaded from the ERA5 model for every 3 hours from 01/01/1996 to 12/31/2020 in a grid around 25.25 to 26.25°N and 80.25 to 79.25°W
- Data was averaged out for each day and outliers were removed
- ERA5 variables besides sea level pressure were normalized

Methods

- Found previous known departure was closely correlated to future departure
- Pressure effect empirically accounted for using 1hPa to 1 cm. relationship
- Other environmental variables were converted to 3-day running average
- Using perfect forecasts, a multilinear regression on future departure was run for each day out to ten days with U10, V10, SST, SWH and previous departure

Running the Regression

- The model was trained on a random 80% of the dataset and tested on the other 20% and the coefficients and p-values for each variable and r-squared value of the model was saved
- A Monte Carlo simulation was made by iterating this process 10,000 times and averaging out the coefficients and p-values
- The ten-day model was created by repeating this process for each day

Results

The following figures show the strength of each of the 5 models (NOAA, +SLR, +SLP, + Persistence, + all variables)

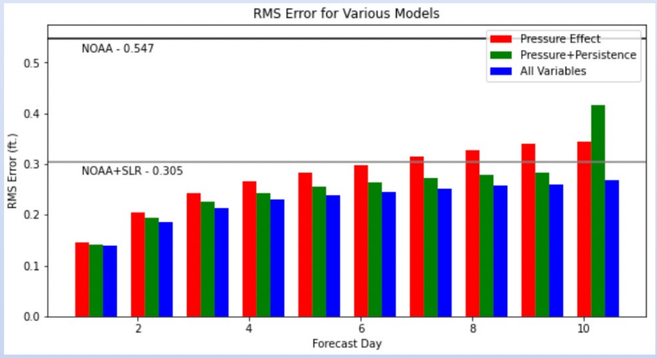


Figure 1: Root mean squared (RMS) analysis for each model

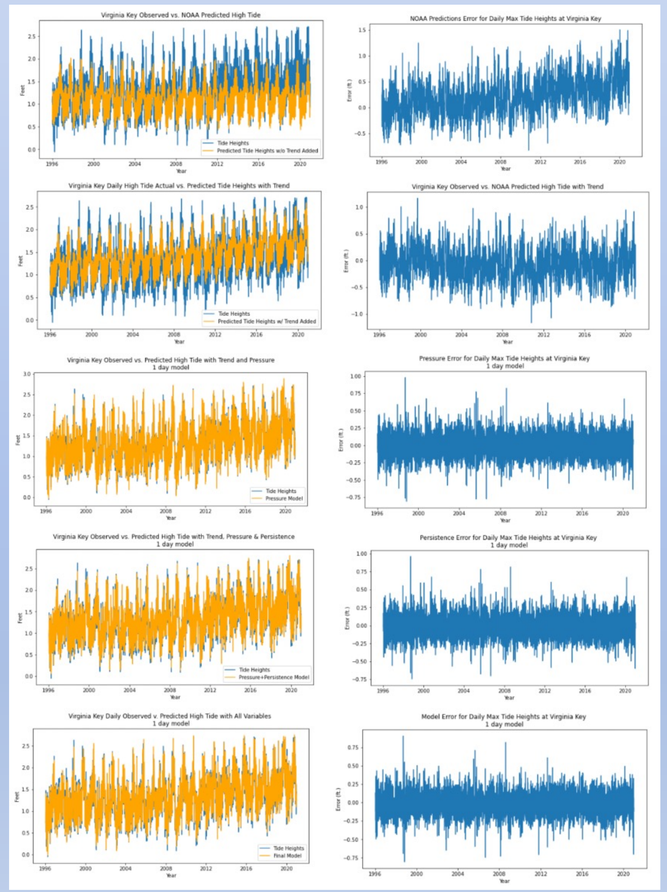


Figure 2: Observed and predicted tides from various models (left) and resulting error (right)

Creating the Forecast

$$Forecasted\ Tide = NOAA\ tide\ prediction + SLR + Pressure\ Effect + Persistence + Environmental\ Effect$$

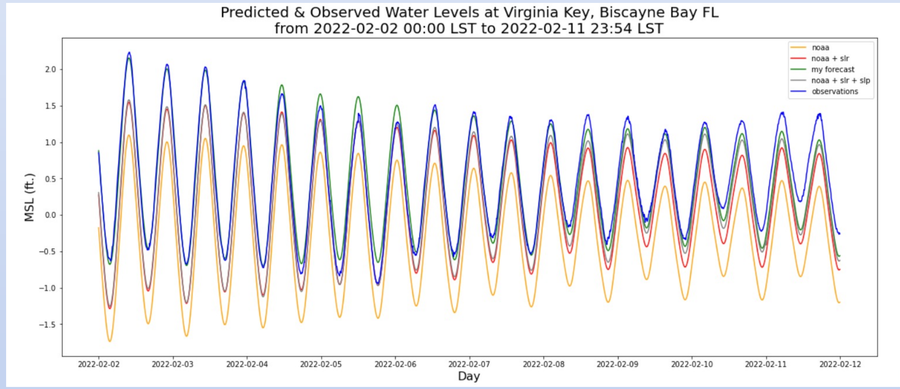


Figure 3: Forecasts from various models plus observed water levels from 02/02/2022 to 02/11/2022

Discussion

- The multilinear regression using all variables resulted in a model that could forecast maximum daily tide heights that was 75% more accurate for one day and over 50% more accurate out to ten days than NOAA's predictions per the RMS analysis
- Model was built using "perfect" forecasts for environmental variables, so the accuracy of the model should be re-evaluated using real forecasts once enough data is collected
- The accuracy of the model could be improved further if forecasts for Florida Current strength or OHC were available
- There exists potential that that the process used to build this model could be recreated at other tide gauges around areas that experience tidal flooding
- NOAA is currently working on updating its predictions using a time period from 2002 – 2020. Once those are released, the model should be updated to account for the change in baseline

Acknowledgements

I would like to give a special thank you to both Dr. Brian Soden and Brian McNoldy for their unending support and guidance with this thesis for the past two years. Without their help, this project never would have been possible. I would also like to thank Dr. Amy Clement for being on my thesis committee and being a resource the past two years. Finally, I would like thank the South Florida Water Management District for allowing me to share this research with them last month and providing support