PennState

Correlation and Regression Analysis on Atlantic Tide Level Data

Beckett Cromer, Benjamin Seiyon Lee, Dr. Murali Haran
Department of Statistics, The Pennsylvania State University

SCRiM

1. Motivation

Extreme storm surges poses significant threats to life and property. Research into these surges is crucial for building and preparing critical infrastructure near the coast. A previous study (Grinsted et al., 2013) examined sea level data from NOAA and concluded that extreme surges were growing in intensity and frequency.

2. Summary

Goals:

- Re-examine the methodology of Grinsted et al. (2012) which combined observations from six different tide gauge stations.
- Develop an alternative pre-processing method and determine if there is a link between more frequent extreme storm surges and rising global temperatures.

Challenges:

- Large data sets, over 800,000 data points per station
- Accounting for overall, seasonal, and daily influences

Our Approach:

- Use harmonic regression and an ARMA model to remove the overall, seasonal, and daily trends.
- Fit a generalized extreme value distribution (GEV) to the pre-processed data and examine how the distribution changes with temperature.

Findings:

- Tide gauge stations exhibit little correlation with one another
- There is a link between rising temperatures and more frequent extreme storm surges, but it is much weaker than reported by Grinsted et al. (2013)

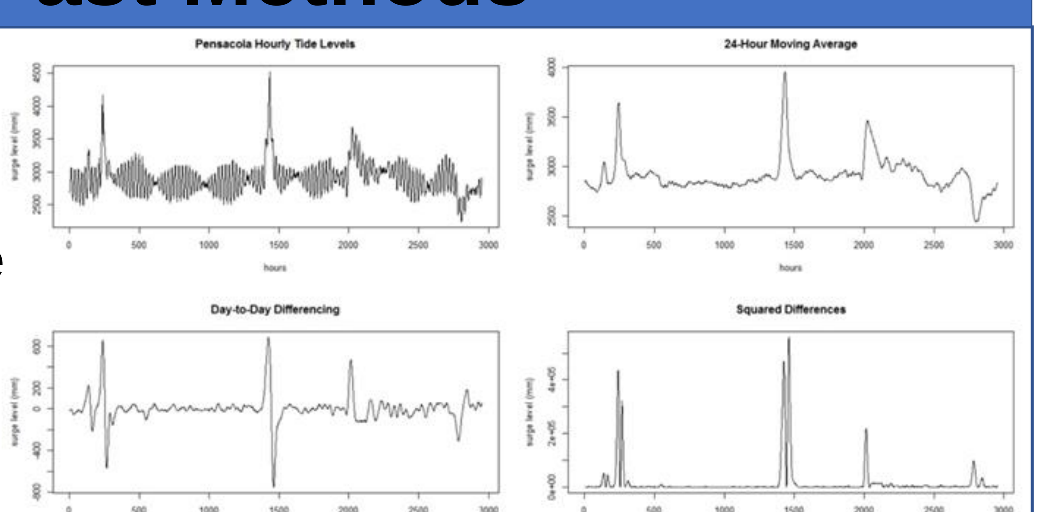
3. Data

- Tide gauge data retrieved from National Oceanic and Atmospheric Administration (NOAA)
- Focus Tide Gauge Stations:
  - Key West (FL), Charleston (SC), Atlantic City (NJ), Pensacola (FL), Galveston (TX)
  - Sea level data from 1923 to 2015

4. Past Methods

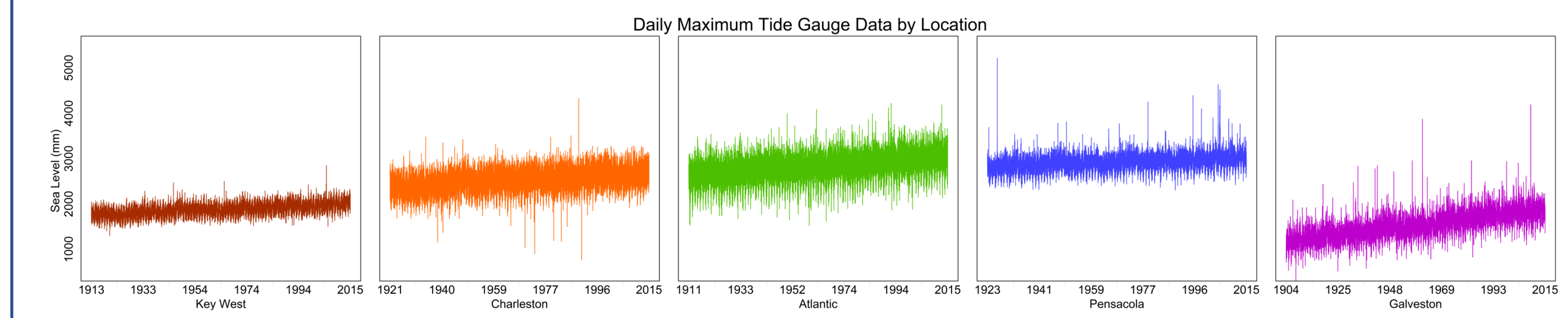
Grinsted et al. (2013):

- Daily max of squared day-to-day differences
- Removed annual cycle by division
- Combined records from 6 stations into one surge index

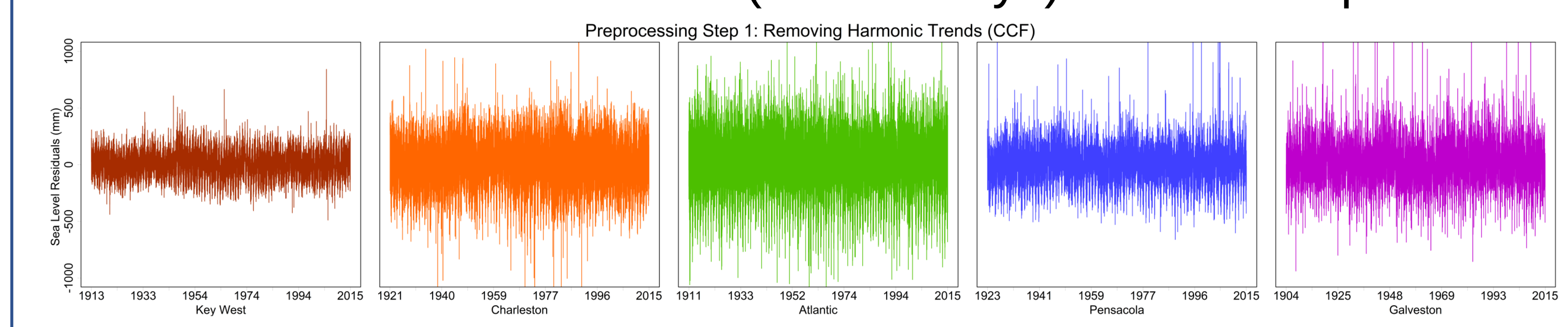


5. Pre-processing Methods

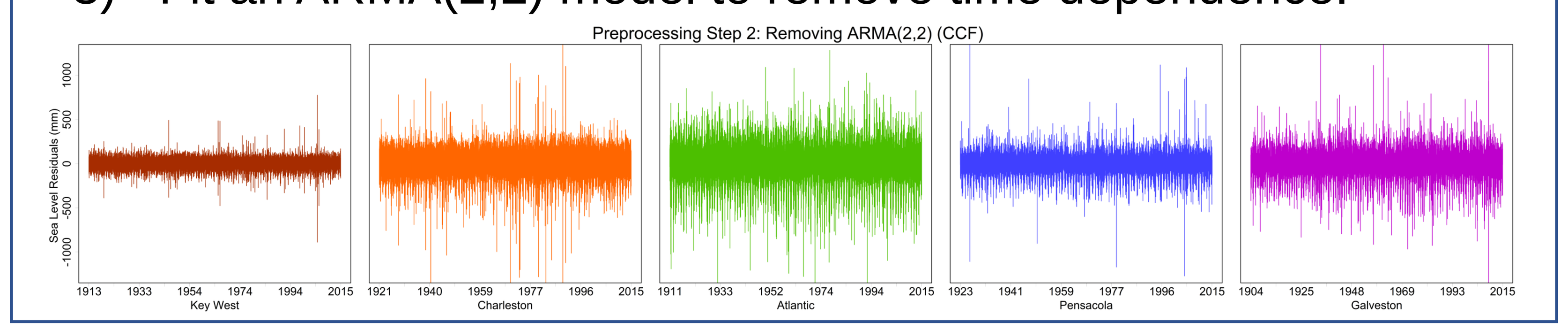
Detrending/Removing Seasonality:



1) Fast Fourier Transformation (FFT) to decompose time series into fundamental frequencies.
2) A harmonic regression with 7 pairs of sines and cosines was used to remove seasonality. Periods used were:
   - A yearly period (365.25 days)
   - Synodic lunar month (29.55.. days) and its half period
   - Sidereal lunar month (27.32.. days) and its half period
   - Anomalistic lunar month (27.55.. days) and its half period



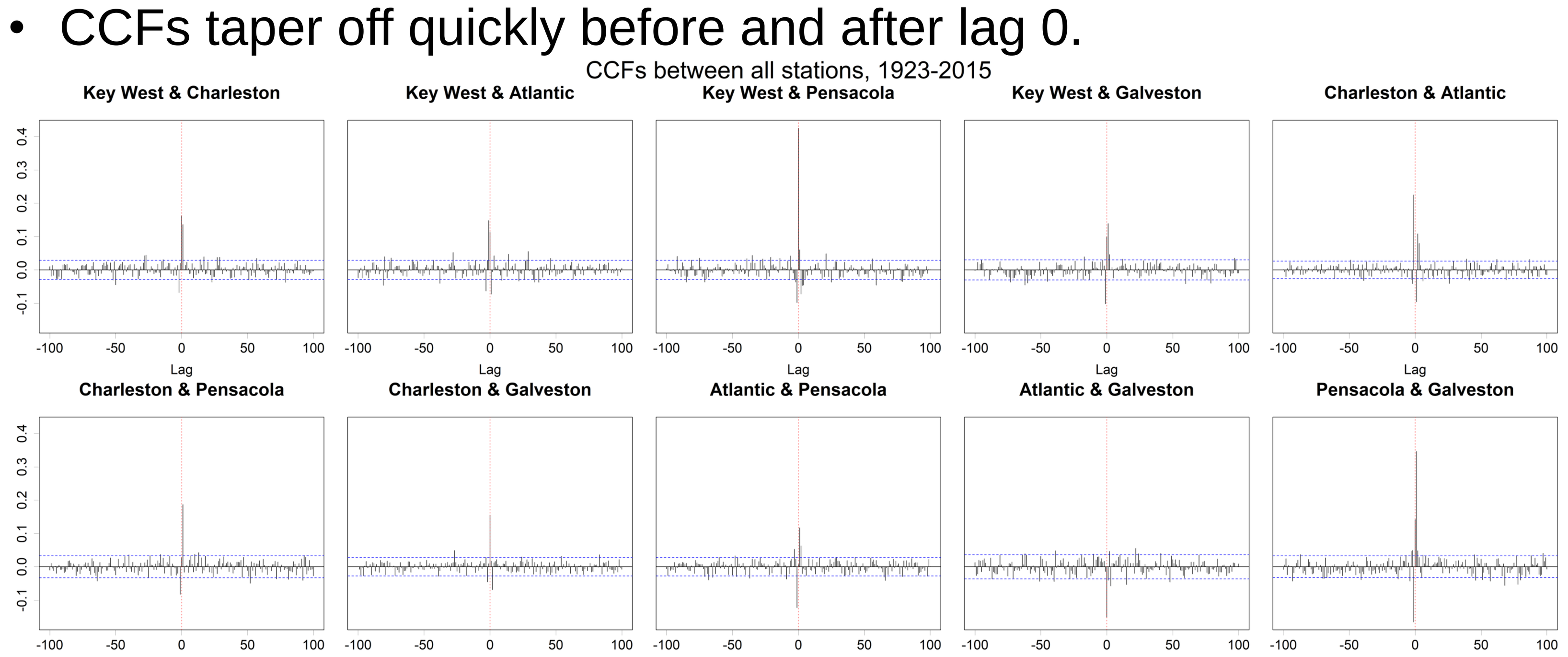
3) Fit an ARMA(2,2) model to remove time dependence.



6. Results

Cross-Correlation Function:

- CCF is a measure of similarity between two series, as a function of lags of one relative to another.
- Strong Correlation between nearby tide gauge stations (~500 mi) at lag 0, maximum of Key West and Pensacola (FL) with 0.42.
- Moderate negative correlation at lag 0 with distant tide gauge locations (~1350 mi), maximum of -0.16 with Galveston (TX) and Atlantic City (NJ).
- CCFs taper off quickly before and after lag 0.



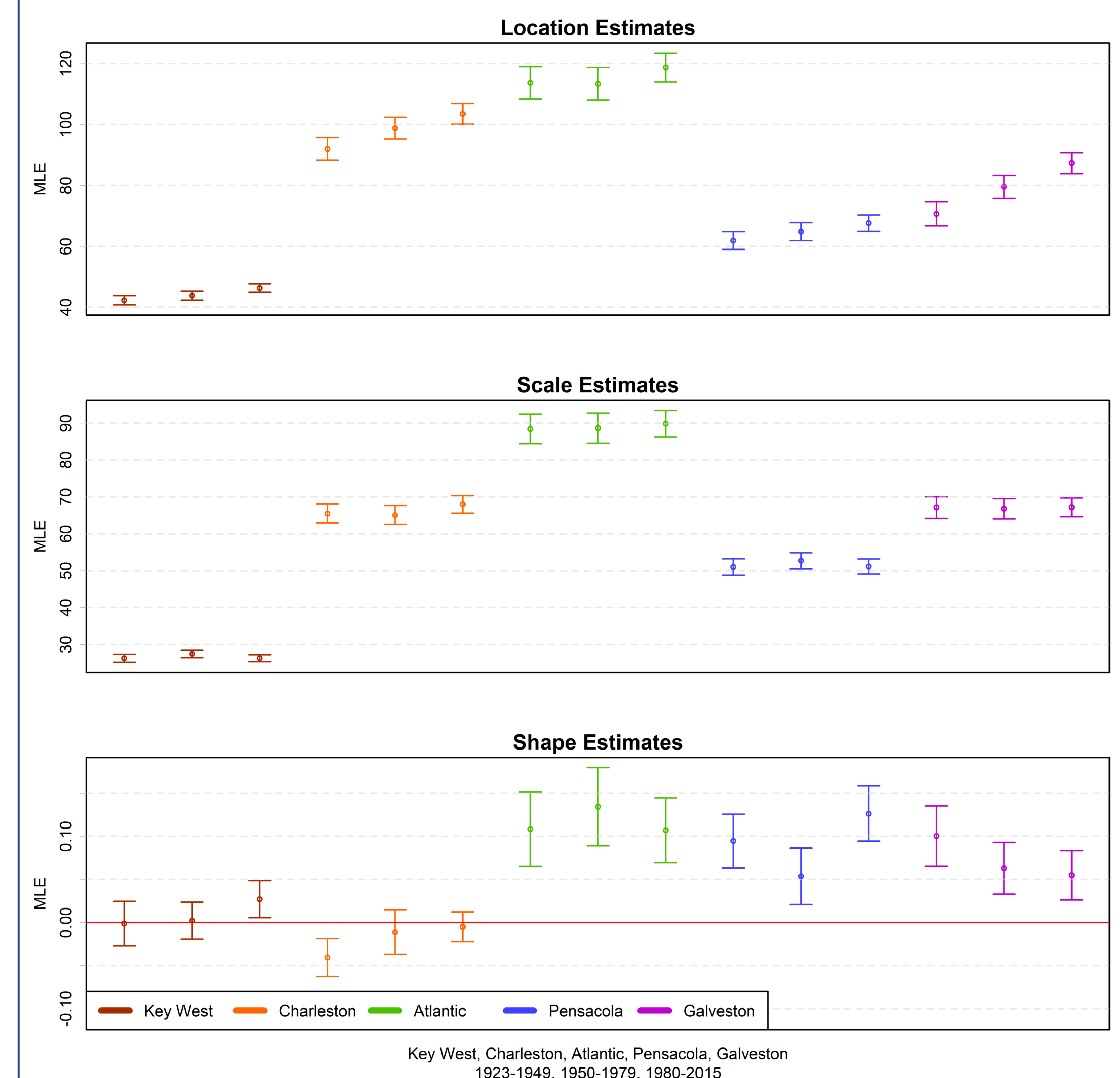
7. Results

Generalized Extreme Value Distribution:

- Location parameters changes with respect to rising temperatures.
- Scale and Shape parameters did not change with respect to temperature.

PDF:

$$f(x;\mu,\sigma,\xi) = \frac{1}{\sigma}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{(-1/\xi)-1} \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$



8. Conclusions and Future Work

Spatial Dependence:

- Tide gauge stations closer in proximity showed a strong correlation to one another while distant stations did not.
- Stations are not similar, so combining them (Grinsted et al., 2013) is not recommended.

Non-Stationary GEV:

- GEV analysis shows a strong increase in the location parameter over time, with no overlap in the 95% CI in four of five tide gauge locations

Future Work:

- Implement estimates for missing stretches of data

Contact:
Beckett Cromer, The Pennsylvania State University
Department of Statistics
Email: bdc5221@psu.edu

Acknowledgements:

- Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.
- This work was supported by the National Science Foundation through the Network for Sustainable Climate Risk Management (SCRiM) under NSF cooperative agreement GEO-1240507

References:

- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/
- M. Chapront-Touzé, J. Chapront (1991): Lunar tables and programs from 4000 B. C. to A. D. 8000. Willmann-Bell, Richmond VA
- Grinsted A, Moore JC, Jevrejeva S. Homogeneous record of Atlantic hurricane surge threat since 1923. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(48):19601-19605. doi:10.1073/pnas.1209542109.