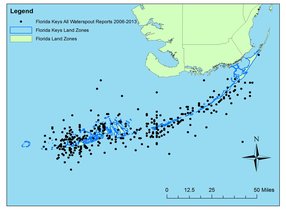
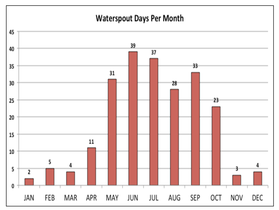
Statistical Prediction of Waterspout Probability for the Florida Keys

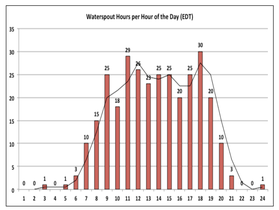
Andrew Devanas*, Lydia Stefanova#, Sean Daida*, Kennard Kasper*, Matt Bloemer@
*National Weather Service WFO Key West
#FSU-COAPS @CWSU Atlanta

Background

- An estimated 300-400 waterspouts occur each year in the waters surrounding the Florida Keys, of which approximately 40 are spotted and reported (no remote sensing).
- During the wet season (summer months), the atmosphere is generally pseudo-barotropic.
- Waterspouts ***are reported*** on approximately 19% of wet season days.
- With little change in the atmospheric profile (outside of tropical cyclones), how are days that are favorable for waterspouts differentiated from days which are not?

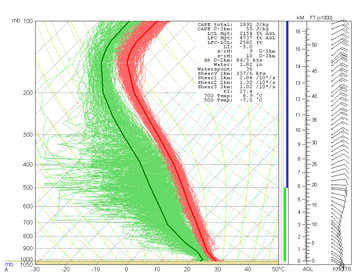






Data

- Nine years of 12Z (8am EDT) rawinsonde soundings at Key West (2006-2014) for the wet season months June – September (1080 days). 12Z was chosen for its predictive potential.
- The Key West sounding was considered representative of the Florida Keys due to the tropical pseudo-barotropic environment.
- NWS Local Storm Reports for the Florida Keys archipelago were used to identify days which waterspouts were reported.
- Waterspouts associated with tropical cyclones were removed.
- Waterspout(s) were reported on 208 of the 1080 days examined.
- 144 variables were pulled from each sounding.
- Data were separated into two sets - days with waterspout reports, and days with no reports.



Question

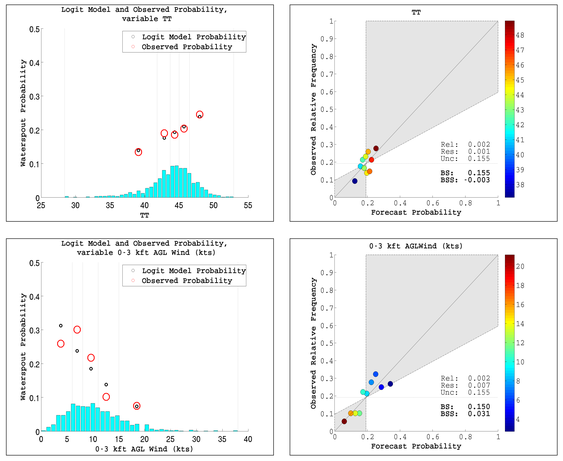
- Is it possible to model the probability of a waterspout report on a given day during the wet season?

Reasoning: a) the probability of waterspout report is proportional to the probability of waterspout existence, i.e. p(report)~p(encounter)*p(existence); b) the probability of waterspout existence is dependent on the environment.

Methodology

- Logistic regression: model probability of waterspout report as a function of predictor variables.
- Preliminary selection of predictor variables: require statistical significance in the difference of means between report vs. no report days; retain wind directions (surface and 100mb) as candidate predictors despite not necessarily satisfying this requirement.
- Examine single-predictor logistic regression for selected variables (examples Fig 1).
- Final selection of predictor variables: Likelihood Ratio testing of multiple logistic regression models.

Fig 1



Results

- Final model: 6 predictor variables (Fig 2).
- Model performance evaluated with 10-fold cross-validation; Results compared to Charleston Waterspout Index (CWI).
- Model probability distribution on Report vs. No Report days (Fig 3).
- Reliability diagrams and Brier Score (Fig 4).
- ROC curves (Fig 5).

Fig 2

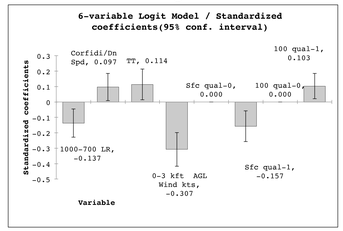


Fig 3

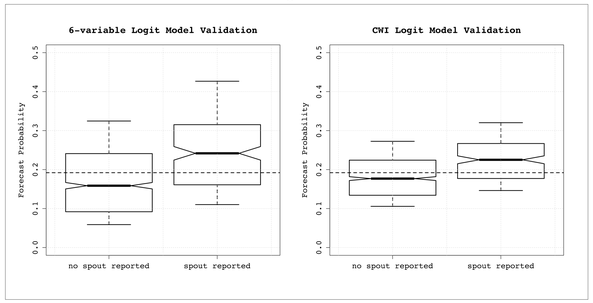


Fig 4

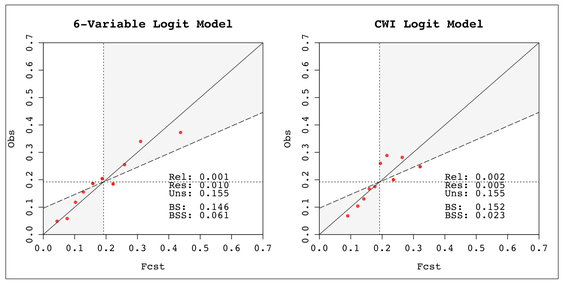
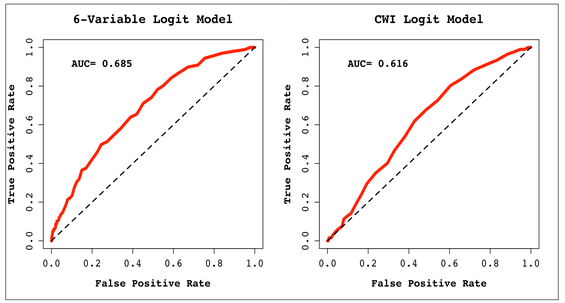


Fig 5



Summary

- Cross validation results suggest that the 6-variable logistic regression model has predictive potential to model the probability of waterspout report on a given day.
- The performance of the model is an improvement over the existing Charleston Waterspout Index for the Florida Keys. The results make no statement about the performance of the Charleston Waterspout Index for the Charleston coastal environment.

Further Work

- Validate model predictions with an independent dataset (soundings and reports for 2014 and 2015).
- Examine the predictive value of 00Z soundings with similar methodology.
- Test the implicit assumption that the model is robust to subsetting (e.g. by calendar month and/or sub-seasonal flow regimes).

Contact:
Andrew Devanas - andrew.devanas@noaa.gov
Lydia Stefanova - lstefanova@fsu.edu