152 STATISTICAL PREDICTION OF WATERSPOUT PROBABILITY FOR THE FLORIDA KEYS

Andrew Devanas[1], Lydia Stefanova[2], Kennard Kasper[1], Sean Daida[1]
[1]NOAA/National Weather Service, Key West, Florida, [2]COAPS/Florida State University, Tallahassee, FL,

1. INTRODUCTION

The frequency of non-supercell waterspouts in the Florida Keys is higher than any other location in the United States and is arguably the highest in the world. Hundreds of waterspouts occur each year in the waters surrounding the Florida Keys. Because waterspouts cannot be detected remotely, only a fraction of these are observed and recorded, underestimating the true number possibly by an order of magnitude (Golden, 1977). Waterspout formation is noticeably more frequent during the summer months, May to October, with a maximum in June. However, waterspouts have been reported in every month of the year, although less frequent from November to April. During the locally defined wet season (June through September), waterspouts are reported on approximately 19% of all days.

The atmospheric environment over the Florida Keys is quasi-barotropic during the wet season, with small day to day variations in the temperature and moisture profile. Because of this, it is difficult to differentiate between days that are more favorable for waterspout development from days that are less favorable simply by examining soundings, sounding climatologies, or individual parameters derived from the soundings. Several studies have been done using sounding parameters to describe conditions favorable for waterspout development (Sioutas and Keul 2007, Sprat and Choy 1994, Brown and Rothfuss 1998), but these studies have tended to be more descriptive than predictive.

The National Weather Service (NWS) forecast office in Charleston, SC developed a waterspout index (CWI) that is available in the commercial RAOB software developed by John Shewchuk, Environmental Research Services, LLC. The CWI assigns risk points based on sounding-derived variables (*e.g.* wind speed, precipitable water, stability indices, *etc.*) meeting certain thresholds. The total number of risk points forms the CWI. The value of the CWI is used to qualitatively classify the risk for waterspout development as high, moderate, low, or none.

In this work, we seek a quantitative representation of the probability of waterspout development as a function of sounding-derived parameters for the Florida Keys wet season. Details of the sounding-derived variables and waterspout observations used here are discussed in the Data section. The statistical model and the conversion of the numerical CWI to quantitative probabilities (needed for comparison) are described in the Methodology section. The performance of our model and a comparison with that of the CWI is

* *Corresponding author:* andrew.devanas@noaa.gov

presented in the Results section, and discussed further in the Summary and Future Considerations sections.

2. DATA

Data were compiled from nine years of 12Z (8am EDT) soundings at Key West (2006-2014) for the wet season months (June through September). Years prior to 2006 were not considered for this study due to the rawinsonde location change in 2005, and inconsistent reporting and waterspout identification practices prior to 2006. The 12Z soundings were selected because of their predictive potential for the daylight hours, which is when most waterspouts occur. The Key West sounding was considered representative of the Florida Keys environment due to the typically quasi-barotropic atmosphere. When tropical cyclones are proximal the atmosphere is not barotropic. Such days were eliminated from the dataset; thus, associated mini-supercell waterspouts were excluded from consideration. In all, that left 1080 days with 144 variables extracted from each sounding.

The NWS Local Storm Reports from the field office in Key West for the Florida Keys were used to identify days on which waterspouts were reported. Days with one or more waterspout reports anywhere along the Florida Keys were classified as waterspout report days. Days with no reports were classified as no waterspout report days. Waterspouts were reported on 208 of the 1080 days examined. It is likely that waterspouts have in fact occurred on no report days but were not observed.

3. METHODOLOGY

Two reasonable assumptions were made: that the probability of waterspout reports is proportional to the probability of waterspout existence; and the probability of waterspout existence is dependent on the environment.

A logistic regression model was employed to model the probability of waterspout report as a function of environmental variables derived from morning soundings. In order to reduce the risk of overfitting, it was necessary to judiciously reduce the number of predictor variables. As a first step towards this reduction, the statistical significance of the difference of means for each variable between waterspout report days versus no report days was calculated. This reduced the number of candidate variables to a more manageable number. Variables associated with wind direction were exempted from this test and were retained for consideration on the basis of examination of individual single-predictor logistic regression. An illustration of an individual single-predictor logistic regression model performance for an individual predictor variable is shown in Figure 1.

The final selection of predictor variables was made using Likelihood Ratio testing of multiple logistic regression models. The final model selected by these criteria contained a total of 6 predictor variables: 1000-700mb lapse rate (LR), Corfidi downshear (speed), Total Totals Index, 0-3 kft AGL average wind speed, surface wind direction, and 100 mb wind direction. Standardized coefficients for the 6-variable logistic regression model (LRM-6) are shown in Figure 2. The surface and 100 mb wind directions were treated as qualitative variables (winds in/out of preferred quadrant). The selected predictors were checked to ensure they were not correlated so that the model was not relying on redundant information.

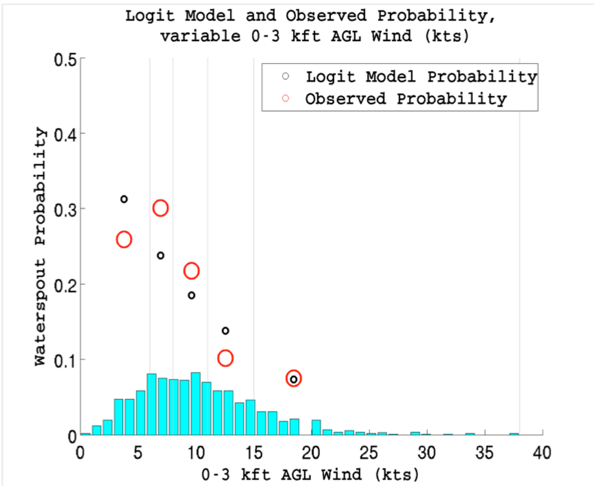

Fig. 1 Example of the results from logistic regression of waterspout probability as a function of a single predictor (low-level average wind speed). predictor distribution (cyan); observed probability (red) within the range of predictor values (gray vertical lines); modeled probability (black) for the same range. The ranges have been selected to have equal number of observations within each bin.

The performance of LRM-6 was evaluated with a 10-fold cross-validation. For comparison, a similar 10-fold cross-validation was performed for the logistic regression built using the CWI and results were compared.

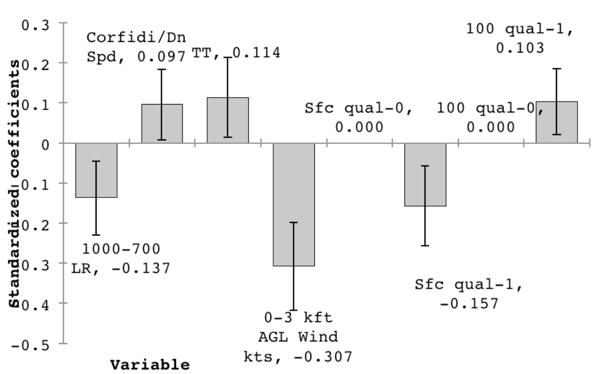

Fig. 2 Standardized coefficients with 95% confidence intervals for LRM-6 using the full dataset.

4. RESULTS

The climatological probability of waterspout reports for the wet season, based on the 2006-2014 data is roughly 19.2%. On no-report days, the cross-validated LRM-6 predicted a mean probability of 17.7% (median 15.8%), while on report days the mean predicted probability was 25.4% (median 23.9%) (Figure 3, top). The difference in means across the two groups is statistically significant (p-value = 5.3e-14).

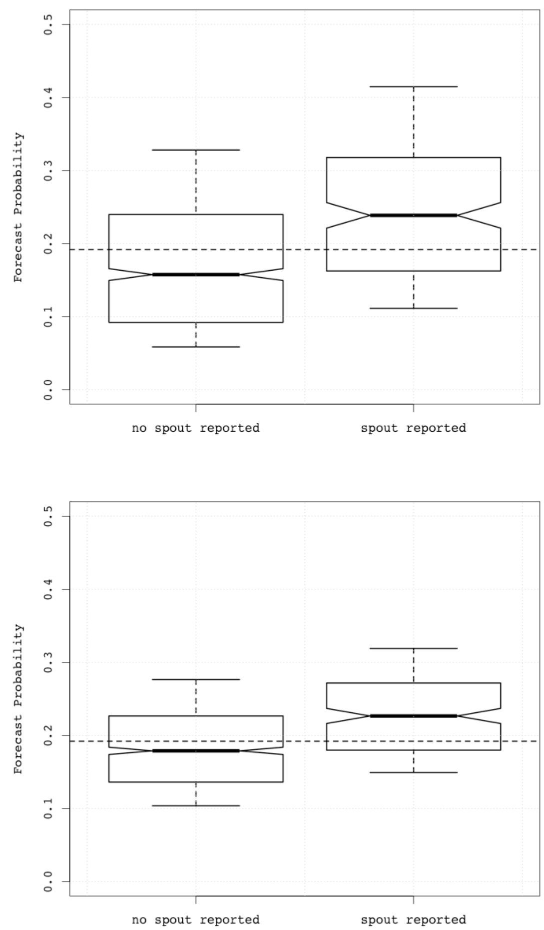

Fig. 3 Boxplot of the cross-validated modeled waterspout probability on report versus no-report days for LRM-6 (top) and the CWI model (bottom). Whiskers show the 10th and 90th percentiles.

The CWI model mean probability on no-report days was 18.4% (median 17.9%), and on report days – 22.8% (22.7%) (Figure 3, bottom), and the difference in means was also statistically significant (p-value = 3.6e-16). The comparison in Figure 3 shows that there is a wider range of probabilities predicted by LRM-6 relative to the CWI model. For a closer look of the two models' performances, reliability diagrams were constructed (Figure 4). Here, the predicted probabilities were separated into 8 evenly spaced bins between 0 and 1 (horizontal axis), and the forecast probability in each bin was assigned its mid-point value. The observed probability (vertical axis) for each bin was calculated as the mean of the corresponding validating observations. The observed probability (vertical axis) for each bin was calculated as the mean of the corresponding validating observations. The Brier Score decomposition (see Wilks 2006) into reliability (Rel), resolution (Res), and uncertainty (Uns) terms, the Brier Score (BS) and the Brier Skill Score relative to climatology (BSS) are indicated on each panel. Note that a better forecast will have a smaller BS, with a smaller reliability term and a larger resolution term. The uncertainty term is determined purely by the observations' distribution, and is not influenced by the forecast. The BSS, on the other hand, is larger for better forecasts. By all these measures, as well as by qualitative visual evaluation, LRM-6 (Figure 4, top) is an improvement over the CWI model (Figure 4, bottom). A comparison between the frequency distribution of forecasts (illustrated by the size of the red dots and the inset histograms) clearly shows that the CWI model is more conservative, with most forecasts near the climatological probability of waterspout occurrence. In contrast, the LRM-6 more frequently, and reliably, predicts probabilities above or below the climatological value.

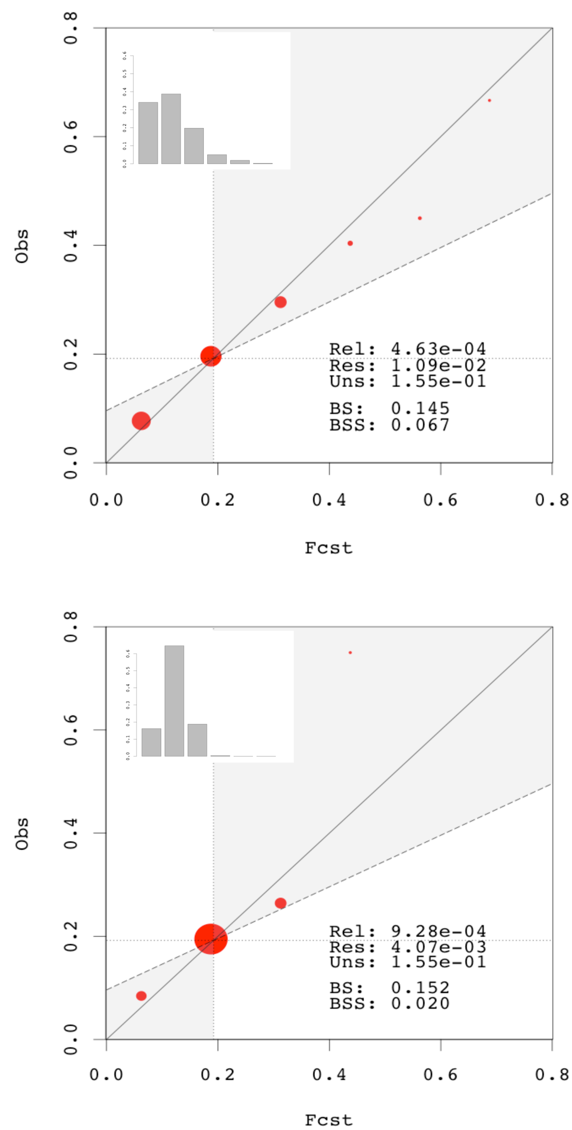

Fig. 4 Reliability diagram for the cross-validated waterspout probability for LRM-6 (top) and for the CWI model (bottom). Point size reflects forecast frequency (also shown as histogram in inset).

Examination of the ROC curves (Figure 5) provides a complementary comparison of the performance of LRM-6 versus the CWI model, in terms of their ability to discriminate between the two outcomes (report versus no report days). Given correctly calibrated forecasts, as indicated by their reliability (Figure 4), the fact that the LRM-6 ROC is entirely above and to the left of the CWI model ROC indicates a greater potential utility of the former (Wilks 2006).

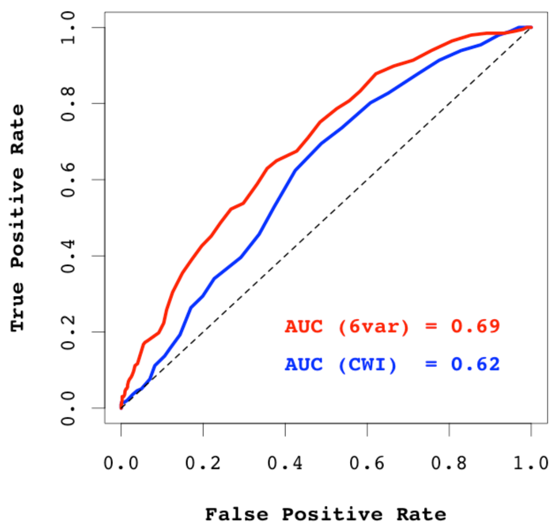

Fig. 5 ROC curves for the cross-validated waterspout probability for LRM-6 (red) and for the CWI model (blue). Area under the curve (AUC) indicated on plot.

5. SUMMARY

The evaluations presented above demonstrate that the multivariable logistic regression, LRM-6, built using a small set of predictor variables derived from morning sounding data at Key West has potential utility in operational forecasting. The LRM-6 produces a superior quantitative probability forecast when compared with an existing operational index, CWI, using the same methodology. The proposed model is an improvement in terms of both its reliability (*i.e.*, statistical accuracy) and its resolution (*i.e.,* ability to distinguish between atmospheric environments with above- or below-climatological probability of waterspout occurrence) over the existing CWI model.

6. FUTURE CONSIDERATIONS

The sounding data will be examined to determine whether temporal and spatial proximity to the sounding location is relevant or significant. This will test the quasi-barotropic assumption implied in the study. The data will also be subsetted to investigate whether intraseasonal variability affects the model performance. Additionally, it will be tested whether including predictors based on prior-day occurrence, and on month of the year would improve the model performance.

The skill of the proposed model for operational forecasting will be further tested by building the logistic regression model with the 2006-2014 wet season data, and applying the resulting model to an independent, unused set constructed with data from the 2015-2016 wet seasons.

7. REFERENCES

Brown D, Rothfuss J (1998) An Approach to Waterspout Forecasting for South Florida and the Keys. *Internal report, National Oceanic and Atmospheric Administration*, Miami, Florida, USA, 1998

Golden J (1977) An assessment of waterspout frequencies along the U. S. East and Gulf Coasts. *J. Applied Met.* **16,** 231-236

Sioutas MV, Keul AG (2007). Waterspouts of the Adriatic, Ionian and Aegean Sea and their meteorological environment. *Atmospheric research* **83**(2), 542-557

Sprat SM and Choy BK (1994) Employing the WSR-88D for waterspout forecasting. Postprints, *NEXRAD Users Conference*, NWS, Norman, OK, 1994

Wilks DS (2006) Statistical methods in atmospheric sciences. *Elsevier*, 627pp.