P 4.8 DEVELOPMENT OF A SOUNDING-BASED PREDICTION MODEL FOR PHOENIX, ARIZONA

Jaret W. Rogers*
National Weather Service, NOAA, Phoenix, Arizona

1. INTRODUCTION

The summer months across the lower deserts of south-central Arizona are notorious for extreme weather. Excessive heat, severe thunderstorms, and flash flooding are all frequent occurrences. Maximum surface temperatures above 100 degrees Fahrenheit are common from May through September. Emphasis on excessive heat awareness has increased over the past several years, making it imperative for the maximum temperature to be predicted with a high degree of accuracy. Moisture associated with the North American Monsoon increases the potential for severe thunderstorms from June through September. Thunderstorms that form during this time period are often extremely difficult to predict, with subtle changes in atmospheric conditions sometimes leading to dramatic convective outbreaks (Maddox et al. 1995). Forecasters in Arizona are presented with unique challenges that may be unmatched anywhere in the country, and rely on several forms of guidance to assist in the decision making process. There are two primary goals of this study: to provide additional guidance to forecasters in the form of a sounding-based prediction model, and to determine what meteorological variables derived from a single sounding are statistically significant predictors of temperature, dewpoint, and precipitation.

2. METHODOLOGY

Regression analysis has been used in past studies as a method to develop prediction equations. Prediction models have been applied to many different weather phenomena, ranging from storm surge prediction (e.g., Harris and Angelo 1963), to maximum temperatures based off numerical model geopotential thickness forecasts (Massie and Rose 1997). Central Arizona may provide the ideal environment during the summer for developing an observation-based prediction model of temperature and moisture parameters. The synoptic environment is typically under the influence of an upper level high pressure, with weak temperature and moisture advection. Exceptions do occasionally occur however, primarily during gulf surge events when moisture advection becomes more prominent (e.g., Dixon 2005). It is believed that routine observations (e.g., soundings) may provide some predictive skill of temperature, dewpoint, and precipitation for the period immediately following the time of observation. The statistical software package R is used for the development of each regression model in the current study.

The regression model developed in the current study is based off data collected from upper-air soundings at Phoenix, AZ. The Salt River Project (SRP) launches soundings two to four times each day during the summer months (June-September). The regression models developed in this study include mandatory level data from the 1200 GMT sounding at Phoenix, AZ (KPSR), between the years 2001 and 2007. Temperature, moisture, and wind observations, as well as other calculated values (e.g., lapse rates, equivalent potential temperature) are included. Stepwise regression methods, using forward selection, are applied to determine statistically significant variables, with a critical alpha level of .05. The prediction model, named the Sounding-Based Regression Model (hereafter, SBRM), was run operationally in 2008.

Four meteorological parameters are used as predictands in this study, including maximum temperature, minimum dewpoint, probability of precipitation, and quantitative precipitation. The maximum temperature and minimum dewpoint, typically occurring in the afternoon, are specific to

**Corresponding author address*: Jaret W. Rogers, National Weather Service, PAB 500, P.O. Box 52025, Phoenix, AZ 85072-2025; e-mail: jaret.rogers@noaa.gov

the Phoenix Sky Harbor ASOS station (KPHX) the day of the sounding. The probability of precipitation and quantitative precipitation regression models are derived from 93 rain gauges, maintained by the Flood Control District of Maricopa County (FCDMC), scattered throughout the Phoenix metropolitan area. The occurrence of precipitation is input into the model as a binary response variable, requiring at least one station within the network to receive precipitation to classify the day as wet. The time period of precipitation occurrence predicted by the model is from 1200 GMT to 0700 GMT the following day. Forward selection logistic regression analysis was applied to generate a prediction equation for the probability of precipitation. Prediction equations for the mean and maximum precipitation amounts across the network were also developed.

3. MODEL

3.1 Maximum temperature

Several variables are included in the maximum temperature prediction model. After a stepwise forward elimination selection process is applied, all sounding variables deemed statistically significant to improving the accuracy of the model are included. The observed low-level temperatures, layer thickness, and lapse rate are all positively correlated to the maximum temperature (Table 1).

The SBRM performed well in the prediction of the maximum temperature at KPHX during the summer of 2008, as indicated by the performance statistics relative to the 0600 GMT MET and MAV Guidance (Table 2). The mean absolute error (MAE) between the observed and predicted maximum temperature is 1.56º F, which is comparable to other guidance output. Generally, a maximum temperature forecast in the summer in Phoenix is generally considered successful if it is within two degrees of the observed maximum temperature (forecaster communication). The mean algebraic error reveals a negative bias in the SBRM prediction, indicating that the maximum temperature is, on average, under-forecast by approximately 0.9º F. A graphical representation illustrates the frequently observed negative bias of the predicted maximum temperature (Fig. 1). The distribution of the differences between the predicted and observed maximum temperature also indicates a negative temperature bias

Variable	**Correlation**
Surface T	+
700 hPa T	+
500 hPa Height	+
300 hPa Tdd	+
Surface-700 hPa Mean RH	-
850-500 hPa Lapse Rate	+
850-700 hPa Layer Thickness	+

Table 1. *A list of the variables included in the maximum temperature prediction model of the SBRM. The correlation between each individual predictor and the maximum temperature is listed.*

Model	SBRM	MET	MAV
Mean Absolute Error	1.56	1.5	1.43
Mean Algebraic Error	-0.92	0.64	-0.57
Standard Deviation	1.71	1.87	1.85
Max. Difference	5	6	6
Min. Difference	-5	-5	-8
Best Model (%)	30.21	28.13	25

Table 2. *SBRM, MET, and MAV guidance maximum temperature prediction results from 2008, compared with the observed maximum temperature. All units are in degrees Fahrenheit, except where noted.*

(Fig. 2). If a bias-corrected model is run, accuracy of the SBRM may be improved even further. Although the MAE of the SBRM relative to the MET and MAV guidance is slightly higher, the standard deviation and other simple dispersion measures suggest that the SBRM may have less variability in its maximum temperature predictions. In addition, the SBRM outperformed both the MET and MAV guidance 30.2% of the time.

3.2 Minimum dewpoint temperature

Dewpoint temperatures are often difficult to predict with a high degree of accuracy (forecaster communication). However, the dewpoint still remains important to forecasting thunderstorms in the summer convective season, as it is used as an indicator of low-level moisture content. Local studies have determined the minimum dewpoint temperature may be significant in determining precipitation chances across the Greater Phoenix Area (e.g., Haro and Bruce 1997).

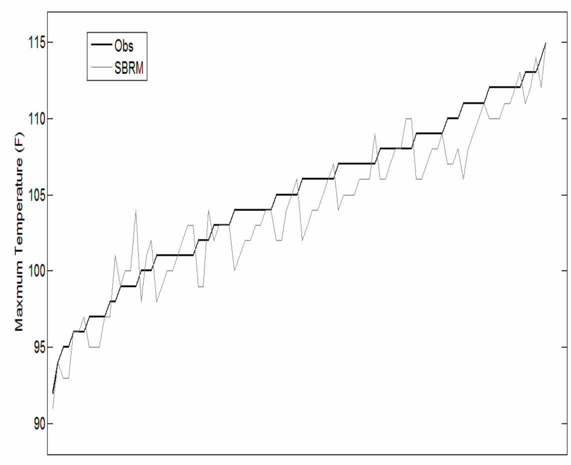

Figure 1. *Comparison of the observed and predicted maximum temperatures in 2008. The observed values are sorted in ascending order.*

A stepwise forward selection procedure is applied to develop the minimum dewpoint model. Variables included in the model are the observed dewpoint values at the surface, 850 hPa, 700 hPa, and 500 hPa, which are all positively correlated. The single non-moisture variable included in the dewpoint model is the 850 hPa to 500 hPa lapse rate, which is negatively correlated with the minimum dewpoint (i.e., smaller lapse rates lead to higher predicted values).

The SBRM prediction of the minimum dewpoint is within approximately 4º F of the observed minimum dewpoint, on average, while the standard deviation is approximately 5.4º F. A persistent slight warm bias is observed. A much higher degree of variability and uncertainty is found with the prediction of the minimum dewpoint, and is evident on a graphical comparison between the observed and predicted minimum dewpoint temperatures (Fig. 3).

3.3 Probability of precipitation

Precipitation events during the summer convective season can often be difficult to predict. Variables derived from the Phoenix, AZ sounding that may be relevant to the prediction of precipitation have been identified. Statistically significant variables included in the SBRM PoP are limited to moisture and wind variables (Table 3). The mean relative humidity between the surface and 500 hPa level is the most statistically significant predictor.

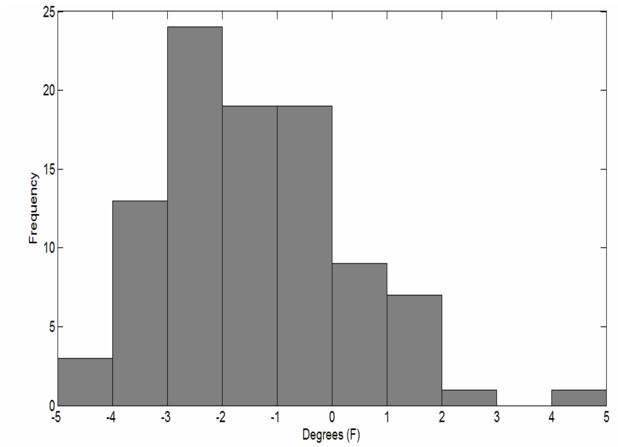

Figure 2. *Histogram of the difference of predicted and observed maximum temperatures.*

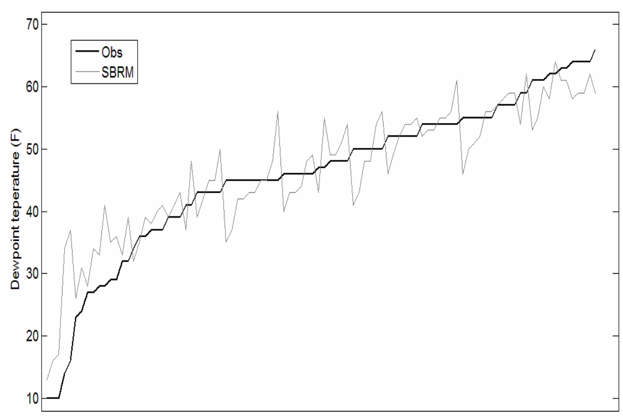

Figure 3. *Comparison of the observed and predicted minimum dewpoint temperatures in 2008. The observed values are sorted in ascending order.*

Stepwise logistic regression analysis was applied to develop the PoP prediction model. The SBRM PoP value represents the probability of precipitation occurring at any location within the network of gauges, based on historical data. For example, if a PoP value of 50% is given by the SBRM, precipitation should occur one out of every two times.

An analysis of the SBRM PoP prediction model performance in 2008 reveals more frequent occurrences of precipitation as the PoP value increases (Fig. 4). Precipitation occurred for every instance of a predicted PoP of 0.82 or greater. A slight over-prediction of PoP occurred at values between 0.1 and 0.4, as well as between 0.7 and 0.9 (Fig. 5).

Variable	Correlation
850 hPa Td	+
300 hPa Wind Direction	-
Surface-500 hPa Mean RH	+
925-500 hPa Mean Wind Speed	+
925-500 hPa Mean Wind Dir.	-

Table 3. *A list of the variables included in the PoP prediction model. The sign of the correlation between each individual predictor and the predicted PoP is listed. Wind directions are in degrees, with westerly (+) and easterly (-).*

3.4 Quantitative precipitation

The average precipitation falling across all included rainfall gauges within the network is predicted daily. However, the relative humidity between the surface and the 500 hPa level was the only variable found to be a statistically significant predictor. As a result, performance of the SBRM average precipitation prediction in 2008 was poor (not shown). Precipitation processes, particularly related to accumulations, are difficult to capture with a single sounding, as many effects (e.g, terrain) may locally enhance precipitation. The maximum precipitation falling at any gauge within the network was also predicted, but produced similar results.

4. SUMMARY

The SBRM attempts to predict four meteorological variables based off the Phoenix, AZ, 1200 GMT sounding. Three of the predicted variables, the maximum temperature, minimum dewpoint, and probability of precipitation, each demonstrated considerable skill. In addition, specific predictors that were found to be statistically significant have been identified.

The SBRM will continue to be run operationally in 2009 and beyond. Additional improvement of the models will likely occur with the inclusion of more historical data. Correction of model biases will be applied to provide more accurate forecasts. The SBRM may also be applied to additional sites in future improvements. Routine soundings are conducted at Flagstaff (KFGZ), Tucson (KTUS), and Yuma (K1Y7). Inclusion of the entire sounding, rather than using only mandatory levels, will be explored in an attempt to improve the accuracy of the SBRM. A desire for increased lead time has been expressed by some forecasters at WFO Phoenix. While it is difficult to modify the existing SBRM regression models based off of a fixed sounding time, generating additional regression models based off the 0000 GMT soundings will provide additional lead time.

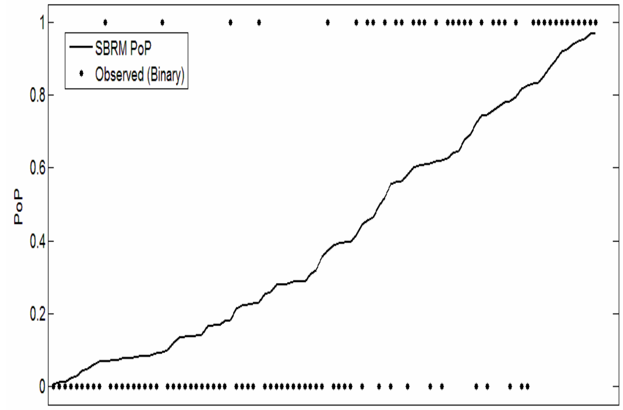

Figure 4. *Predicted PoP and observed precipitation events in 2008. The predicted values are sorted in ascending order.*

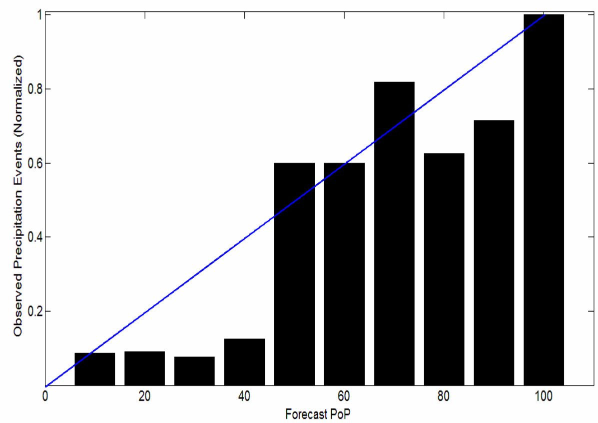

Figure 5. *Plot of cumulative predicted PoP values with respect to observed precipitation events in 2008.*

5. REFERENCES

Dixon, P. Grady, 2005: Using sounding data to detect gulf surges during the North American Monsoon. *Mon. Wea. Rev.*, **133**, 3047-3052.

Haro, Jesus A. and Michael A. Bruce, 1997: A brief climatology of the relationship between afternoon dewpoints and monsoonal rainfall at Phoenix Sky Harbor International Airport. WR- Technical Attachment 97-32.

Harris, D. Lee and Aldo Angelo, 1963: A regression model for storm surge prediction. *Mon. Wea. Rev.*, **91**, 710-726.

Maddox, Robert A., Darren M. McCollum, and Kenneth W. Howard, 1995: Large-scale patterns associated with severe summertime thunderstorms over central Arizona. *Wea. Forecasting*, **10**, 763-778.

Massie, Darrell R. and Mark A. Rose, 1997: Predicting daily maximum temperatures using linear regression and Eta geopotential thickness forecasts. *Wea. Forecasting,* **12,** 799-807.

Ryan, T. P., 1997: Modern Regression Methods. John Wiley and Sons, Inc., New York, NY.