IMPROVED SHORT-TERM PROBABILISTIC FORECASTS OF CEILING AND VISIBILITY

Stephen M. Leyton * and J. Michael Fritsch
Pennsylvania State University, University Park, Pennsylvania

1. INTRODUCTION

For many years and still today, statistically post-processed numerical model output has been a primary source of guidance for forecasting surface weather conditions. In particular, Model Output Statistics (MOS) derived from the Nested Grid Model (NGM) and, more recently, the Aviation Model (AVN) provide forecasts of weather parameters at three-hour intervals out to 60 hours. These models are run four times daily, allowing MOS forecasts to become six hours old before an updated product is made available. Therefore, it may be possible to improve conventional guidance for short-term (< 6 h) forecasts of selected parameters if *observations* taken after the time that the MOS guidance is derived are used to produce updated forecasts.

Such an "observations-based" approach for producing short-term forecasts was developed by Vislocky and Fritsch (1997; hereafter VF). Specifically, a network of surface observations was used as predictors in a multiple linear regression technique. It was demonstrated that this approach could improve the accuracy of ceiling and visibility forecasts for the hours in between the times that the output from the twice-daily operational runs is released.

Following this development of observations-based forecasting systems using reports from the standard synoptic observing network, increasing amounts of surface weather data became available as a result of upgrades in the coverage of Automated Surface Observation Systems (ASOS) and Automated Weather Observation Systems (AWOS) by the Federal Aviation Administration (FAA) and National Weather Service (NWS). It is possible that this availability of higher density surface weather observations may yield further improvements in observations-based statistical forecasting systems. Therefore, a study was undertaken to quantify whether or not the increased spatial

* *Corresponding author address:* Stephen M. Leyton, 503 Walker Building, University Park, PA 16801; e-mail: leyton@psu.edu

resolution of surface observing networks yields improvements in short-term forecasts of ceiling and visibility.

2. DATA

2.1 *Data Sets*

Two data sets are used in the development of the statistical forecast equations. The first data set is an archive of standard hourly surface (SAO) data from the mass storage system of the National Center of Atmospheric Research (NCAR). This data set spans from October to March, 1982-83 through 1995-96, and includes reports of temperature, dew point, wind speed and direction, cloud cover, ceiling, visibility and precipitation. This data set is used solely to compute the climatology at each site in the domain. The second data set is an archive of standard hourly surface (ASOS) data, also acquired from NCAR. This data set spans the October to March period, 1996-97 through 2000-01, and includes reports of temperature, dew point, wind speed and direction, cloud cover, visibility and precipitation. This data set is also used in development of climatology as well as the forecast equation development.

2.2 *Data Preprocessing*

Data preprocessing requires checking the data for any inconsistencies, often referred to as "bad" data, identifying instances of missing data, and developing methods to replace the bad or missing data with surrogate values. Of particular importance are the procedures to replace bad or missing data. A primitive scheme for deriving surrogate values was applied by VF, using an adjusted climatology from the nearest neighbor for bad or missing temperature, dew point and wind observations. All other parameters were replaced by that parameter at the nearest neighbor.

However, there are many occasions when this replacement system would fail to provide the best surrogate value. Such scenarios include a nearest neighbor with very different topographical surroundings or the passage of a

squall line between the two sites. Therefore, a more sophisticated replacement system was deemed necessary to provide more accurate surrogate values for a range of scenarios. Methods involving climatology, conditional climatology, persistence, and nearest-neighbor techniques are considered for all weather parameters.

Following the development of climatology and conditional climatology data sets for each site, the bad or missing data can be replaced with surrogate values. The replacement is accomplished in the following manner:

i) determine the departure from climatology (conditional climatology for instances with bad/missing ceiling and visibility) at each reporting station,

ii) for a station with missing or bad data, use the five nearest neighbors in a Cressman distance-weighting scheme to estimate the departure (for the conditions at that time) at the site requiring replacement data, and,

iii) add the estimated departure value to the appropriate climatology (conditional climatology) value.

From the surrogate values for temperature and dew point, the dew point depression and relative humidity are calculated. Persistence is used to replace bad or missing cloud cover in the same manner except that the departure value is added to the previous hour's observation at the site requiring replacement data. Bad or missing precipitation is replaced by precipitation information from the nearest neighbor.

2.3 *Predictands*

The problem of low ceilings and visibilities is one that affects all locations in each of the domains throughout the year, most notably during the winter months being studied. However, developing and testing forecast equations for all sites in each domain would be a tedious task. Nevertheless, forecast equations must be developed and tested for a sufficient number of stations in order to properly identify the forecast improvements provided by these new data sets.

For this study, 10 stations were selected; Waterloo, IA (ALO), Cedar Rapids, IA (CID), Des Moines, IA (DSM), Fort Dodge, IA (FOD), Fairmont, MN (FRM), Mason City, IA (MCW), Mankato, MN (MKT), Worthington, MN (OTG),

Ottumwa, IA (OTM), and Rochester, MN (RST). These sites were chosen because they are centrally located within the high-density network and have a history of reliable data.

Several thresholds of ceiling and visibility are used as predictands at each forecast site. Threshold ceiling heights of 500 ft, 1000 ft and 3000 ft are considered. Threshold visibilities of 1 mi. and 3 mi. are also considered. These threshold values were selected because the FAA uses them to determine flight rules.

2.4 *Predictors*

Consistent with the work of VF, the set of potential predictors is comprised of 27 observational terms in either binary or continuous format. A binary predictor is one in which if the variable exceeds a desired threshold, then a "1" is assigned. Otherwise, a "0" is used. A continuous predictor simply equals the observational value. Also, following the work of VF, two classes of predictors are developed. The first class consists of the 27 observational terms at the forecasting site and each of a selected group of nearest neighbors. The other class consists of smoothed observational values for all non-predictand variables. These smoothed values are calculated by averaging the observational value at the forecasting site and a selected number of nearest neighbors.

Pilot studies revealed many similar results to those found by VF:

- The optimal number of nearest neighbors to be considered in equation development increases with lead time: For a 1-hr forecast, predictor observations from the 10 nearest stations to the forecast site were optimal. For a 3-hr and 6-hr forecast, observations at the 30 and 45 nearest stations were optimal, respectively.

- Observations of the predictand variable at the forecast site and its nearest neighbors offered by far the most predictive information

- Some additional accuracy was gained by offering each of the observational predictors at the forecast site as well as those in the spatially smoothed format described previously

As a result of these pilot studies, the final predictor set considered for the final baseline system consisted of observations of ceiling and visibility at the forecast site and the 10, 30, and

45 nearest neighbors for the 1-, 3-, and 6-hr forecasts, respectively. In addition, the observations of the predictor terms from the forecast site as well as the spatially smoothed values of these predictor terms were also included in the final predictor set. Lastly, several climatic terms were also included in the final list of potential predictors, including the sine and cosine of the day of year along with the climatology of the predictand variable at the forecast site, similar to VF.

When expanding the pool of potential predictors to include observations from the high spatial resolution network, the only additional consideration is the number of nearest neighbors to be included. It is important to note that the radius encompassing the 10 nearest neighbors in the standard observing network contains many more stations in the high spatial resolution network. Thus, a significantly larger number of nearest neighbors must be considered. For example, in the baseline system at DSM, the nearest 10 neighbors are within a radius of 120 miles. In the high spatial resolution system, 38 neighbors are within that same radius of 120 miles. By using these similar radii, the high spatial resolution forecast equations should have no less skill than the baseline system because even if the high-density sites are not beneficial, the same stations used for the baseline system are still available for potential predictors. Therefore, on average, for the 1-hr (3-hr, 6-hr) forecasts, 40 (75, 100) nearest neighbors will be offered in the pool of potential predictors.

2.5 *Equation Development*

Ideally, forecast equations should be developed for each hour of each day to provide the optimal guidance for ceiling and visibility. Unfortunately, due to limitations in observational data available, this process would result in too few cases and even fewer events. So, the next best method is to develop forecast equations for each hour of the day, with no regard given to day of the year (e.g. 0300 and 1500 UTC). With a dependent data set that only spans four winter periods, however, this yields only approximately 600 cases. While this is an improvement, the ratio of cases to potential predictors is insufficient for reliable statistical equation development.

Therefore, consideration is given to computing one forecast equation for all hours of the day at each forecast site. Since station

climatologies show that there is little diurnal variation in any of the thresholds, it was decided that all hours of the day could be used for producing a single group of best predictors for each predictand at various lead times. This process yields approximately 14,500 cases and a more statistically acceptable ratio of cases to predictors.

Therefore, separate regression equations were developed for each predictand and for three lead times (1-, 3-, 6-hr) at each forecast site. All equations were developed independently of one another. Thus, each regression formula contains its own assortment of predictors, developed for each of the 10 sites in the domain.

3. RESULTS

3.1 *Application to Independent Data*

All of the developed equations were applied to independent data sets. A cross-validation technique was applied such that one winter period was used as the independent data set for the equations created from the other four winter periods combined (the dependent data set). Any observation in the independent data set that did not possess a real value (i.e. not a surrogate value) for the predictand was eliminated from the verification data set. However, observations with surrogate predictors remained.

3.2 *Comparison to Persistence Climatology Forecasts*

Consistent with the work by VF, the performance of the baseline and high-density observations-based forecast systems were compared to the performance by persistence climatology. Persistence climatology is the climatology of the relationship between an initial time condition and the presence or absence of that condition at a later time. In order to make this comparison, the mean squared errors (MSE) were computed as benchmarks to the performance of all models in each station, predictand and lead-time combination. The MSE is computed by first calculating the squared difference between the forecast probability (0 to 1, inclusive) and the verification (0 or 1) for each of the cases in the independent data sets. An average of all these squared differences results in the MSE.
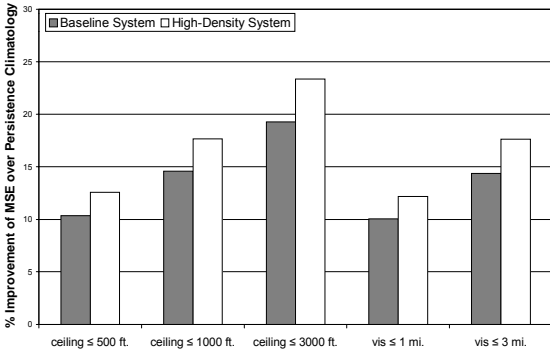
Figure 1. Summary of 1-hour forecasts

Figure 2. Summary of 3-hour forecasts

Figure 1 is a plot of the percent improvement of the MSE over persistence climatology for both the baseline forecasting system and the high-density forecasting system when considering a 1-h lead time. The graph shows a consistent improvement by incorporating the high-density observing sites into the standard synoptic network of stations. Overall, the high-density forecasting system provided an additional 2%-4% reduction in MSE to that of the baseline forecasting system. This reduction in MSE corresponds to a 20%-25% increase in skill.

Figure 2 is a plot of the forecast system performance when considering a 3-h lead time. The graph shows a lesser improvement when including the high-density sites. Overall, the high-density forecasting system provided up to an additional 1.5% reduction in MSE to that of the baseline forecasting system; i.e. up to an 8% improvement over the baseline forecasting system.

As expected, the high-density observing network provided little to no improvement over the baseline forecasting system when considering a 6-h lead time. This result was anticipated since VF found that observations-based forecasts using a 6-h lead time were at the limit of their success. Thus, the inclusion of additional sites within the same radius should not improve upon the forecasts developed from a less-dense network of stations.

4. SUMMARY

Accurate short-term forecasts of low ceiling and visibility are vital to air traffic operations, which try to maximize the use of an airport at all times. Therefore, when impeding phenomena occur, it is desirable to have the best guidance available. It was shown by VF that an observation-based forecast system improved
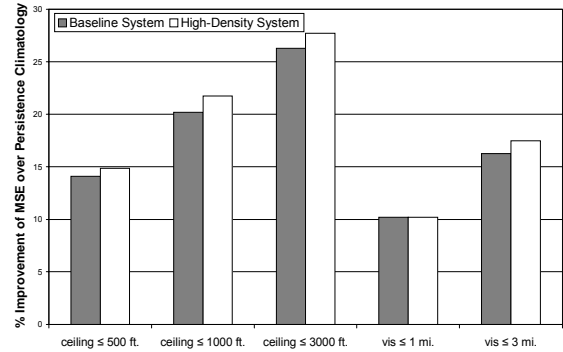
upon those techniques previously used for short-term forecasts of ceiling and visibility. The goal of this study was to examine whether an increased density of observing sites would improve the performance of that system.

The results presented from this study are encouraging. It was found that the inclusion of these additional sites into the forecasting system provided an additional 2%-4% reduction in MSE when forecasting with a 1-h lead time. With a 3-h lead time, the improvement was reduced to 0%-1.5% reduction in MSE. At the 6-h lead time, the higher density of observing sites provided little forecast improvement.

Inspection of individual events indicates that the improvements in the very short-term forecasts stem from the ability of the higher-density observing network to more correctly characterize a given event. For example, while it is true that there are instances wherein large areas are *uniformly* and *persistently* blanketed with low ceiling and visibility, analyses of individual events revealed that some events are more aptly described as having intermittent, patchy or isolated low ceilings/visibilities. Therefore, the chances of mischaracterizing an event are greater when there are only a few observations. Thus, a higher density of predictors (i.e., more observing stations) can more clearly distinguish among the various types of events and provide more accurate probabilistic categorical forecasts.

5. REFERENCES

Vislocky, R. L., and J. M. Fritsch, 1997: An automated, observations-based system for short-term prediction of ceiling and visibility. *Wea. Forecasting*, **12**, 31-43.