J5.5 AUTOMATED FORECASTING OF CEILING AND VISIBILITY FOR AVIATION: EXPLORATION OF A DATA MINING-BASED FORECAST METHOD

Paul H. Herzegh*, Gerry Wiener, Richard Bateman and Beth Weekley

National Center for Atmospheric Research, Boulder, CO 80303

Special Note: At the time of final preparation of this manuscript, the authors determined that subtle data processing errors had degraded the statistics used to document the performance of the data mining-based forecast method under development. We have chosen to withhold the Results section until corrected statistical information is available.

Please direct your browser to ftp://ftp.rap.ucar.edu/incoming/irap/herzegh/ to download this paper in complete form reflecting corrected statistics. We expect our completion to become available on or before 1 December. Thank you for your interest in this work.

1. INTRODUCTION

Observations-based forecast techniques utilize site-specific data archives to relate current and recent conditions to those that can, based on past history, be expected in the near future. These techniques are well known to have particularly good skill from 1 to 6 h, and thus offer significant practical forecast value. Once predictive relationships have been established for a particular site, for specific target variables, and for specific forecast lead times, practical use and maintenance of the methods are relatively straightforward. Forecasts can be quickly generated in real time using observations of current and recent past conditions.

The practical benefits of observations-based (obs-based) techniques strongly encourage full exploration of means to efficiently develop and apply them. One well-proven approach (Leyton and Fritsch, 2003) utilizes multiple linear regression to relate multiple predictors to a specific target predictand. Other techniques are built around the use of neural networks and logistic regression. Recently, the widespread commercial development of data mining techniques has broadened capability for computationally-intensive searching and processing, and has yielded practical opportunity to explore new obs-based forecast methods based upon powerful new classification and estimation algorithms.

Obs-based classification techniques are well suited to categorical prediction of flight conditions, where improved forecasts over the range 1-6 h carry significant benefit to both aviation safety and airport efficiency. U.S. flight regulations break ceiling and surface visibility values into specific intervals that dictate certain restrictions on flight operations. Instrument Flight Rule (IFR) conditions are associated with ceiling values less than 1000 feet and/or surface visibility values less than 3 n.mi. Visual Flight Rule (VFR) conditions are associated with ceiling and visibility values both greater than or equal to the IFR thresholds outlined above.

In this study we begin assessment of obs-based classification methods for use in an algorithm yielding categorical forecasts of IFR vs VFR conditions over 1, 3, 5 and 7 h forecast lead times. A prototype forecast system is implemented using C5.0, a commercially-available rule-based decision tree classification algorithm (Rulequest Research, 2004).

2. TECHNIQUE DESCRIPTION

The data mining (DM) process begins with an algorithm training stage described in Table 1. In this stage a predictive ruleset is derived through a data mining exercise focused on the long-term archive of observations at a specific site, a chosen target variable and a specific forecast lead time. Since the objective of this work is to begin assessment of the skill of the DM forecast method outlined, we next define a corresponding test stage described in Table 2. In the test stage, the DM forecast process is systematically applied for the same target site at the top of each hour throughout the duration of the test period.

The training period utilizes data from 1980 to present, but excludes the test period chosen to assure independence. Since the test period duration used here is 2 years (2003-2004 unless otherwise noted), the effective duration of the training period is 21.7 years.

The resulting DM forecasts specify either IFR or VFR conditions for ceiling and visibility. Specification of IFR conditions is considered an IFR event, while specification of VFR conditions is considered a null event. Accumulated forecast results over the

* *Corresponding author address*: Paul H. Herzegh, NCAR, P.O. Box 3000, Boulder, CO 80307;
Email: herzegh@ucar.edu

Table 1. Training Stage Procedure

Step	Comment
1. Specify target forecast site.	The forecast site must offer a long-term archive of observations at the site and ongoing access to real-time observations.
2. Select duration of training period.	Longer is generally better.
3. Perform quality control of training data archive for forecast site.	Assures removal of faulty or questionable observations. Quality control in the present study is limited to confirmation that data is present.
4. Select desired forecast dependencies.	- Forecast period (in this study, either 1, 3, 5 or 7 h). - Forecast initiation times correspond to the top of the hour. - Input times: use data at initiation time plus optional preceding times to capture tendencies, etc. - Single-site or multi-site? Specify any neighboring site(s) to be used. For example, taking into account conditions at one or more neighboring sites may improve forecast performance at the target site. - Select data parameters to include in the rulesets. For example, current (and optionally past) values of ceiling and visibility, wind speed and direction, dewpoint depression, date, time of day, etc. - Establish data mining options. Each can be invoked or ignored. 'Boosting' applies user-defined weights used by the internal rule evaluation process to tune ruleset selection to favor a desired result. For example, weights can be set to favor detection of IFR conditions over missing an IFR forecast. This preferentially yields a higher probability of IFR detection at the expense of a corresponding higher false alarm rate. 'Winnowing' reduces the number of input observation types selected by the user by ignoring individual data types that are shown to be less effective elements within the derived ruleset.
5. Populate training period database with qualified data.	- Assures complete forecast initiation data (data from all required times for target site and any neighboring sites selected are present). - Assures presence of necessary valid time data needed for verification.
6. Perform data mining.	Execute C5.0 classification algorithm to produce a forecast ruleset.

Table 2. Test Stage Procedure

Step	Comment
1. Select duration of independent test period.	To gain a reasonable sampling of conditions at a given site, a 2-year test period is used unless otherwise noted. This interval lies outside the training period.
2. Perform quality control on test data archive for forecast site.	As in Step 3 of Table 1.
3. Populate test period database with qualified data.	As in Step 5 of Table 1, assures that all necessary data are available for forecast and verification.
4. Perform data mining (DM) forecast exercise.	DM forecasts are generated for all valid initiation times within the test period using the ruleset derived in Step 6 of Table 1.
5. Perform corresponding persistence forecast exercise.	For each DM forecast generated in Step 4, a corresponding persistence forecast is produced and recorded.
6. Assess DM and persistence forecast skill through statistical analysis of forecast hits, misses and false alarms.	Statistical analysis describes the results accumulated over all forecasts within the test period for both DM and persistence methods. Utilizing data on forecast hits, misses, false alarms and null events, calculate quantities such as bias, probability of detection, false alarm ratio and Pierce skill score.

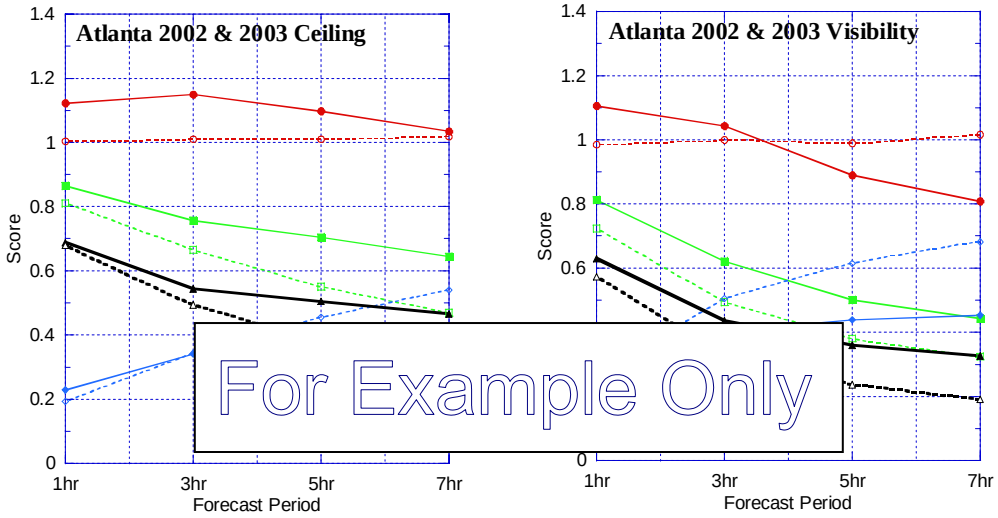

*Figure 1. Plots of **bias** (red), **probability of detection** (green), **false alarm ratio** (blue) and **critical success index** (black) for single-site DM forecast lead times of 1, 3, 5 and 7 hours for the occurrence of IFR conditions in ceiling and visibility at the sites shown. Solid lines show the DM forecast scores. Dashed lines show the corresponding persistence forecast scores.*

test interval (e.g., frequency of correct forecasts, false alarms, missed forecasts and null events) are used to calculate the skill measures and statistical quantities discussed in Sec 3. We choose to compare DM forecast skill with that of persistence, which is itself a particularly skillful forecast method over short lead times such as those examined

3. FORECAST TEST RESULTS

Single-Station Trials

To be revised.

Multi-Station Trials

To be revised.

5. SUMMARY AND DISCUSSION

We describe early trials of a new technique in observations-based forecasting that utilizes data mining of long-term data archives at selected sites in conjunction with a classification model to produce rulesets for 1-7h forecasting of ceiling and visibility. Early results are encouraging in that the data mining forecasts generally exceed the performance of persistence at 1-7h.

Discussion to be continued.

Acknowledgements

This research is in response to requirements and funding by the Federal Aviation Administration (FAA). The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA.

REFERENCES

*Leyton, S.M. and J.M. Fritsch, 2003: Short-term probabilistic forecasts of ceiling and visibility utilizing high-density surface weather observations. Weather and Forecasting, **18**, 891-902.*

Rulequest Research, 2004: C5.0 and Cubist. http://www.rulequest.com