

Data Mining Numerical Model Output for Single-Station Cloud-Ceiling Forecast Algorithms

RICHARD L. BANKERT AND MICHAEL HADJIMICHAEL

Naval Research Laboratory, Monterey, California

(Manuscript received 15 September 2006, in final form 29 November 2006)

ABSTRACT

Accurate cloud-ceiling-height forecasts derived from numerical weather prediction (NWP) model data are useful for aviation and other interests where low cloud ceilings have an impact on operations. A demonstration of the usefulness of data-mining methods in developing cloud-ceiling forecast algorithms from NWP model output is provided here. Rapid Update Cycle (RUC) 1-h forecast data were made available for nearly every hour in 2004. Various model variables were extracted from these data and stored in a database of hourly records for routine aviation weather report (METAR) station KJFK at John F. Kennedy International Airport along with other single-station locations. Using KJFK cloud-ceiling observations as ground truth, algorithms were derived for 1-, 3-, 6-, and 12-h forecasts through a data-mining process. Performance of these cloud-ceiling forecast algorithms, as evaluated through cross-validation testing, is compared with persistence and Global Forecast System (GFS) model output statistics (MOS) performance (6 and 12 h only) over the entire year. The 1-h algorithms were also compared with the RUC model cloud-ceiling (or cloud base) height translation algorithms. The cloud-ceiling algorithms developed through data mining outperformed these RUC model translation algorithms, showed slightly better skill and accuracy than persistence at 3 h, and outperformed persistence at 6 and 12 h. Comparisons to GFS MOS (which uses observations in addition to model data for algorithm derivation) at 6 h demonstrated similar performance between the two methods with the cloud-ceiling algorithm derived through data mining demonstrating more skill at 12 h.

1. Introduction

Numerical weather prediction (NWP) models provide guidance for operational forecasters with useful insight into future atmospheric conditions. However, generating accurate cloud-ceiling forecasts from these models remains a difficult and complex problem. Accurate cloud-ceiling analysis and forecasts are important for aviation operations, both in terms of monetary costs and human safety. The Federal Aviation Administration's (FAA's) Aviation Weather Research Program (AWRP) is investigating this problem through a research and development effort of the National Ceiling and Visibility (NCV) Product Development Team (PDT). The forecast product being developed by the NCV PDT involves the integration of numerical mod-

els, guidance products, and observation-based methods (Herzegg et al. 2006).

Model output statistics (MOS) (Glahn and Lowery 1972) was one of the first methodologies employed to find a statistical relationship between NWP model output and cloud-ceiling height (Bocchieri and Glahn 1972; Bocchieri et al. 1974). MOS schemes are still applied to many of today's numerical models. Along with NWP model data, observations are also used to develop MOS regression equations for ceiling and visibility, temperature and dewpoint, wind speed and direction, probability of precipitation, precipitation amount, and cloud cover. MOS techniques, applied to various NWP models, have also been used to develop specific algorithms to forecast the occurrence of Levante wind regimes during the warm season (Godfrey 1982), marine fog probability for the northern Pacific Ocean (Koziara et al. 1983), maximum and minimum temperatures for seven Australian cities (Woodcock 1984), and the probability of precipitation and rain amount in Australia (Tapp et al. 1986). More recently, Norquist (1999) pre-

Corresponding author address: Richard Bankert, Naval Research Laboratory, 7 Grace Hopper Ave., Monterey, CA 93943-5502.

E-mail: rich.bankert@nrlmry.navy.mil

DOI: 10.1175/WAF1035.1

sented a MOS approach to diagnose cloud characteristics, including cloud-base height, from a mesoscale NWP model. Weiss and Ghiradelli (2005) describe the redevelopment of the Localized Aviation MOS Program (LAMP). LAMP focuses on the predictions of probabilities of eight ceiling-height and five total-sky categories.

Model data and MOS schemes are not the only resources being used to predict cloud cover and ceiling conditions. Vislocky and Fritsch (1997) investigated an observation-based system for short-term ceiling and visibility prediction using a network of surface stations. Other observation-based ceiling estimation and prediction schemes recently developed include a statistical system for low-ceiling forecasts at San Francisco International Airport using surface and upper-air observations (Hilliker and Fritsch 1999), an algorithm that combines satellite-derived cloud classification data and surface observations to improve cloud-base height estimates (Forsythe et al. 2000), and the use of total-sky imagers for the retrieval of cloud-base heights (Kasianov et al. 2005). Leyton and Fritsch (2003, 2004) describe the utilization of high-density and high-frequency surface observations for short-term probabilistic forecasts of ceiling and visibility. A cloud-ceiling and visibility forecast system that applies fuzzy logic and case-based reasoning to historical data is presented in Hansen and Riordan (2003).

MOS and observation-based methods rely on observations as predictors within the algorithms or regression equations. An algorithm that can be applied to an NWP model grid directly (without using observations) would be beneficial, particularly when and/or where observations are not available. Other research and applications of NWP model data for specific weather phenomena include the use of pattern recognition in locating fronts in model output (Fine and Fraser 1990), model output calibration for temperature forecasts (Mao et al. 1999), the training of neural networks for temperature forecasts for 31 stations (Marzban 2003), maximum and minimum temperature forecasts at 12 locations in India based on the perfect prog method (Maini et al. 2003), the application of quantile regression for probabilistic precipitation forecasts (Bremnes 2004), and turbulence estimations (Frehlich and Sharman 2004).

With the work in Hansen and Riordan (2003) being one example, machine learning, data mining, and other artificial intelligence tools are being used more often in the diagnosis and forecasting of meteorological phenomena. Abdel-Aal and Elhadidy (1995) applied a machine-learning modeling tool for forecasting daily maximum temperatures in Dhahran, Saudi Arabia. Induc-

tion of rules has been applied to develop maritime fog forecasts (Tag and Peak 1996), and, as mentioned earlier, neural networks have been used for temperature forecasts (Marzban 2003). There has been much research and many algorithms developed in the area of cloud classification that included the use of various machine-learning tools. One of the more recent classification methodologies has involved the use of support vector machines (Lee et al. 2004). Classification trees and machine learning are used to develop rules for predicting the onset of Australian winter rainfall (Firth et al. 2005). Related to this effort, Bankert et al. (2004) demonstrated the use of data-mining methods in the diagnosis of cloud-ceiling height through satellite and numerical model data. Data-mining techniques have also been applied to the discovery of associations between drought and oceanic parameters (Tadesse et al. 2005).

Data mining is applied here to NWP model output to create a station-specific (John F. Kennedy International Airport, New York, New York, in this experiment) translation algorithm for the Rapid Update Cycle (RUC) model. The potential of this approach will be analyzed and future enhancements and application of this methodology to an entire model grid will be discussed.

2. Data and data mining

RUC is a regional mesoscale high-frequency data assimilation and short-range numerical prediction system (Benjamin et al. 2004a,b). RUC runs on a 1-h assimilation cycle to produce frequently updated 3D mesoscale analyses and short-range forecasts. The version of the RUC model that produced the data used here is RUC20. The RUC20 horizontal resolution is 20 km (301×225 grid points) and the model uses a hybrid isentropic-sigma coordinate (50 vertical levels) in which most of the atmosphere is resolved on isentropic surfaces except for layers near the ground where sigma coordinates are used. The RUC20 microphysics scheme (Brown et al. 2000) explicitly predicts mixing ratios of cloud water, rainwater, snow, cloud ice, and graupel. Ice particle number concentration is also predicted. The model also employs the Grell–Devenyi ensemble-based convective parameterization scheme (Grell and Devenyi 2001).

RUC data from the entire year of 2004 were made available for use in the development of single-station cloud-ceiling forecast algorithms. The hourly RUC model output was saved in a database for data-mining exploration. Using a bilinear interpolation of neighboring grid points, the values of RUC variables at various routine aviation weather report (METAR) locations

TABLE 1. RUC output variables stored in database of hourly records. The lowest model level is at 5 m.

Temperature at lowest model level
Dewpoint temperature at lowest model level
RH at lowest model level
u wind component at lowest model level
v wind component at lowest model level
Tot wind speed at lowest model level
Sensible heat flux at surface
Latent heat flux at surface
Bowen ratio
Cloud–no cloud (yes–no)
Cloud-top height
Cloud-top temperature
Ground temperature
Ground moisture availability
Snow cover–no snow cover (yes–no)
PBL depth
Vertically averaged TKE in PBL
Height of model level with max vapor mixing ratio
u wind component 0–30 mb AGL
v wind component 0–30 mb AGL
Tot wind speed 0–30 mb AGL
Net longwave radiation
Net shortwave radiation
Avg RH in lowest 150 mb
Temperature difference (top – bottom) in lowest 150 mb
Precipitable water
PWR (lowest 150 mb)
Richardson number (using lowest four model levels)
LCL

compose the database. Developed algorithms, applied to RUC model output, will produce ceiling estimations at specific locations (or grid points). The cloud-ceiling height is defined here as the lowest level at which broken or overcast cloud cover exists. Ceilometer measurements provide the data for METAR cloud coverage (NWS 1998). KJFK is the METAR station at John F. Kennedy International Airport and was the focus of this study.

The RUC variables selected for inclusion in the database are listed in Table 1. They were chosen based on a priori assumptions of their potential influence or relationship to cloud-ceiling height. The available RUC data are 1-h forecast data valid at nearly every hour in 2004. Not all variables are actually used by the final classification algorithms.

After establishing the database, a data-mining tool was applied to find the relationships between the RUC data and 1-, 3-, 6-, and 12-h cloud-ceiling observations. C5.0 is the supervised inductive-learning tool used here to generate classifiers from data in the form of decision trees (Quinlan 1993; Bankert et al. 2004). It recursively breaks down a training set into smaller sets, using a single independent variable test at each iteration. Vari-

TABLE 2. Cloud-ceiling-height (AGL) classes.

Class	Height (ft)
VFR1	No ceiling
VFR2	>12 000
VFR3	>5000 and \leq 12 000
VFR4	>3000 and \leq 5000
MVFR	\geq 1000 and \leq 3000
IFR	\geq 500 and <1000
LIFR	<500

ables and test conditions on those variables are chosen such that each of the resulting subsets is as homogeneous as possible. The result is a set of classification decisions, organized as a decision tree, where each leaf of the tree denotes a final classification. With these empirically developed classification models expressed as decision trees, they are less of a “black box” and more open to interpretation than classifiers from other data-mining methods (e.g., neural networks). The decision tree size and complexity are determined by the training data (used to generate the trees) and the balancing of training set classification accuracy with the ability of the classifier to generalize to events not adequately represented in the training data. Decision tree “pruning” is implemented to approach this balance. Pruning removes decision boundaries and introduces an acceptable degree of error in order to compact, and thus generalize, the tree. Additional details on decision tree generation can be found in Quinlan (1993).

Data mining is initially performed to create a seven-class classifier algorithm. The seven classes, listed and defined in Table 2, were chosen based on the FAA flight regulations regarding cloud ceilings: visual flight rules (VFR), marginal visual flight rules (MVFR), instrument flight rules (IFR), and low instrument flight rules (LIFR). Cloud-ceiling heights that fall into the VFR category were further distributed into four classes as defined in Table 2. The total number of database

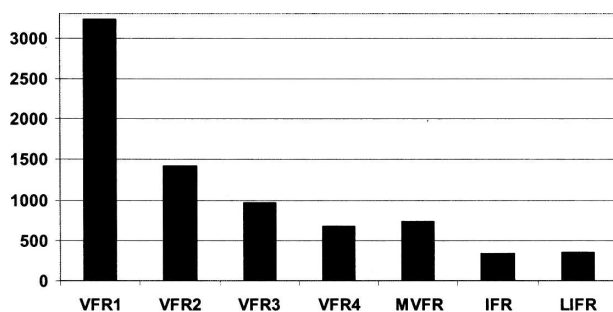


FIG. 1. Total number of KJFK hourly database records for each cloud-ceiling class. Both RUC and METAR data were available for each of these records.

TABLE 3. The 1-h cloud-ceiling forecast classification accuracies (overall and each class %) for various seven-class classifiers. Bottom numbers for overall accuracy column represent the six-class overall accuracy (VFR1 and VFR2 combined). Explanations of abbreviations can be found in the text and in Table 2.

	Overall	VFR1	VFR2	VFR3	VFR4	MVFR	IFR	LIFR
DMCC	66.7 75.2	88.6	55.1	45.8	39.2	56.3	42.0	67.3
Persistence	76.7 82.0	88.6	72.8	63.8	61.0	68.8	66.9	79.1
RUC: qc + qi	53.9 64.6	88.3	50.1	16.4	9.3	12.2	9.5	72.0
RUC: RH	47.5 59.5	84.9	8.4	24.5	12.3	20.0	5.3	90.8
RUC: SW	46.7 63.1	98.6	9.9	12.5	4.5	9.2	10.7	4.9

records used in this study was 7682 and these records were distributed among the classes as displayed in Fig. 1. To evaluate the algorithms developed through C5.0, 10 tenfold cross-validation (CV) tests were performed. For each tenfold CV, the data are randomly split into 10 data record subsets with each subset held out for testing and the data mining performed on the remaining data. The average skill and accuracy of the 10 tenfold CV tests are reported here.

3. Performance evaluation

a. 1-h forecasts

Using 1-h RUC data, 1-h cloud-ceiling-height classification algorithms were created first. Arbitrary discretization of continuous cloud-height data into the seven aviation-defined height categories blurs any natural “clusters” of the height data. These artificial class boundaries make for a more difficult classification problem. Initial results are presented for the seven-class problem, but additional evaluations for a two-class grouping are also shown.

The 1-h forecast overall accuracy and individual class accuracies are presented in Table 3 for the C5.0-developed algorithm [referred to as data-mining cloud ceiling (DMCC) from here on] for persistence (Murphy 1992; Wilks 2006), which is used as a standard of reference, and for three cloud-ceiling estimations from RUC output: the mixing ratio threshold (RUC: qc + qi), the relative humidity threshold (RUC: RH), and the Stoelinga–Warner translation algorithm (Stoelinga and Warner 1999) applied to RUC data (RUC: SW). The cloud-ceiling-height estimate from the mixing ratio data is the lowest level at which the combined cloud water (qc) and cloud ice (qi) mixing ratios exceed 10^{-6} g g^{-1} . For the RH threshold, the cloud-ceiling height is the lowest level of greater than 90% RH. The

Stoelinga–Warner translation algorithm is based on empirical and theoretical relationships between hydrometeor attributes and light extinction. The DMCC algorithm produced higher overall accuracy than did the current RUC estimates of cloud-ceiling height at 1 h. With the DMCC algorithm depending solely on RUC data (i.e., no historical or current observations used as predictors), persistence had a higher accuracy for 1-h forecasts. Including observations and/or satellite data should improve DMCC accuracies, as demonstrated in previous and related work for cloud-ceiling-height diagnosis (Bankert et al. 2004).

Note some of the relatively low individual class accuracies. Three of the seven classes had less than 50% accuracy for the DMCC algorithm. Specific RUC algorithms had more classes with even lower accuracies. This outcome is related, at least in part, to the problem of classifying continuous data. However, very few of the misclassifications for the DMCC algorithm (1.3% of all testing samples) would be considered very bad misclassifications, with a majority of all misclassifications fall-

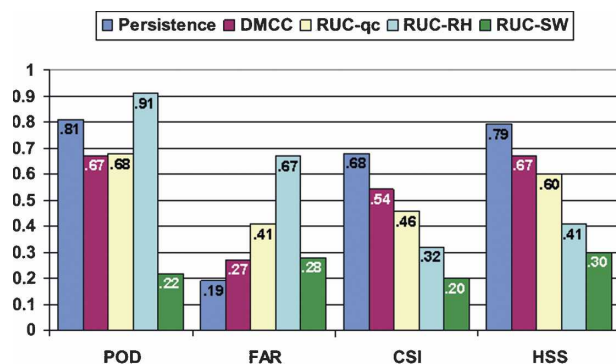


FIG. 2. Two-class (<1000 ft; ≥ 1000 ft) 1-h cloud-ceiling forecast accuracy and skill levels for different classification methods.

TABLE 4. The 3-h cloud-ceiling forecast classification accuracies (overall and each class %) for the DMCC and persistence seven-class classifiers.

	Overall	VFR1	VFR2	VFR3	VFR4	MVFR	IFR	LIFR
DMCC	65.3	87.7	51.3	45.7	37.6	57.1	41.1	63.4
Persistence	60.7	78.6	53.9	44.0	39.5	48.8	42.9	57.7

ing into neighboring height classes. As displayed in Table 3, the overall accuracy for each method improves substantially when combining VFR1 (no ceiling) with VFR2 (>12 000 ft ceiling) to produce a six-class outcome. As a baseline climatology, a 1-h forecast of “no ceiling” for all samples produces an overall accuracy of 42.0%.

To get a more extensive evaluation on the accuracy and skill of the DMCC algorithm compared with other methods, the output data were distributed into two classes: less than 1000 ft (<1000 ft) and greater than or equal to 1000 ft (≥ 1000 ft). The resulting 2×2 contingency tables produced for each classification were analyzed with low ceilings (<1000 ft) being the event of interest. As noted in Wilks (2006) and Doswell et al. (1990), three or more performance parameters are needed to fully represent the forecast performance. Probability of detection is a measure of the discrimination (or detection) capability. False alarm ratio is a reliability attribute and is the proportion of forecasted events (ceilings <1000 ft) that are incorrect. The low-ceiling class (<1000 ft) occurs much less frequently in this data source than does the high-ceiling class (≥ 1000 ft); therefore, critical success index (CSI) is a more useful accuracy measure than the total percent correct. The Heidke skill score (HSS) is computed to measure the performance of each algorithm relative to random chance. A positive HSS value indicates a performance better than random chance (perfect score is 1.0), an HSS value of zero indicates no better (or worse) than random chance, and a negative HSS value indicates forecast skill that is worse than random chance. These results are presented in graphical form in Fig. 2. The DMCC algorithm produced the closest performance level, when all statistics are considered, to persistence at 1 h.

b. 3-h forecasts

With 1-h RUC output being the only available data for this study, C5.0 data mining had to be applied to those data with appropriate 3-, 6-, and 12-h METAR observation times. For example, the 1-h RUC data with a 0300 UTC initial time (data valid at 0400 UTC) are paired with the METAR observation at 0600 UTC for a 3-h forecast from the initial time. It is not known

whether the DMCC algorithm developed in this manner would be more or less accurate than one developed using 3-h RUC data (valid at 0600 UTC in this example). The resulting forecast skill would be dependent on the skill of the 3-h RUC forecast, as well as the ability of the C5.0 algorithms to find the relationships.

Without 3-h RUC data, the DMCC 3-h cloud-ceiling forecast algorithm is compared with persistence only. The seven-class results are presented in Table 4. DMCC algorithm class accuracies compare very favorably to persistence at 3 h. In fact, it produces higher (nearly 5%) overall accuracy. As expected, the persistence accuracies at 3 h are much lower than those produced at 1 h. For the DMCC algorithm, there is not much difference between the 1- and 3-h accuracies. The VFR1 and VFR2 classes are combined into one class, increasing the overall accuracy by approximately 10% for both DMCC and persistence.

As was done with the 1-h results, the seven-class output for 3-h forecasts is regrouped into two classes (<1000 ft and ≥ 1000 ft). The performance measures are presented in Fig. 3 with the DMCC algorithm demonstrating slightly higher accuracy (CSI) and skill (HSS) than persistence.

c. 6- and 12-h forecasts

For both 6- and 12-h cloud-ceiling forecasts, Global Forecast System (GFS) MOS output was available for comparison. The seven classes output by GFS MOS are

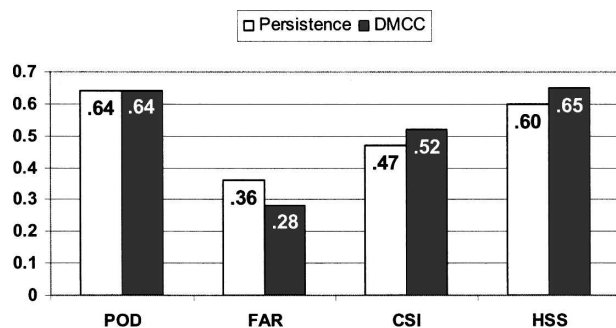


FIG. 3. Two-class (<1000 ft; ≥ 1000 ft) 3-h cloud-ceiling forecast accuracy and skill levels for persistence and DMCC classification algorithms.

TABLE 5. The 6-h cloud-ceiling forecast classification accuracies (overall and each class %) for DMCC, persistence, and GFS MOS six-class classifiers. The boundary between VFR3 and VFR4 is 5000 ft for DMCC and persistence and 6500 ft for GFS MOS.

	Overall	VFR1/VFR2	VFR3	VFR4	MVFR	IFR	LIFR
DMCC	72.9	91.3	43.1	36.9	50.2	37.4	62.0
Persistence	61.1	81.3	28.9	29.0	34.8	28.1	36.4
GFS MOS	68.3	83.5	35.3	52.4	49.0	37.2	54.2

slightly different than those used for DMCC and persistence. The VFR1 and VFR2 classes have been combined for DMCC and persistence to match GFS MOS. The two lowest-height classes used in GFS MOS have been combined to match the LIFR class. These class mergers create a six-class system. The comparison still is not perfect, with GFS MOS having a class boundary at 6500 ft as opposed to 5000 ft for persistence and MOS. However, a representative comparison is possible.

Six-class accuracy results for the three methods can be found in Table 5 for 6-h forecasts and Table 6 for 12-h forecasts. The two-class (<1000 ft and \geq 1000 ft) accuracy and skill evaluations are presented in Figs. 4 and 5 for 6 and 12 h, respectively. Not surprisingly, persistence fairs poorly at these longer forecast times. DMCC and GFS MOS have comparable performance at 6 h, but the DMCC algorithm performs noticeably better at 12 h as evidenced by higher accuracy and skill scores for the lower-ceiling-height class in the two-class test (Fig. 5).

4. Model analysis

As outlined in Bankert et al. (2004), C5.0 produces decision trees by selecting those attributes (RUC variables in this study) that best partition the data. Attributes selected at each tree node are selected using an information gain principle (Shannon 1948). The attributes in the top branches of the decision tree are assumed to be the most significant contributors to the classification. Examining the top three branches of each decision tree for the 1-, 3-, 6-, and 12-h cloud-ceiling-height classification produced some interesting findings. The top contributing RUC variables (in order) for each of the forecast times are listed in Table 7.

Precipitable water ratio (PWR) was computed for the lowest 150 mb and is the ratio of precipitable water to precipitable water assuming vapor saturation. While the PWR variable is prevalent at all forecast times, the average relative humidity (in the lowest 150 mb) becomes less significant at longer forecast times and the u wind component in the lowest 30 mb becomes more significant. Because the 1-h RUC variables were used to create all decision trees, the 1-h moisture variables appear to play a more significant role in classifying cloud-ceiling heights at 1 and 3 h, and the 1-h momentum (e.g., the u wind component) variables contribute more at the longer forecast times.

Other significant contributors in the top three branches included the dewpoint temperature (at lowest RUC level) and the lifting condensation level (LCL). Both of these variables played a larger role at shorter forecast times (along with yes or no cloud, as noted in Table 7). Average turbulent kinetic energy (TKE) and planetary boundary layer (PBL) depth were strong contributors at the longer forecast times.

While a detailed analysis of the physical reasoning of how these variables—valid at 1-h forecast time—relate to current (defined here as 1 h) and future (3, 6, and 12 h) cloud-ceiling heights is beyond the scope of this research, a general analysis is provided here. At 1 h, one would expect the current state of the atmosphere, in terms of temperature and moisture, to play a prominent role. This is indeed the case when considering the strongest contributing variables selected: average RH, PWR, yes–no cloud, dewpoint temperature (lowest level), and LCL height. At 3 h, the current moisture state of the atmosphere is still important, but the current wind speed and direction (u wind component in the lowest 30 mb) also appears to be a factor. In the coastal environ-

TABLE 6. The 12-h cloud-ceiling forecast classification accuracies (overall and each class %) for DMCC, persistence, and GFS MOS six-class classifiers. The boundary between VFR3 and VFR4 is 5000 ft for DMCC and persistence and 6500 ft for GFS MOS.

	Overall	VFR1/VFR2	VFR3	VFR4	MVFR	IFR	LIFR
DMCC	71.3	90.6	40.2	35.1	51.4	31.2	54.7
Persistence	52.8	73.7	17.9	21.2	27.1	20.6	20.8
GFS MOS	64.3	83.6	27.1	46.0	42.0	20.5	29.2

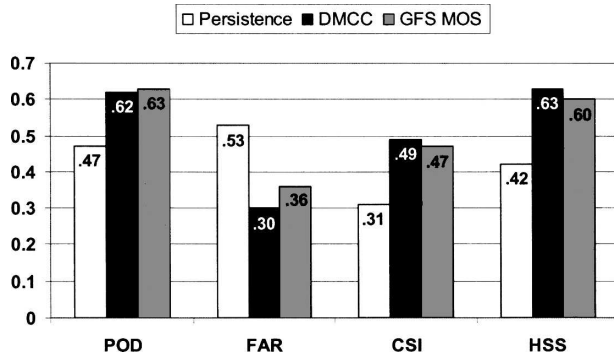


FIG. 4. Two-class (<1000 ft; ≥ 1000 ft) 6-h cloud-ceiling forecast accuracy and skill levels for persistence, DMCC, and GFS MOS classifiers.

ment in which KJFK exists, the magnitude of the temperature and moisture advection are keys in determining future cloud-ceiling conditions, possibly explaining the relationship of future cloud-ceiling height with the current wind data. The 1-h *u* wind (lowest 30 mb) is an even stronger contributor in the 6- and 12-h forecasts and the vertically averaged TKE appears as a significant variable. The dynamic elements (wind, TKE) at 1 h appear to be stronger factors in determining the 6- and 12-h cloud-ceiling-height forecasts than the current moisture state. If the state of the environment is changing from 1 to 6 or 12 h, having current data that relates to change should play a more prominent role.

5. Discussion

Application of a data-mining technique to 1-h RUC model data and METAR observations at a single station (KJFK) has resulted in RUC-based cloud-ceiling-height classification algorithms producing estimates at

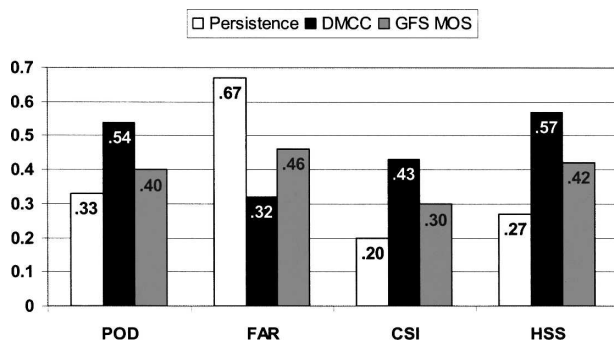


FIG. 5. Two-class (<1000 ft; ≥ 1000 ft) 12-h cloud-ceiling forecast accuracy and skill levels for persistence, DMCC, and GFS MOS classifiers.

TABLE 7. Top contributing RUC variables in the decision tree algorithms for each forecast time. The variables are listed in order (left to right).

1 h	Avg RH	PWR	Yes-no cloud
3 h	PWR	Avg RH	<i>u</i> wind (layer)
6 h	<i>u</i> wind (layer)	PWR	Avg RH
12 h	<i>u</i> wind (layer)	PWR	<i>u</i> wind (level) Longwave radiation

1-, 3-, 6-, and 12-h forecast times. The DMCC algorithms were able to maintain six-class (no ceiling and >12 000 ft ceiling combined into one class) overall accuracy from 1 to 12 h (Fig. 6) indicating the ability to find useful relationships between RUC output and cloud-ceiling observations out to 12 h. Persistence and GFS MOS both have a steeper decrease in accuracy at the longer forecast times. Similar results can be seen in the two-class (<1000 ft and ≥ 1000 ft) HSS statistics (Fig. 7).

Improvements to the DMCC model could be accomplished by taking into account day-night differences and warm-cold seasonal variations. Additionally, a more complex algorithm, developed through data mining, that also includes station observations, satellite data, and climatological parameters should produce a more robust cloud-ceiling forecast.

Future work includes generating algorithms for multiple locations within the contiguous United States and applying those algorithms in postprocessing RUC data for appropriate grid points. Some locations will prove to be more challenging than others. Mountainous terrain, microclimates, and other local factors not adequately represented in the RUC model can contribute to the development of cloud-ceiling algorithms with unsatisfactory results. Also, the sensitivity of developed algorithms to RUC model upgrades needs to be considered.

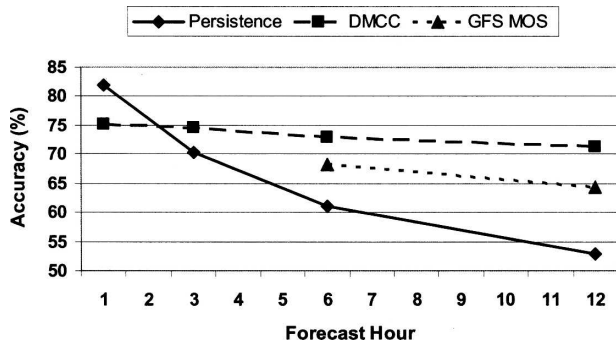


FIG. 6. Six-class overall accuracy for cloud-ceiling forecasts produced by persistence, DMCC algorithm, and GFS MOS over all forecast times.

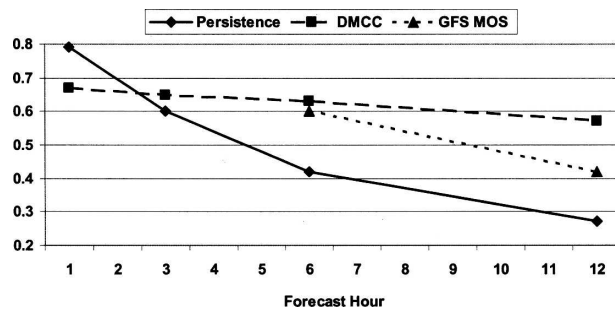


FIG. 7. Two-class (<1000 ft; \geq 1000 ft) HSS for cloud-ceiling forecasts produced by persistence, DMCC algorithm, and GFS MOS over all forecast times.

Acknowledgments. The support of the sponsor, the FAA through the AWRP, is gratefully acknowledged. The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA. Without the assistance, advice, and support of Paul Herzegh, Gerry Weiner, and Jim Cowie at NCAR RAL and John Brown and Stan Benjamin at the Global System Division of ESRL, this work would not have been possible. Their assistance is greatly appreciated.

REFERENCES

- Abdel-Aal, R. E., and M. A. Elhadidy, 1995: Modeling and forecasting the daily maximum temperature using abductive machine learning. *Wea. Forecasting*, **10**, 310–325.
- Bankert, R. L., M. Hadjimichael, A. P. Kuciauskas, W. T. Thompson, and K. Richardson, 2004: Remote cloud ceiling assessment using data-mining methods. *J. Appl. Meteor.*, **43**, 1929–1946.
- Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004a: Mesoscale weather prediction with the RUC hybrid isentropic–terrain-following coordinate model. *Mon. Wea. Rev.*, **132**, 473–494.
- , and Coauthors, 2004b: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.
- Bocchieri, J. R., and H. R. Glahn, 1972: Use of model output statistics for predicting ceiling height. *Mon. Wea. Rev.*, **100**, 869–879.
- , R. L. Crisci, H. R. Glahn, F. Lewis, and F. T. Globokar, 1974: Recent developments in automated prediction of ceiling and visibility. *J. Appl. Meteor.*, **13**, 277–288.
- Bremnes, J. B., 2004: Probabilistic forecasts of precipitation in terms of quantiles using NWP model output. *Mon. Wea. Rev.*, **132**, 338–347.
- Brown, J. M., T. G. Smirnova, S. G. Benjamin, R. Rasmussen, G. Thompson, and K. Manning, 2000: Use of a mixed-phase microphysics scheme in the operational NCEP Rapid Update Cycle. Preprints, *Ninth Conf. on Aviation, Range, and Aerospace Meteorology*, Orlando, FL, Amer. Meteor. Soc., 100–101.
- Doswell, C. A., III, R. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576–585.
- Fine, S. S., and A. B. Fraser, 1990: A simple technique for multiple-parameter pattern recognition with an example of locating fronts in model output. *J. Atmos. Oceanic Technol.*, **7**, 896–908.
- Firth, L., M. L. Hazelton, and E. P. Campbell, 2005: Predicting the onset of Australian winter rainfall by nonlinear classification. *J. Climate*, **18**, 772–781.
- Forsythe, J. M., T. H. Vonder Haar, and D. L. Reinke, 2000: Cloud-base height estimates using a combination of meteorological satellite imagery and surface reports. *J. Appl. Meteor.*, **39**, 2336–2347.
- Frehlich, R., and R. Sharman, 2004: Estimates of turbulence from numerical weather prediction model output with applications to turbulence diagnosis and data assimilation. *Mon. Wea. Rev.*, **132**, 2308–2324.
- Glahn, H. R., and D. A. Lowery, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.
- Godfrey, R. A., 1982: An application of model output statistics to the development of a local wind regime forecast procedure. *J. Appl. Meteor.*, **21**, 1786–1791.
- Grell, G. A., and D. Devenyi, 2001: Parameterized convection with ensemble closure/feedback assumptions. Preprints, *Ninth Conf. on Mesoscale Processes*, Fort Lauderdale, FL, Amer. Meteor. Soc., 12–16.
- Hansen, B. K., and D. Riordan, 2003: Fuzzy case-based prediction of cloud ceiling and visibility. Preprints, *Third Conf. on Artificial Intelligence Applications to the Environmental Sciences*, Long Beach, CA, Amer. Meteor. Soc., CD-ROM, 4.4.
- Herzegh, P., and Coauthors, 2006: Development of FAA national ceiling and visibility products: Challenges, strategies and progress. Preprints, *12th Conf. on Aviation, Range, and Aerospace Meteorology*, Atlanta, GA, Amer. Meteor. Soc., CD-ROM, P1.17.
- Hilliker, J. L., and J. M. Fritsch, 1999: An observation-based statistical system for warm-season hourly probabilistic forecasts of low ceiling at the San Francisco International Airport. *J. Appl. Meteor.*, **38**, 1692–1705.
- Kassianov, E., C. N. Long, and J. Christy, 2005: Cloud-base height estimation from paired ground-based hemispherical observations. *J. Appl. Meteor.*, **44**, 1221–1233.
- Koziara, M. C., R. J. Renard, and W. J. Thompson, 1983: Estimating marine fog probability using a model output statistics scheme. *Mon. Wea. Rev.*, **111**, 2333–2340.
- Lee, Y., G. Wahba, and S. A. Ackerman, 2004: Cloud classification of satellite radiance data by multicategory support vector machines. *J. Atmos. Oceanic Technol.*, **21**, 159–169.
- Leyton, S. M., and J. M. Fritsch, 2003: Short-term probabilistic forecasts of ceiling and visibility utilizing high-density surface weather observations. *Wea. Forecasting*, **18**, 891–902.
- , and —, 2004: The impact of high-frequency surface weather observations on short-term probabilistic forecasts of ceiling and visibility. *J. Appl. Meteor.*, **43**, 145–156.
- Maini, P., A. Kumar, L. S. Rathore, and S. V. Singh, 2003: Forecasting maximum and minimum temperatures by statistical interpretation of numerical weather prediction model output. *Wea. Forecasting*, **18**, 938–952.
- Mao, Q., R. T. McNider, S. F. Mueller, and H.-M. H. Juang, 1999: An optimal model output calibration algorithm suitable for objective temperature forecasting. *Wea. Forecasting*, **14**, 190–202.
- Marzban, C., 2003: Neural networks for postprocessing model output: ARPS. *Mon. Wea. Rev.*, **131**, 1103–1111.
- Murphy, A. H., 1992: Climatology, persistence, and their linear

- combination as standards of reference in skill scores. *Wea. Forecasting*, **7**, 692–698.
- Norquist, D. C., 1999: Cloud predictions diagnosed from meso-scale weather model forecasts. *Mon. Wea. Rev.*, **127**, 2465–2483.
- NWS, 1998: ASOS user's guide. National Weather Service, 61 pp. [Available from ASOS Program Office, NWS, 1325 East-West Highway, Silver Spring, MD 20910.]
- Quinlan, J. R., 1993: *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 302 pp.
- Shannon, C. E., 1948: A mathematical theory of communication. *Bell Syst. Tech. J.*, **27**, 379–423.
- Stoelinga, M. T., and T. T. Warner, 1999: Nonhydrostatic, meso-beta scale model simulations of cloud ceiling and visibility for an East Coast winter precipitation event. *J. Appl. Meteor.*, **38**, 385–404.
- Tadesse, T., D. A. Wilhite, M. J. Hayes, S. K. Harms, and S. Goddard, 2005: Discovering associations between climatic and oceanic parameters to monitor drought in Nebraska using data-mining techniques. *J. Climate*, **18**, 1541–1550.
- Tag, P. M., and J. E. Peak, 1996: Machine learning of maritime fog forecast rules. *J. Appl. Meteor.*, **35**, 714–724.
- Tapp, R. G., F. Woodcock, and G. A. Mills, 1986: The application of model output statistics to precipitation prediction in Australia. *Mon. Wea. Rev.*, **114**, 50–61.
- Vislocky, R. L., and J. M. Fritsch, 1997: An automated, observation-based system for short-term prediction of ceiling and visibility. *Wea. Forecasting*, **12**, 31–43.
- Weiss, M., and J. Ghiradelli, 2005: A summary of ceiling height and total sky cover short-term statistical forecasts in the Localized Aviation MOS Program (LAMP). Preprints, *21st Conf. on Weather Analysis and Forecasting and 17th Conf. on Numerical Weather Prediction*, Washington, DC, Amer. Meteor. Soc., CD-ROM, 13B.6.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 627 pp.
- Woodcock, F., 1984: Australian experimental model output statistics forecasts of daily maximum and minimum temperature. *Mon. Wea. Rev.*, **112**, 2112–2121.