Reprinted from Preprint Volume, Sixth Conference on Aviation Weather Systems, January 15-20, 1995, Dallas, AMS

6.4

THE ACCURACY OF CEILING AND VISIBILITY FORECASTS PRODUCED BY THE NATIONAL WEATHER SERVICE

J. Paul Dallavalle and Valery J. Dagostaro

Techniques Development Laboratory
Office of Systems Development
National Weather Service, NOAA

1. INTRODUCTION

In the AFOS-Era Verification (AEV) system, the National Weather Service (NWS) verifies ceiling and visibility forecasts produced by the local NWS forecaster and issued in the aviation terminal forecasts (FT's). A sample of the ceiling and visibility forecasts is extracted from the body of the FT to represent prevailing conditions at specific forecast times. Forecasts from two of the three scheduled FT issuances are verified.

In this paper, we present a verification of the local forecasts of ceiling and visibility for 90 stations in the contiguous United States. The latest hourly surface observation available before the FT issuance time (persistence) is also verified. Forecasts made during the cool (October - March) and warm (April - September) seasons since October 1983 are verified. Verification scores are computed by categorizing the forecasts and verifying observations into four by four contingency tables based on the definitions of LIFR (low instrument flight rules), IFR (instrument flight rules), MVFR (marginal visual flight rules), and VFR (visual flight rules) meteorological conditions (Table 1). We also compare local forecasts for the 1993-94 cool season to Model Output Statistics (MOS) guidance produced from the Nested Grid Model (NGM).

2. DESCRIPTION OF FT'S

The FT's describe expected weather conditions affecting operations at an airport. Conditions predicted in the FT's include clouds, visibility, weather, obstructions to vision, winds, and nonconvective low-level wind shear. Scheduled FT's are issued three times daily based on the local time zone of the issuing office. These scheduled FT's consist of specific weather element forecasts for a 24-h period. Amended FT's are issued if changing weather conditions require modifications to the FT.

The FT is comprised of one or more forecast groups describing expected conditions during specific periods of time. The first group contains conditions predicted at the start of the FT valid period ("valid time"), namely, 15 minutes after the scheduled issuance time. Subsequent groups in the FT's are required to indicate significant weather changes. Each FT group is comprised of the forecast body and, possibly, forecast remarks. The forecast body indicates prevailing conditions during the period, while remarks indicate conditions that differ from those predicted in the body. Because remarks are designed to indicate operationally significant, but transitory, features, the remarks can be more important than the body of the forecast. NWS Operations Manual Chapter D-21 (National Weather Service 1994) describes the FT's.

3. AFOS-ERA VERIFICATION (AEV) SYSTEM

Since October 1983, the NWS has verified public and aviation weather forecasts made by local

Table 1. Definitions of critical aviation categories for ceiling and visibility conditions.

Category	Ceiling (ft)	Visibility (mi)
LIFR	< 500	< 1
IFR	500 - < 1000	1 to < 3
MVFR	1000 - 3000	3 to 5
VFR	> 3000 or none	> 5

Corresponding author address: 1325 East-West Highway, Silver Spring, MD 20910

NWS forecasters through use of an automated data collection and transmission system. The AEV system collects subjective forecasts of weather elements such as maximum/minimum temperature, probability of precipitation, wind, ceiling height, and visibility; obtains forecasts of the same elements from the operational MOS guidance message; and collates appropriate verifying observations. Messages containing the AEV data are transmitted daily by the local NWS Forecast Offices to the National Meteorological Center for processing and archival by the Techniques Development Laboratory. Data for approximately 100 of the over 500 FT stations in the U.S. are archived. Dagostaro et al. (1989) describe the NWS national verification system.

For ceiling and visibility, the AEV data consist of a sample of local NWS forecasts extracted from two of the three daily scheduled FT's. Specifically, forecasts valid 3, 6, 9, and 15 hours after the beginning of the FT forecast period are collected and verified (Table 2). Only the forecasts of prevailing conditions are considered. Thus, remarks which may indicate significant, anomalous conditions are not verified though these remarks might indicate conditions of ceiling height or visibility valid at the projection of interest. For comparison, the skill of the local FT's is checked against that of persistence. Note that prior to December 1986, the verified FT issuance times were fixed at 0930 and 2130 UTC, and the persistence observations used for comparison were 0900 and 2100 UTC, respectively, for all stations, regardless of time zone.

4. EARLIER STUDIES

Previous studies have shown that local forecasts of ceiling height and visibility for LIFR/IFR conditions seldom improve upon persistence at the very short-range projections (up to 6 hours). Zurndorfer et al. (1979) found that persistence had

Table 2. Issuance time of scheduled FT's verified in the AEV program. LT denotes local time. DT (ST) denotes issuance time during Daylight Savings (Standard) Time. The persistence observations used in comparative verification are indicated, as are the valid times of the verified FT. These times correspond to 3-, 6-, 9-, and 15-h projections after the start of the FT period.

Time Zone	Issuance Time	Persistence Obs. (Z)	FT Valid Time (Z)
Eastern	0345LT - 0745Z DT	0700	1100, 1400, 1700, 2300
	0845Z ST	0800	1200, 1500, 1800, 0000
	1245LT - 1645Z DT	1600	2000, 2300, 0200, 0800
	1745Z ST	1700	2100, 0000, 0300, 0900
Central	0245LT - 0745Z DT	0700	1100, 1400, 1700, 2300
	0845Z ST	0800	1200, 1500, 1800, 0000
	1145LT - 1645Z DT	1600	2000, 2300, 0200, 0800
	1745Z ST	1700	2100, 0000, 0300, 0900
Mountain	0245LT - 0845Z DT	0800	1200, 1500, 1800, 0000
	0945Z ST	0900	1300, 1600, 1900, 0100
	1145LT - 1745Z DT	1700	2100, 0000, 0300, 0900
	1845Z ST	1800	2200, 0100, 0400, 1000
Pacific	0245LT - 0945Z DT	0900	1300, 1600, 1900, 0100
	1045Z ST	1000	1400, 1700, 2000, 0200
	1145LT - 1845Z DT	1800	2200, 0100, 0400, 1000
	1945Z ST	1900	2300, 0200, 0500, 1100

greater skill scores (separating LIFR conditions from the other categories) for the 3-h prediction of ceiling height than the local forecasts. German and Hicks (1981) showed that from 1968 to 1979 USAF forecasters had skill relative to persistence for combined ceiling and visibility forecasts for the 3- and 6-h projections. The increase in improvement relative to persistence was quite dramatic over the 12-year period for the 3-h projection, but was less pronounced for the 6-h projection. German and Hicks concluded that most of the improvement was in the ability to forecast good weather, and that the ability to forecast the lower categories of ceiling height and visibility was low and had not changed. An internal NWS report (National Weather Service 1991) noted that local short-range forecasts of LIFR/IFR ceiling were no better than persistence. The same report noted that, after the change in the scheduled and verified FT release time in December 1986, the forecast skill of ceiling height in the LIFR/IFR categories deteriorated relative to persistence. Analysis showed that the local forecasters tended to end LIFR/IFR conditions too early. Lerner and Polger (1989) described a prototype FT verification system that considered every hour of the FT as well as the remarks. In one case of verifiying visibility forecasts, the bias scores were better (closer to 1.0) for LIFR/IFR conditions when sampling was done every hour. Goldsmith (1993) verified 2 months of enhanced terminal forecasts made by NWS forecasters in Denver. Remarks and prevailing conditions were verified by using all available surface observations, including specials. Goldsmith found that during the first 3 hours of the FT, persistence of the latest available observation was a more skillful forecast of IFR conditions than the local FT. He also noted that the skill of the FT remarks was surprisingly disappointing and that his verification of prevailing conditions corroborated the level of skill documented in the NWS AEV program.

5. VERIFICATION OF CEILING/VISIBILITY

In the verification scores that follow, we verified the FT ceiling height and visibility taken from the AEV data. Scores shown were based on forecasts for 90 stations in the contiguous U.S. that have participated in the AEV program since its inception in 1983. Thus, for most seasons and projections, approximately 15,000 cases were available. Because of the change in the FT release time and problems with the AEV program, data for the 1986-87 cool season were available only from October 1 through December 16 while data for the 1987 warm season were available only from July 1 through September 30. The FT forecasts were categorized into one of four groups (Table 1) and were verified by computing scores from four by four contingency tables. As verification measures, we used the Heidke skill score (HSS) and the critical success index (CSI) for LIFR/IFR conditions. Comparisons were made with the skill of the persistence observation defined previously.

Fig. 1 shows the HSS for combined ceiling/visibility forecasts made from the scheduled noon release time during the cool season. In the Eastern time zone, for example, this release time coincides with 1245 LT (Table 2). In Fig. 1, scores for forecasts valid 3 and 9 hours after the start of the FT period are shown. Fig. 2 shows the CSI for the same forecasts. In the cool seasons from 1983-84 through 1993-94, no increase in skill of the local forecasts is evident. In fact, the skill of the 3-h local forecasts appears to decline after the change of the FT issuance time in the 1986-87 cool season. This change in skill indicates that forecasting LIFR/IFR conditions from a noon FT release time is more difficult than from a late afternoon release time. At noon, ceiling/visibility conditions are often in a transition state, and predicting the exact time that changes from LIFR/IFR to MVFR/VFR will occur is difficult. At the 3-h projection, persistence is more skillful than the local forecasts in terms of both HSS and CSI. In contrast, for the 9-h projection, local forecasts are more skillful in terms of HSS, and, to a lesser extent, in terms of CSI. LIFR/IFR conditions in the cool season occur about 9% and 10% of the time for this 3- and 9-h projection, respectively.

Similar time series (not shown) indicate the skill of the warm season forecasts is less than that of the cool season forecasts. In addition, the skill of the warm season local forecasts appears to decline after the FT release time is changed, persistence has greater skill at 3 hours than the locals in terms of both HSS and CSI, and the skill of the 9-h

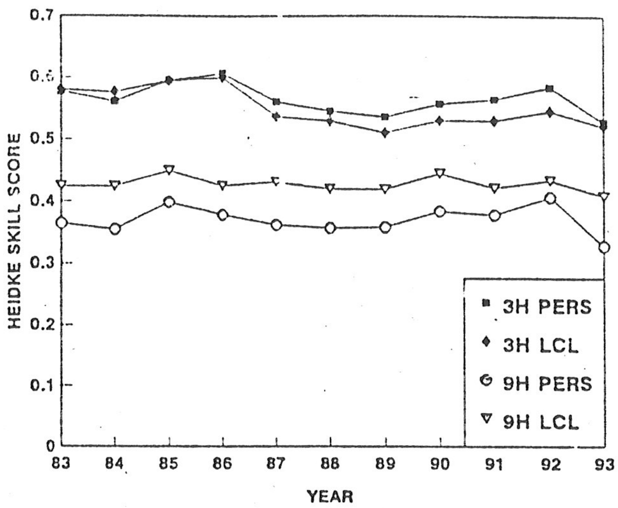

Figure 1. Heidke skill scores of persistence (PERS) and local (LCL) forecasts of combined ceiling/visibility conditions made during the cool season (late afternoon release until 1986-87, noon release afterwards).

local forecast exceeds that of persistence in terms of the HSS. Differences in the CSI of the local forecasts and persistence vary at the 9-h projection. Note that LIFR/IFR conditions in the warm season occur about 3% and 4% of the time for this 3- and 9-h projection, respectively.

Though we have not shown the time series here, a consistent picture emerges in the individual ceiling/visibility verifications. In short:

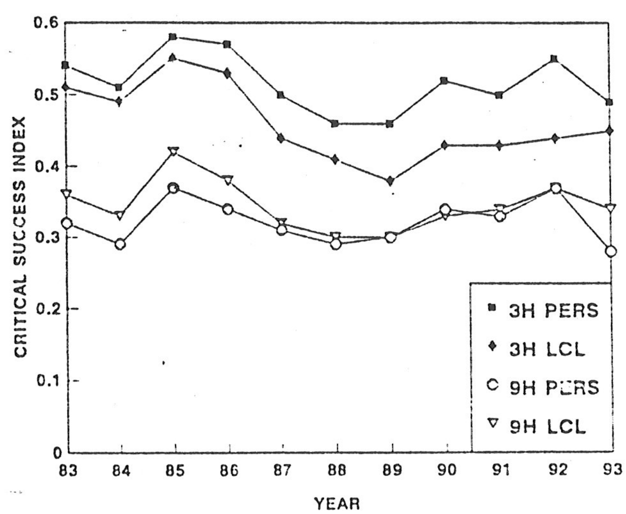

Figure 2. Same as Fig.1, except for CSI.

-- HSS and CSI for the local ceiling forecasts exceeded those for comparable visibility forecasts;
-- the cool season verification scores were better than those for the warm season;
-- no increase in skill of the local forecasts was evident over the years; in fact, the CSI of the local forecasts declined for the daytime release after the December 1986 change in the release schedule; this decline reflected the difficulty of predicting transitions in ceiling and visibility conditions that occurred in the afternoon;
-- for the late night release, the CSI declined during the cool season for both the ceiling and visibility forecasts after December 1986;
-- for all seasons, release times, and elements, the skill of persistence exceeded that of the locals for the 3-h forecasts;
-- for the 6-h forecasts, the HSS for the locals was greater than or about equal to that of persistence; however, the CSI of persistence usually exceeded that of the locals;
-- for the 9-h forecasts, the HSS for the local forecasts exceeded that of persistence; differences in the CSI between the local forecasts and persistence for the noon release varied substantially; for the late night forecast, persistence generally had higher CSI's;
-- for the 15-h forecasts, the skill of the local forecasts exceeded persistence.

Data for the cool seasons of 1987-88 through 1993-94 (approximately 108,000 cases for each projection) were stratified into four subsets according to whether the persistence observation for the combined ceiling/visibility condition was LIFR, IFR, MVFR, or VFR. As before, the CSI was computed for combined ceiling/visibility conditions of LIFR and IFR. Thus, for example, when the persistence observation was LIFR, a verifying observation of LIFR or IFR was considered a hit. For persistence observations of MVFR or VFR, no hits were possible and so the CSI equalled 0. Figs. 3 and 4 show the CSI of the combined ceiling/visibility forecasts for the 3- and 9-h projections, respectively, from the noon release when the persistence observation was LIFR, IFR, MVFR, or VFR. The four subsets amounted to approximately 3%, 7%, 16%, and 74%, respectively, of the total number of cases. Note that for the 3-h projection,

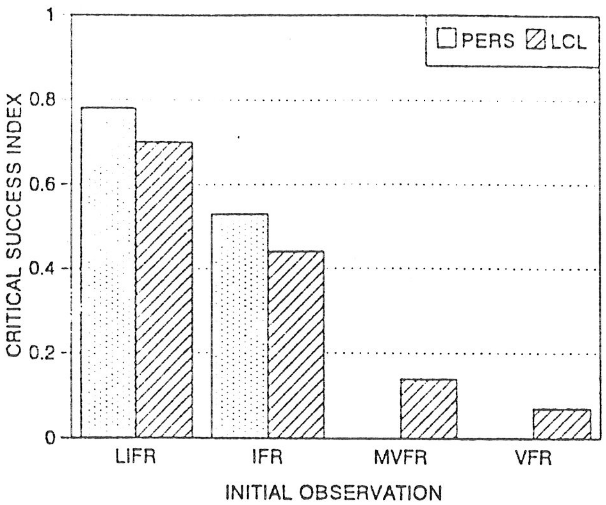

Figure 3. CSI of combined 3-h ceiling/visibility forecasts (noon release) for the cool season.

the CSI of persistence exceeded that of the local forecasts when initial conditions were LIFR or IFR. Apparently, the local forecasts predicted a change to prevailing conditions of MVFR/VFR too quickly. Moreover, when initial conditions were MVFR or VFR and changed to LIFR/IFR conditions at the verifying time, the CSI of the local forecasts was very low. Note that this scenario occurred for MVFR or VFR conditions only 10% and 1% of the time, respectively. For the 9-h projection, the CSI of the local forecasts and persistence were about equal for initial LIFR or IFR conditions. For initial MVFR or VFR conditions, the locals had higher CSI values than at 3 hours. Note that MVFR or VFR conditions reverted to LIFR/IFR conditions 13% and 3% of the time, respectively.

6. MOS CEILING/VISIBILITY GUIDANCE

NGM-based MOS guidance for ceiling and visibility has been available to forecasters since January 1993 and July 1993, respectively. The MOS guidance is generated for the probability of certain ceiling and visibility categories. These probability forecasts are then converted to categorical forecast values before issuance to the forecasters. Guidance valid every 3 hours from 6 to 36 hours after 0000 or 1200 UTC is available in alphanumeric form (Dallavalle et al. 1992). Figs. 5 and 6 present verifications comparing persistence, local, and MOS forecasts for the 1993-94 cool season. MOS forecasts from the 1200 UTC run of the NGM were compared to the local forecasts issued at 1745 UTC. Only stations in the Eastern and Central time zones were verified since the forecasts made by the local forecasters in these areas were valid at the same time as the MOS guidance. Note that the forecast projections shown in Figs. 5 and 6 refer to hours after the start of the FT. In fact, the MOS forecasts were produced from 1200 UTC NGM data and were valid at the same time as the FT's issued later. Some of the MOS

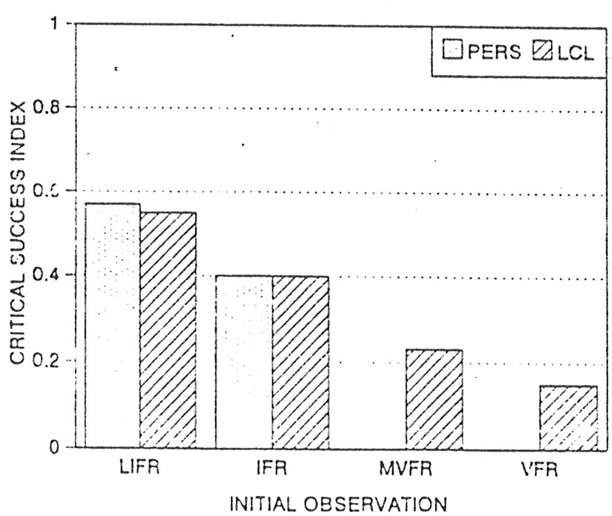

Figure 4. Same as Fig. 3, except for the 9-h projection.

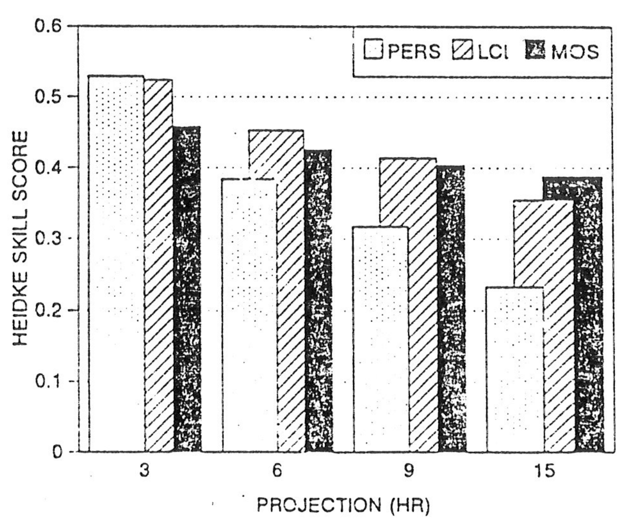

Figure 5. Heidke skill scores of persistence (PERS), locals (LCL), and NGM MOS guidance for the 1993-94 cool season (noon release).

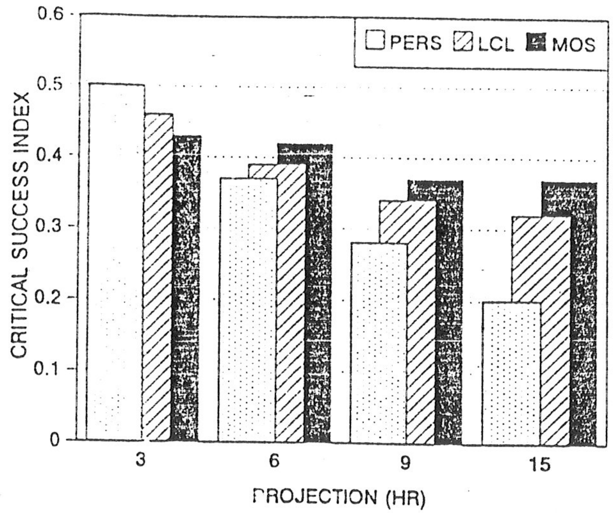

Figure 6. Same as Fig. 5, except for the CSI.

forecasts may have also used 1400 UTC surface observations as predictors. For the 1993-94 cool season, the local forecasts were more skillful than persistence at all but the 3-h projection. Note, too, that the MOS guidance was more skillful in forecasting the threat event (LIFR/IFR) than persistence or the locals at 6-, 9-, and 15-h projections.

7. CONCLUSIONS

Verifications of the local ceiling and visibility FT's since 1983 have been compiled by using the AEV dataset. Although no increase in skill during that time was evident, our assessment was limited by the inability to verify the FT remarks. Generally, persistence forecasts appeared to have higher skill than the local forecasts for the 3-h projection. For the 6-h forecasts, the local forecasts had higher Heidke skill scores, but lower CSI's than persistence. By the 15-h projection, the local forecasts were superior to persistence. The NGM-based MOS guidance for ceiling/visibility combined appeared to offer some help to the forecaster in predicting LIFR/IFR conditions.

8. REFERENCES

Dagostaro, V. J., C. L. Alex, G. M. Carter, J. P. Dallavalle, and P. D. Polger, 1989: Evolution of the NWS national verification system: Past, present, and future. Preprints Eleventh Conference on Probability and Statistics, Monterey, Amer. Meteor. Soc., J41-J46.

Dallavalle, J. P., S. A. Gilbert, and J. S. Jensenius, Jr., 1992: NGM-based MOS guidance - The FOUS14/FWC message. NWS Technical Procedures Bulletin No. 408, National Oceanic and Atmospheric Administration, NOAA, U.S. Department of Commerce, 15 pp.

German, K. E., and P. Hicks, Jr., 1981: Air Weather Service ceiling and visibility verification. Bull. Amer. Meteor. Soc., 62, 785-790.

Goldsmith, B. S., 1993: Objective verification of an enhanced terminal forecast experiment at Denver, Colorado. Fifth International Conference on Aviation Weather Systems, Vienna, Amer. Meteor. Soc., 249-253.

Lerner, W., and P. D. Polger, 1989: An evaluation of aviation terminal forecasts based on an enhanced verification system. Preprints Eleventh Conference on Probability and Statistics, Monterey, Amer. Meteor. Soc., 133-137.

National Weather Service, 1991: From the forecast desk. Verification Viewpoint No. 1, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 10 pp.

____, 1994: Aviation Terminal Forecasts. NWS Operations Manual Chapter D-21, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 45 pp.

Zurndorfer, E. A., J. R. Bocchieri, G. M. Carter, J. P. Dallavalle, D. B. Gilhousen, K. F. Hebenstreit, and D. J. Vercelli, 1979: Trends in comparative verification scores for guidance and local aviation/public weather forecasts. Mon. Wea. Rev., 107, 799-811.