Very Short Range Weather Forecasting Using Automated Observations

Robert G. Miller *National Weather Service*

Pilot Joe Lang was using his medical-evacuation helicopter to take an accident victim from the beltway to a nearby trauma center, and after 10 miles of his 20-mile trip, the weather changed at his destination, leaving him no visibility to make a landing. Precious time was lost finding an alternate landing site at the last moment because there had been no short range forecast. Helicopter pilots want and need help in preventing this loss of time and the life-threatening disaster of a crash.

This essay describes a statistical effort to produce very short range weather forecasts to help helicopter pilots and others. The work was conducted in conjunction with the automated weather observing system (AWOS) currently under test. An AWOS observation is made at an unstaffed airfield by taking minute-by-minute readings from meteorological instruments (sensors that measure temperature, wind, and so forth). The most recent 30 minutes of these readings are sent electronically to a computer programmed to emulate human observations in creating a weather observation. There are three alternative ways to make short-range forecasts of, say, 10–120 minutes:

1. *Persistence forecasts.* These are forecasts that the current conditions will prevail into the short-term future. This method provides good results and is not easy to beat. We know that weather changes infrequently over short time periods, so persistence forecasting is bound to be rather accurate.
2. *Human judgment.* Human forecasts are an important alternative because all current and past weather information is at the disposal of the forecaster. This information includes observations from weather radar and satellites, spatial and temporal weather conditions at the surface and aloft, large-scale prognoses, and, of course, the local AWOS observation. Combined with the forecaster's experience and training, this information looks like the best source for 10–120-minute forecasts. Experience has shown however, that such forecasts rarely show any gain over persistence forecasting for less than 120 minutes.
3. *Statistical prediction.* The statistical approach employed here, that of basing forecasts on empirical data, represents another alternative. It has the advantages of using only the most recent locally available AWOS observation as input, of being easily automated, and of possessing a rapid response time. The approach is founded on almost a million past observations that generate estimated relationships.

Because short-term forecasts based on human judgment that show improvement over persistence are hard to provide, can statistical prediction help?

CREATING PREDICTION EQUATIONS

For purposes of illustration, we describe a statistical approach for making a 10-minute prediction of visibility that is directly compared with persistence. The approach is called *regression analysis* (or two-group discriminant analysis; see Tatsuoka, 1971). (See also the essay by Howells on the discriminant function.) It is a popular way of estimating the value of an unknown variable that is given in terms of weighted values of known variables. (In particular, the weights are determined in such a way as to minimize the sum of squares of the differences between the estimates and the variable's true values.) We choose these weights by examining situations where we do know the values of the usually unknown variable.

Visibility is measured in miles and fractions of miles by a visibility sensor. The AWOS visibility is derived from the 10 most recent visibility sensor readings. We could attempt to predict this derived value but instead it turns out to be more convenient to report in which of the following six categories of 10-minute visibility the observation and the prediction lie:

- 0 up to but not including 1/2 mile
- 1/2 up to but not including 1 mile
- 1 up to but not including 3 miles
- 3 up to but not including 5 miles
- 5 up to but not including 7 miles
- 7 miles or more

In a given forecast situation, an AWOS observation is represented by a series of 1s and 0s. For example, the weather element wind speed has five velocity categories, only one of which can occur in the observation. That category gets a 1 and the other four get 0s. Similarly, wind direction has eight direction categories of which only one can occur in the observation. Altogether there are 26 weather elements represented by 166 categories. A given observation of the 166 categories has exactly 26 1s and 140 0s. This form of representation has two important purposes. First, although some weather elements are ordered, they are not on a numerical scale (visibility is an example in which there are no numeric values for the category "unlimited" and "7 miles or more"). Second, the mathematical processing of the categorical data is done very efficiently by logical computer operations, replacing slower arithmetical operations. (For example, it takes a microcomputer 20 times longer to obtain the multiplication of two real numbers than for a logical operation between two integers where each is representing many 0/1 numbers. All told, the comparison works out to be about two orders of magnitude in computation time and about one order of magnitude in storage space in favor of the 0/1 scheme.)

The statistical analysis is a regression procedure that selects 30 predictors from the 166 variables (see Draper and Smith, 1981, and Miller, 1962). A few examples of potential variables whose categories are used as predictors are: lowest height observed by cloud sensor in one minute, pressure, wind direction, and precipitation amount. On the basis of success in estimating the outcome of 818,953 previous observations, the statistical procedure selects the 30 predictors, or categories, and also computes weights for them that would use their outcome to good effect. The overall method produces a separate estimate or score for each of the categories of 10-minute visibility listed above. Ideally the outcome would be a 1 for the category that is going to happen and a 0 for all others, but actually the numbers vary from somewhat below 0 to somewhat above 1. Each category has its own equation.

The first regression estimate gives us for 10-minute visibility the number associated with its first category listed above—0 up to but not including 1/2 mile:

Regression estimate =

$$
\begin{aligned}
0.047 \; & + (-0.001) \times (\text{visibility sensor 7 miles or more}) \\
& + 0.003 \times (\text{visibility sensor 1 mile to less than 3 miles}) \\
& + \\
& \;\cdot \\
& \;\cdot \\
& \;\cdot \\
& + (-0.009) \times (\text{precipitation accumulation } 0.002 - 0.100 \text{ in 1 minute})
\end{aligned}
$$

The terms to the right of the equal sign are algebraically added according to whether the corresponding predictor condition within the parentheses is occurring or not. For example, we start with 0.047, then add −0.001 if the visibility sensor is 7 miles or more; but if the visibility is not 7 miles or more, the term −0.001 is deleted. The remaining 28 coefficients are added or not added in a similar fashion depending on the observed condition of the predictor. The quantity obtained as the regression estimate is an index from which the final visibility forecast is made and will be described later.

Recall that regression analysis computes equation weights that minimize the sum of squares of the error between the forecast value and the condition being forecast. Table 1 sums up how well the two methods perform. It compares the sums of squares of deviations between the observed and forecast values for each category and for each method with the sums of squares using just the mean performance for each category. Thus, persistence forecasting had a sum of squares of deviations for the 0 to 1/2 mile category that was 50.2% as big as that for the sum of squares of deviations from the average. The complement, 100 − 50.2 = 49.8, is called the *predictability.*

The persistence percentage of total predictability was obtained as follows: Persistence probability forecasts were obtained by applying a regression approach similar to the statistical method described above, except that only persistence predictors (the six categories of visibility at time 0 listed above) were selected for inclusion into their (persistence's) equations. This is shown in column 1. The improvement of regression over persistence is given in the last column of Table 1.

The amount of improvement shown by regression over persistence indicates that there is a sizable contribution being made by the predictors in the regression equation both in absolute percentage of total predictability and over and above what the persistence terms in the equation are contributing. These quantities indicate that we can expect better forecasts when the regression equations' index values are applied to the final step in the forecast process. We turn now to the actual process used and how well it works.

COMPARATIVE RESULTS

The goal of this entire effort is to utilize the index values produced by the regression analysis and to predict the category within which visibility will be observed to occur 10 minutes hence. Numerous approaches could be taken. However, the category with the highest index value has been found to produce the largest number of correct forecasts. Unfortunately, this criterion tends to provide little chance of forecasting categories that occur less often and favors those that occur more often. This phenomenon is most evident as the projection time of the forecast is extended. What is desired by both practicing meteorologists and operational users is a balance in the frequency of *each* category—balance in that the forecast frequencies agree closely with the observed frequency for each category of the event. For example, visibilities of 0 to 1/2 mile occur only about 0.4% of the time while visibilities of 7 miles or more occur about 75.5% of the time. Our scheme must predict with roughly these percentages or we will fail to meet the standards of balance that have evolved within the meteorological profession.

Table 1 Percent of total predictability

Predictor Category (Miles)	*Persistence (%)*	*Regression (%)*	*Regression improvement over persistence (%)*
0 < 1/2	49.8	54.8	5.0
1/2 < 1	31.5	40.6	9.1
1 < 3	56.5	64.3	7.8
3 < 5	54.9	62.4	7.5
5 < 7	44.5	53.3	8.8
7^+	83.9	87.6	3.7

A method has been devised by Klein et al. (1959) to satisfy the variability condition preferred by meteorologists and users. We shall not describe it here, but will turn instead to the proof of the pudding.

We can evaluate the stability of our statistical procedure by applying the forecasts to a test sample of data. Tallies for a sample of 369,802 10-minute visibility forecasts are given in Table 2 for the statistical procedure. Table 3 shows the corresponding set of results obtained using persistence.

The results from the test-sample obtained in this effort are extremely gratifying. The number of correct forecasts for the statistical method is obtained from Table 2 by starting with 1,032 and going down the diagonal 1,032 + 1,549 + 14,562 + · · · + 274,368 = 345,240. From Table 3, persistence had 1,061 + 1,382 + 13,926 + · · · + 273,849 = 342,736 correct forecasts. Thus, the statistical method gave 2,504 more correct forecasts than did persistence. Since persistence missed in 369,802 − 342,736 = 27,066 forecasts, regression has succeeded in correcting more than 9% of those misses. Another encouraging fact is that the statistical scheme changed 180 (1,450 − 1,270) of those situations where the visibility was 0 to 1/2 mile at forecast time and only had 29 fewer hits in that category than persistence did. The percentage of correct forecasts of this category is 1,032/1,270 or 81.3%, which is better than the 1,061/1,450 or 73.2% for persistence.

Obviously forecasting high visibility and observing low visibility is undesirable. Table 3 has 13,297 persistence situations below the diagonal (the undesirable forecasts), while regression (Table 2) has 11,389, or 1,908 fewer. The crucial area of three or more below the diagonal in this area of the table is even more impressive, with 153 for persistence and 85 for regression, or 68 fewer very bad forecasts for regression.

The final conclusion is that statistics has succeeded in improving on persistence forecasts of a very difficult meteorological element, visibility, and has done so at the very short range of 10 minutes. The value of correctly forecasting changing conditions, where persistence by its very nature does not make such forecasts, is an accomplishment that cannot be overemphasized.

Table 2 Statistically based forecasts versus observations for 10-minute visibility predictions on a test sample of 369,802 observations

Forecast Category (Miles)				*Verifying Observations (Miles)*			
	0 < 1/2	*1/2 < 1*	*1 < 3*	*3 < 5*	*5 < 7*	*7⁺*	*Totals*
0 < 1/2	1,032	162	52	13	6	5	1,270
1/2 < 1	261	1,549	657	90	31	16	2,604
1 < 3	52	476	14,562	2,884	196	103	18,273
3 < 5	5	34	2,341	27,660	4,312	382	34,734
5 < 7	1	3	133	3,405	26,069	4,264	33,875
7^+	0	0	76	351	4,251	274,368	279,046
Totals	1,351	2,224	17,821	34,403	34,865	279,138	369,802

Table 3 Persistence versus observations for 10-minute visibility predictions on a test sample of 369,802 observations

Forecast Category (Miles)				*Verifying Observations (Miles)*			
	0 < 1/2	*1/2 < 1*	*1 < 3*	*3 < 5*	*5 < 7*	*7⁺*	*Totals*
0 < 1/2	1,061	249	110	17	7	6	1,450
1/2 < 1	197	1,382	590	40	11	5	2,225
1 < 3	80	530	13,926	2,998	207	112	17,853
3 < 5	8	47	2,828	26,673	4,251	564	34,371
5 < 7	2	11	243	4,205	25,845	4,602	34,908
7^+	3	5	124	470	4,544	273,849	278,995
Totals	1,351	2,224	17,821	34,403	34,865	279,138	369,802

PROBLEMS

1. Joe Lang can't land his helicopter if the visibility is, say, less than 1 mile. If he had followed the statistical forecasting approach during the sampling period covered in Tables 2 and 3 how many times would he have had to cancel his flights? How many times would he have had to cancel if he had used persistence? He has to divert to another airport when the forecast is incorrect. Compare the two approaches in light of the actions required, and state a case for choosing one method over the other.
2. Why do you think users of weather forecasts represented by categories prefer them to be issued with the same frequency as they are observed? Is this requirement reasonable? Give the pros and cons as you see them.
3. Does it make you uncomfortable to realize that very short range forecasts might best be issued by a computer program and not from a human being's judgment? Express your opinion.
4. Why do you suppose forecasts are not accurate enough to satisfy our needs? Is it that our demands are too high, our understanding of the atmosphere too inadequate, our data too incomplete, or our analytical methods too limited?

REFERENCES

N. R. Draper and H. Smith. 1981. *Applied Regression Analysis.* New York: Wiley.

W. H. Klein, B. M. Lewis, and I. Enger. 1959. "Objective Prediction of Five-Day Mean Temperatures during Winter." *Journal of Meteorology* 16: 672–681.

R. G. Miller. 1962. *Statistical Prediction by Discriminant Analysis.* Meteorology Monograph No. 25. Boston: American Meteorological Society.

M. M. Tatsuoka. 1971. *Multivariate Analysis: Techniques for Educational and Psychological Research.* New York: Wiley.