STATISTICAL FORECASTS OF SURFACE TEMPERATURE FROM NUMERICAL WEATHER PREDICTIONS

BY

W. KLEIN

NATIONAL WEATHER SERVICE, NOAA

SILVER SPRING, MD.

U.S.A.

ABSTRACT

Several techniques for automatically converting numerical weather prognoses into forecasts of surface temperature in the United States have been developed by the author during the past twenty years. The evolution of objective methods for forecasting 5-day mean temperature, daily mean temperature, maximum/minimum temperature, and temperature at specific times (every 3 hours) is described. The "Perfect Prog" method is compared to "Model Output Statistics" with the aid of sample forecast equations, operational output, and comparative verification figures. The synoptic climatology of 5-day mean temperature is also discussed, especially the relation of surface temperature to the 700-mb height pattern.

1. Introduction and statistical background

During the past 20 years, I have developed several techniques for automatically converting numerical weather predictions into surface temperature. In this lecture I shall describe the evolution of these techniques and their synoptic climatological background.

1.1 Correlation fields

In order to relate a given variable to the field of any potential predictors, such as 700-mb height, it is convenient to compute the simple linear correlation coefficient between that variable and the height at each of a large number of grid points. The sample correlation coefficient may be expressed as:

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} \qquad (1)$$

where x_i and y_i are the ith values of the two variables being correlated, $\bar{x}$ and $\bar{y}$ are their mean values over the sample, and the summation is performed over n cases. The coefficients computed from this equation are plotted on a base map at the respective grid points, and the resultant field is analyzed by drawing lines of equal correlation coefficient. Little smoothing is necessary in drawing isopleths to the plotted data, which usually yield smooth and regular patterns with a few well-defined centers of maximum or minimum correlation.

Correlation fields may be readily interpreted in terms of anomalous flow patterns. The soundness of this analogy was demonstrated by Stidd (1954), who combined the geostrophic wind equation with equations for the correlation coefficients among precipitation, wind

velocity, and height, and then summarized his findings as follows:

> "The regression coefficients of rainfall on pressure (or height) satisfy the geostrophic equation in the same manner as the pressures themselves. Furthermore, if there were no spatial variation in the standard deviation of pressure, the correlation coefficients would, themselves, satisfy this equation. Thus, it may be said that, disregarding the spatial variations in standard deviation of pressure, the correlation field is strictly analogous to an anomalous flow pattern. A tight gradient of the isopleths in the correlation field indicates a strong correlation of the wind tangent to those isopleths."

This interpretation of the correlation field will form the basis for much of the climatological portion of this lecture.

1.2 Linear regression

The maximum correlation coefficient (without regard to sign) in a correlation field indicates how well the given variable can be predicted by a simple linear regression equation of the type

$$\hat{y} = a + bx \tag{2}$$

where a and b are regression coefficients, $\hat{y}$ is the forecast value of the predictand or dependent variable, and x is the predictor or independent variable. The values of the constants a and b are determined by the method of least squares, which minimizes the sum of the squares of the differences between forecast values $\hat{y}$ and observed values y. The coefficient b is the slope of the line of regression and is given by the formula:

$$b = r\frac{s_y}{s_x} \tag{3}$$

while a depends on the mean values of the two variables in accordance with the formula:

$$a = \bar{y} - b\bar{x} \tag{4}$$

where $\bar{y}$ is the mean of the predictand, $\bar{x}$ is the mean of the predictor, r is the correlation coefficient defined in equation (1), s_y is the standard deviation of the predictand, and s_x is the standard deviation of the predictor. An unbiased estimate of s_x is given by the formula:

$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{5}$$

where x_i represents any individual value of x from the first to the nth case, and $\bar{x}$ is the sample mean. The square of the standard deviation is called the variance.

Improved results can be obtained by using additional predictors in a multiple linear regression equation of the type:

$$\hat{y} = a_1 + b_1x_1 + b_2x_2 + \ldots b_kx_k \qquad (6)$$

where a and b are regression constants determined by least squares, and $x_1, x_2 \ldots x_k$ are k different predictors. The closeness of the relation between values of $\hat{y}$ forecast from equation (6) and values of y observed in the dependent data is measured by the multiple correlation coefficient R. This can be defined as follows:

$$R=\sqrt{1-\frac{w}{w_{11}}}=\sqrt{1-\frac{S^2}{s^2}} \qquad (7)$$

where w is the determinant of the complete matrix of simple correlation coefficients including those between the predictand and each of the predictors (in the first column and first row) and those among the various predictors (in the remainder of the matrix); w_{11} is the co-factor of the element in the first column and first row, S is the standard error of estimate of forecasts made from the regression equation; and s is the standard deviation of the predictand.

From equation (7) it follows that the percentage of total variance of the predictand explained (EV) by the predictors in a multiple regression equation equals the square of the multiple correlation coefficient, or:

$$\frac{EV}{100}=R^2=1-\frac{\sum_{i=1}^{n}(\hat{y}_i-y_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \qquad (8)$$

where y_i is the observed value of the predictand, $\hat{y}_i$ is the value forecast from the regression equation (6), and $\bar{y}$ is the mean value over the dependent sample. EV is also known as the reduction of variance RV.

1.3 Screening technique

The object of the screening procedure is to select from a large set of possible predictors only those few which contribute significantly and independently to the forecast of a predictand. This is accomplished by a forward method of multiple regression in which significant predictors are picked in a stepwise fashion, one by one. As a result, a small number of predictors can be selected which contain practically all the linear predictive information of the entire set with respect to a specific predictand. The importance of using a small set of predictors to prevent redundancy and instability of the multiple regression equation and to insure good results when applying it to new data has been emphasized many times.

The steps in the screening procedure have been described by Miller (1958) as follows:

1. "Select that predictor x_1 which explains the largest amount of total variability of the predictand y." (This step coincides with choosing the highest simple linear correlation coefficient.)

2. "Orthogonalize all remaining predictors with the first selected predictor x_1. This transformation is performed by expressing each of the remaining variables as a function of x_1." (This is done by a simple linear regression equation such as equation (2).)

3. "Select the predictor x_2 which explains the largest amount of variability (of y) left unexplained by x_1. The selection of x_2 is made using the procedure of step 1 except that now the orthogonalized transformations of the predictors are used in place of the original predictors."

4. "Orthogonalize all remaining transformed predictors with the second selected predictor x_2. Continue the alternating procedure of selecting and orthogonalizing until one of the selected variables fails to explain a preassigned percentage of the total variability. This percentage is based on the criterion of significance used."

Thus the screening procedure simply orthogonalizes each predictor with respect to the previously selected ones and then chooses as the next predictor that one which correlates highest with the predictand. This procedure is equivalent to computing partial correlation coefficients between the predictand and each of the remaining predictors, holding the first selected predictor constant, and selecting as the second predictor the one giving the highest partial correlation. The third predictor chosen is the one with the highest partial correlation coefficient after removal of the effect of the first two predictors picked, and additional predictors are selected in a similar fashion. An alternate but equivalent interpretation is to define the second predictor as the one giving the highest possible multiple correlation coefficient in combinations of two variables, when the first one has been specified. Similarly, the third predictor is the one giving the maximum multiple correlation in conjunction with the first two picked, etc. Thus at each step, that variable is chosen which, when used in conjunction with all previously selected predictors, will add the most to the multiple correlation coefficient.

1.4 Verification statistics

In ordinary linear regression, standard statistical tests of significance can be performed to determine whether the contribution of a given predictor is significantly greater than that which would be expected by chance. However, the significance level used to reject a variable is quite arbitrary, and the assumptions upon which the tests are based, i.e., normal distribution and independent cases, are seldom satisfied in meteorological data. Furthermore, these tests do not apply precisely in screening because the predictors are not chosen at random, and the highest correlation is being selected from among many. Hence, I have not applied tests of significance giving exact confidence intervals or probability levels. Instead, all prediction equations were tested on independent data samples.

Predictions made from independent data were verified by means of several different statistics because meteorologists have seldom agreed on a single measure for the correctness of a weather forecast. One of the most useful verification statistics is the simple correlation coefficient, which was computed from equation (1) by letting x be the

forecast value and y the observed value. Another statistic frequently used in verification is the root mean square error. This was computed in the form:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m} (\hat{y}_i - y_i)^2}{m-1}} \tag{9}$$

where $\hat{y}_i$ is the ith forecast, and y_i is the corresponding observed value in a series of m forecasts on independent data.

Tests on independent data were also verified in terms of the standard skill score for elements like 5-day mean temperature which are customarily expressed in terms of classes. The skill score SS is defined as:

$$\frac{SS}{100} = \frac{C - E}{T - E} \tag{10}$$

where C is the number of forecasts in the correct class, T is the total number of forecasts, and E is the number expected correct by chance. SS is 100 for perfect forecasts, and positive, zero, or negative for forecasts better, equal, or worse than chance.

In expressing the objective temperature forecasts in terms of classes, the forecasts were first multiplied by the reciprocal of the multiple correlation coefficient valid for each city's regression equation since:

$$R = \frac{s_f}{s_o} \tag{11}$$

where R is the multiple correlation coefficient, s_f is the standard deviation of the forecasts, s_o is the standard deviation of the observed quantity, and all parameters are expressed as departures from their mean values. The regression forecasts were inflated in this way in order to equate their variability with that observed and thus prevent an excessive number of predictions in the central classes.

2. Synoptic climatology of 5-day mean temperature

2.1 Correlation fields

This section discusses correlation fields between the anomaly of 5-day mean surface temperature and the anomaly of simultaneous 5-day mean 700-mb height at 70 grid points for 140 winter cases. Figure 1 presents the charts for 15 cities in the United States, arranged so as to progress from northeast to southwest.

The correlation fields for every city reveal two widely separated centers of maximum correlation, one positive and within 800 mi. of the reference station; the other negative and about 2,000 mi. to the northwest. The positive correlation reflects the fact that the troposphere usually approaches an equivalent barotropic state, with contours tending to be in phase with isotherms. The negative correlation reflects an advective effect, whereby the amplitude of the ridge upstream controls the deployment of cold air southeastward.

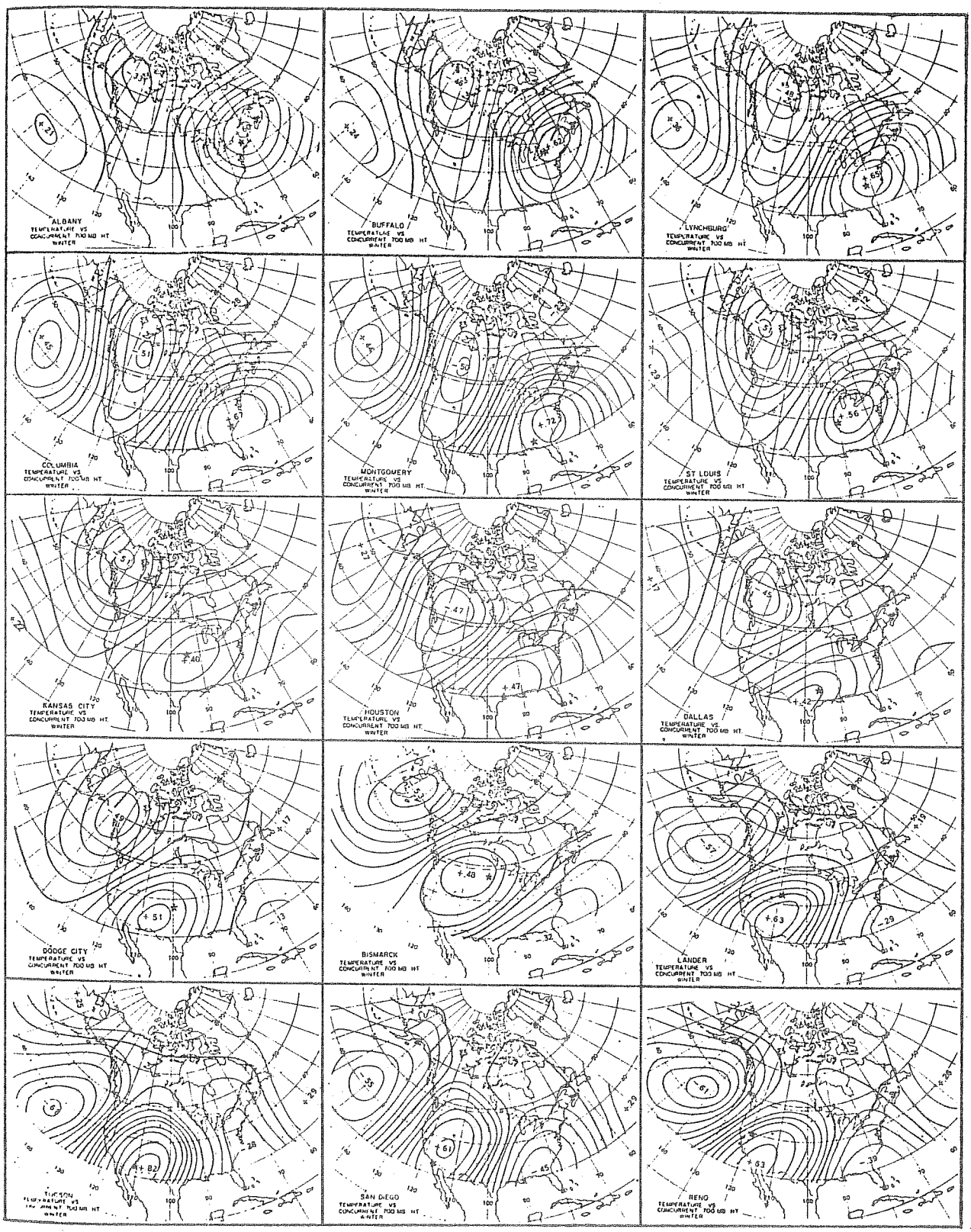

Figure 1. Correlation fields between anomaly of 5-day mean surface temperature at starred city and simultaneous anomaly of 5-day mean 700-mb height at 70 grid points for 140 winter cases. Centers labeled with maximum (interpolated) correlation coefficient; isopleths drawn for units of 0.1, with zero line heavy.

The magnitude of the maximum positive correlation varies from 0.82 at Tucson to 0.40 at Kansas City and averages 0.59 for all 15 cities. There is a distinct tendency for a minimum positive correlation to be found between the Rocky Mountains and the Mississippi River, with higher values in both eastern and western thirds of the country. The distance to the center of maximum negative correlation northwest of the reference city varies from about 2,500 mi. at Dallas to 1,500 mi. at Reno. The magnitude of this negative correlation ranges from -0.61 at Tucson and Reno to -0.33 at Albany and averages -0.51 for all 15 cities. The magnitude of the maximum negative correlation exceeds that of the maximum positive correlation at Bismarck, Kansas City, and Dallas; the two correlations are equal at Houston; and the positive correlation is greater at the 11 other cities.

Figure 1 may be compared with the corresponding diagram for the maximum correlation between 5-day mean anomalies of surface temperature and 700-mb height 2 days earlier during the winter season, previously published for eight cities by Klein et al. (1959). The centers of maximum correlation are generally located in the same areas, but the magnitudes of the correlation coefficients are different. The positive correlations are slightly higher for simultaneous heights, but the negative correlations are slightly greater in magnitude for heights 2 days earlier. This difference implies that temperatures lag heights upstream, probably as a result of the time required for advection of air masses from distant source regions, but there is no appreciable lag between temperature and local height.

2.2 Regional characteristics

The properties of the correlation fields of Fig. 1 are further summarized in Figs. 2-5. Figure 2 shows the geographical distribution of the local correlation coefficient between simultaneous anomalies of surface temperature and 700-mb height directly overhead. The correlations are positive everywhere, with maxima over the Southeast and Southwest and minima over the Northwest and Great Plains. Likewise, the correlation fields indicate that cyclonic curvature of 700-mb contours favors cold weather and anticyclonic curvature accompanies warm weather.

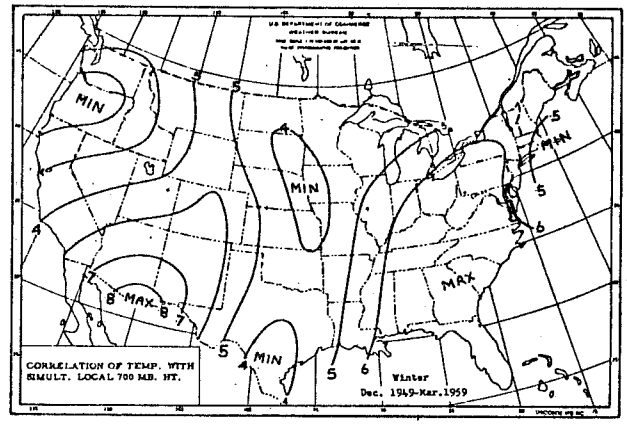

Figure 2. Simple linear correlation coefficient between 5-day mean anomaly of surface temperature in winter and simultaneous 700-mb height at the same point, with centers labeled as maximum or minimum.

Figure 3 was prepared by assuming as before that the isopleths in the correlation fields are analogous to lines of equal height departure from normal. The direction of the anomalous (geostrophic) flow at 700 mb for below normal temperatures at each city was then obtained. Four homogeneous areas are delineated in this figure. Along the Atlantic and Gulf coasts, northwesterly anomalous flow accompanies cold weather, while southeasterly components produce high temperatures. In most of the Mid-west and Far West, northeasterly anomalous flow favors cold conditions, and southwesterly components go with mild weather. The Northern and Central Plains are rather surprising in indicating that southeasterly components are associated with abnormal cold, and northwesterly flow with warmth. This can be explained in terms of polar outbreaks and upslope cooling with easterly winds, in contrast to Pacific air and foehn warming with a Basin High.

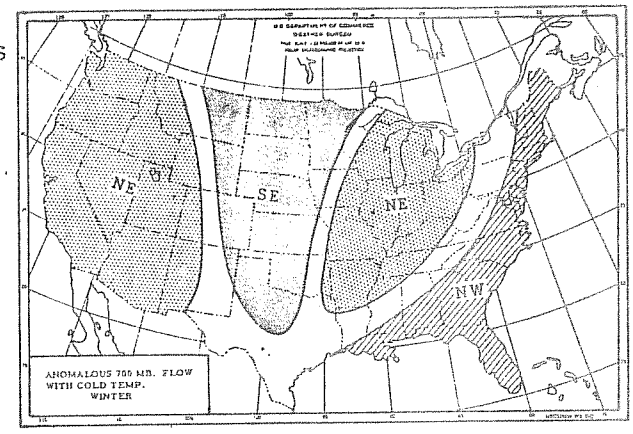

Figure 3. Local component of anomalous 700-mb flow conducive to below normal temperature in winter with symbols as follows: NE=northeasterly, SE=southeasterly, NW=northwesterly. In unshaded areas no single wind direction predominates.

Figure 4 shows the zonal distance from each city to the nearby axis of maximum positive correlation. Distances are negative in the Great Plains because here the positive axis is west of the reference station, so that local (anomalous) northerly components go with warmth. In the remainder of the nation southerly components favor high temperature since the axis of positive correlation is everywhere east of the station. The distances, however, range from less than 100 mi. along the east coast to over 800 mi. in the Pacific Northwest.

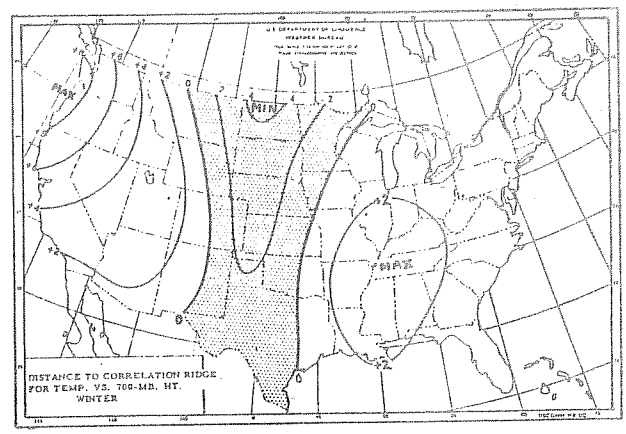

Figure 4. Zonal distance in hundreds of miles to the nearest axis of positive correlation between simultaneous 5-day mean anomaly of surface temperature and 700-mb height for 140 winter cases. Distances are considered positive when the correlation ridge is east of the reference city, negative when it is to the west (stippled).

2.3 Schematic models

The preceding material has been summarized in the form of three schematic models illustrated in Fig. 5. These were prepared by assuming zonal flow at 700 mb on the normal winter map. The first model applies to Washington, Oregon, northern California, western

Idaho, and northwestern Nevada. In this region the optimum portion of the wave train for unseasonable warmth is about 400 to 800 mi. west of a large-amplitude ridge under strong southwesterly flow. Conversely, the most favorable location for low temperatures is from 400 to 800 mi. west of a deep trough. In view of the low correlations in this area between local height and temperature (Fig. 2), wind direction (southwesterly or northeasterly) is probably just as important as proximity to the ridge or trough in determining surface temperature.

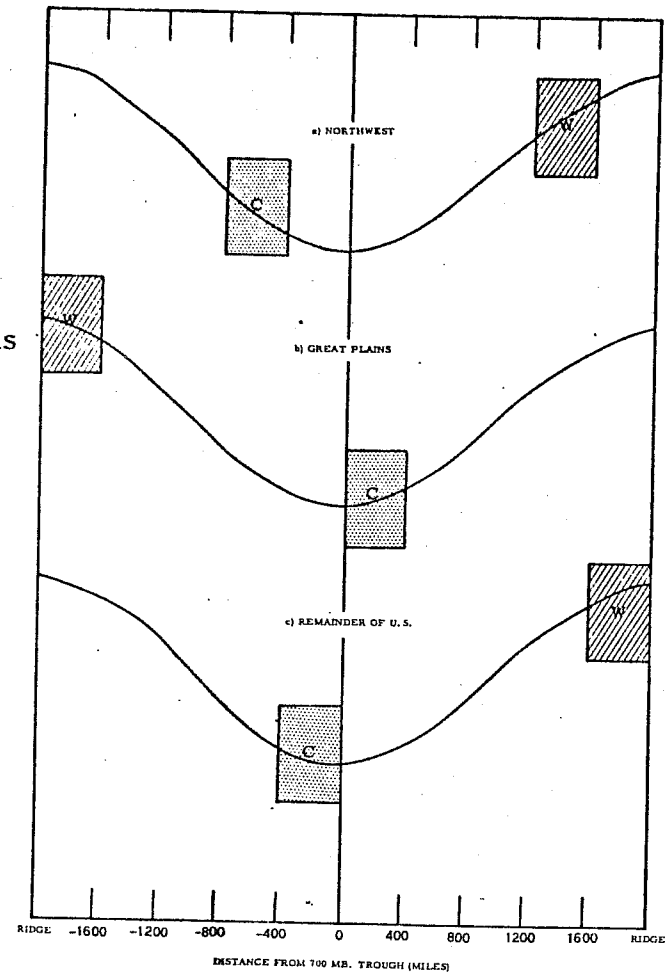

Figure 5. Optimum regions for cold (C) and warm (W) temperatures in winter in different parts of the United States relative to a sinusoidal 700-mb contour. Distance from the 700-mb trough given along the abscissa is assumed to be 2,000 miles to the ridges up- and downstream, with positive values when the trough is west of the reference area and negative values when the trough is east of the area.

The second model is for most of the Great Plains. Here the optimum location for mild weather is from the ridge line to 400 mi. east of it, while cold conditions prevail in and a few hundred miles east of the trough. Since there is a low correlation in this area between local height and temperature (Fig. 2) and a high negative correlation with the remote center (Fig. 1), it is likely that the flow components (westerly or easterly) are just as important as dynamic effects associated with the local ridge or trough in determining surface temperatures. In this regard, cold weather in the Plains is favored by an out-of-phase blocking-type circulation at 700 mb, with a ridge at high latitudes surmounting a trough at low latitudes; while warm temperatures are likely with the opposite, high-index, type of circulation.

In all other portions of the country the third model is applicable. This shows warmest weather under anticyclonic curvature from the ridge line to 400 mi. west of it. Conversely, cold conditions occur with cyclonic curvature under a deep trough or a few hundred miles to its west. Although southerly flow favors warmth and northerly flow cold, wind direction is probably of less significance here than proximity to the local ridge or trough. This is suggested by the high correla-

tions between local height and temperature (Fig. 2), the small distance to the positive correlation axis (Fig. 4), and the higher correlations with the positive local center than with the negative remote center (Fig. 1).

3. Prediction of 5-day mean surface temperature

In previous papers by Klein et al. (1959,1960), an objective method of forecasting 5-day mean surface temperature in the United States was derived for winter, fall, and spring seasons. This section will review the method and extend it to summer.

3.1 Relation between temperature and 700-mb height

The first step in deriving the objective method was to obtain simple linear correlation coefficients between 5-day mean anomalies of surface temperature and of 700-mb height two days earlier. The heights were taken at 70 grid points, the temperatures at 39 cities, and the dates covered a 10-year period from 1949 to 1958.

The resulting correlation fields were similar to those shown in Fig. 1 for simultaneous values and therefore will not be reproduced here. It is interesting to note, however, that the summer season exhibits two special characteristics. Whereas the centers of positive correlation in the vicinity of the reference city and of negative correlation about a half wavelength upstream are approximately equal in absolute magnitude during fall and winter, in summer the former is consistently greater. This indicates an increased importance of local height, at the expense of heights upstream, during summer. Secondly, the magnitude of the maximum positive correlation in summer is less than in the other three seasons, varying from 0.76 at Spokane to as low as 0.18 at Brownsville and 0.24 at Miami. These results indicate that during the warm season local or low-level factors, like cloudiness, sea breeze, humidity, precipitation, and stability, have a greater effect upon surface temperature than advection of air masses or the large-scale pattern of the upper air circulation.

The second step in the derivation was to apply the screening technique to the same temperatures and heights used for the correlation fields. Multiple regression equations for predicting temperature from height were first derived for winter and later for the other three seasons of the year. In most cases they incorporate one height in the vicinity of the station (positive correlation), one height near the center of negative correlation a half wavelength upstream, and about two additional heights which contribute significantly to the temperature variance. The average number of heights used in each season, as well as the mean explained variance, is shown in Table 1. There is little difference among winter, fall, and spring, but the explained variance for summer is about 10 percent less than for the other seasons. For the year as a whole, about 52 percent of the variance is explained by four heights.

Although the equations described above contain a 2-day lag between temperature and height, they can also be applied to specification. For example, during 40 winter cases chosen at random between 1948 and 1957

Table 1. Explained variance and number of variables in equations for predicting 5-day mean surface temperature from 5-day mean 700-mb heights two days earlier. Results are averaged for 39 cities in the United States on the developmental sample.

Season	EV	No. of heights
Fall	54.7	3.9
Winter	54.5	3.6
Spring	54.6	3.9
Summer	45.0	4.0
Year	52.2	3.9

as part of another experiment, in which 5-day mean temperature classes were specified from observed simultaneous 5-day mean 700-mb height anomalies, the multiple regression equations yielded a skill score of 23, averaged over 100 cities in the United States. This compares favorably with the average skill score of 13 obtained by four experienced forecasters who were asked to estimate the temperature pattern from the observed height anomalies for the same 40 cases.

Although the above equations were derived from 5-day mean data, they have been found to give better results with 30-day mean data, probably because the latter smooths the effects of small-scale and short-period phenomena. For example, in the test described in the preceding paragraph, the regression equations specified the correct class of the 5-day mean temperature 40 percent of the time. When the same equations were applied to specify monthly mean temperatures, they were correct 46 percent of the time in winter, 50 percent in fall, and 39 percent in spring and summer (Klein, 1962). In a later experiment they produced an average skill score of 31 when applied to 35 winter 30-day periods (overlapping) between December 1959 and March 1964.

As a result of these tests, the equations for predicting temperature from heights only are used routinely in the National Meteorological Center as an aid in preparing 30-day outlooks. In 5-day mean prediction, however, they have been superseded by another set of equations which includes the effect of initial surface temperature.

3.2 Incorporation of initial surface temperature

It is well known that mean values of surface temperature exhibit marked persistence. As a result, better 5-day mean temperature forecasts were obtained from multiple regression equations containing initial local temperature, and two 700-mb heights, than from the equations described previously containing 3 or 4 heights (Klein et al., 1959).

It seems reasonable to expect that improved results can be secured by including the entire temperature field, in addition to local temperature. For example, Fig. 6 shows the correlation between anomalies of 5-day mean temperature at Dallas, during 10 winters from

1948 to 1957, and 5-day mean surface temperature for a period centered 4 days earlier at other cities. Whereas the auto-correlation for Dallas is only 0.47, the maximum cross-correlation is 0.60, in an area about 400 mi. northwest of the station. Similar correlation fields were constructed for each of the other cities and seasons, and their characteristics may be summarized as follows:

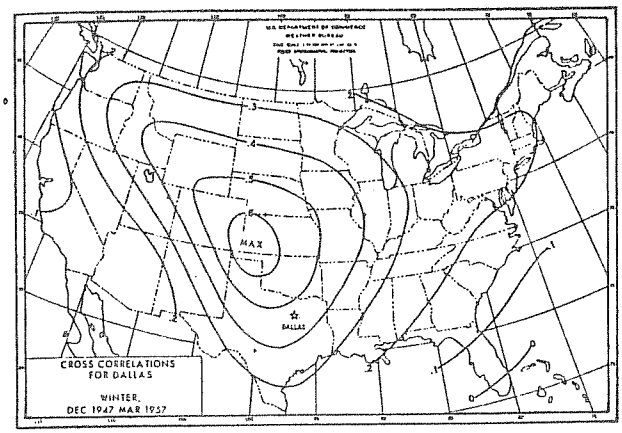

Figure 6. Simple correlation coefficients between 5-day mean temperature anomaly at Dallas (starred) and 5-day mean temperature anomaly four days earlier at a grid of 30 cities for winter season.

(a) The correlations are nearly all positive and generally higher in winter than in the other seasons.

(b) The distances to the points of maximum correlation are generally greater in fall and spring than in winter and summer and greater east of the Continental Divide than west of it.

(c) Except for a few cities in the West in summer, most centers of maximum positive correlation are located north or west of the reference city.

The above findings have the following implications for the climatology of 5-day mean temperature:

(a) Persistence of local temperature is greater in winter than in the other seasons.

(b) Advection of temperature from remote areas is more important than local persistence in fall and spring and in the eastern two-thirds of the Nation.

(c) Mean temperature anomalies tend to move southeastward over intervals of four days.

The screening program was applied to derive prediction equations for the anomaly of 5-day mean surface temperature as a function of two basic parameters: the anomaly of 5-day mean 700-mb height at 70 grid points two days earlier, and the anomaly of 5-day mean surface temperature at 39 stations four days earlier. This was done separately for each of the four seasons and for each of the cities. In nearly all cases at least two heights were selected, one located in the vicinity of the remote center of negative correlation to the northwest, and the other near the local center of positive correlation. In addition, at least one temperature was usually selected, generally from the center of maximum positive correlation near or northwest of the city.

At most cities the screening procedure was continued after the third step, with one to three additional heights and temperatures (generally to the west of the station) being selected as significant predictors. The final equations are summarized in Table 2 for each season separately.

Table 2 shows that the properties of the final prediction equations, when averaged over all cities in the network, are quite uniform in most respects from season to season. The average number of heights selected varies only from 3.0 to 3.2, the number of temperatures from 1.3 to 1.8, and the explained variance from 52 percent (in summer) to 62 percent (in winter). There is a much greater variation, however, in the percent of cities for which local temperature was selected as a predictor (last column). This ranges from 23 percent in spring to 57 percent in winter. For the year as a whole, the regression equations explain about 59 percent of the variance by means of approximately 3 heights and 1.5 temperatures, which include the local temperature at about 2 out of every 5 cities.

Typical results are presented in Fig. 7 for Cleveland, Ohio during the winter season. The most important single predictor of Cleveland's mean temperature during the next 5 days is the 5-day mean temperature centered on forecast day at Indianapolis, Ind., and the correlation between the two variables is 0.60. The positive sign of the regression coefficient preceding Indianapolis in the prediction equation written at the top of the figure indicates that low temperatures in that city tend to be followed by cold weather in Cleveland, and conversely for warm conditions.[1]

Table 2. Explained variance and number of variables in equations for predicting 5-day mean surface temperature anomaly from fields of both 700-mb height two days earlier and surface temperature four days earlier. Results averaged for 39 cities in the United States on the developmental sample.

Season	EV	No. ht.	No. temp.	No. var.	Percent local temp.
Fall	60.5	3.2	1.8	4.9	35.9
Winter	62.3	3.0	1.6	4.7	56.6
Spring	60.3	3.0	1.3	4.4	23.0
Summer	52.2	3.1	1.3	4.4	48.6
Year	58.8	3.1	1.5	4.6	41.0

[1]Strictly speaking, the coefficients in the multiple regression equation should not be interpreted as simply as has been done in this section since they reflect the joint, rather than the individual, contribution of the various predictors. However, inspection of numerous temperature prediction equations containing from 1 to 8 variables indicates that, at least for this type of data, the regression coefficients may fluctuate in magnitude as additional terms are added, but they rarely change in sign.

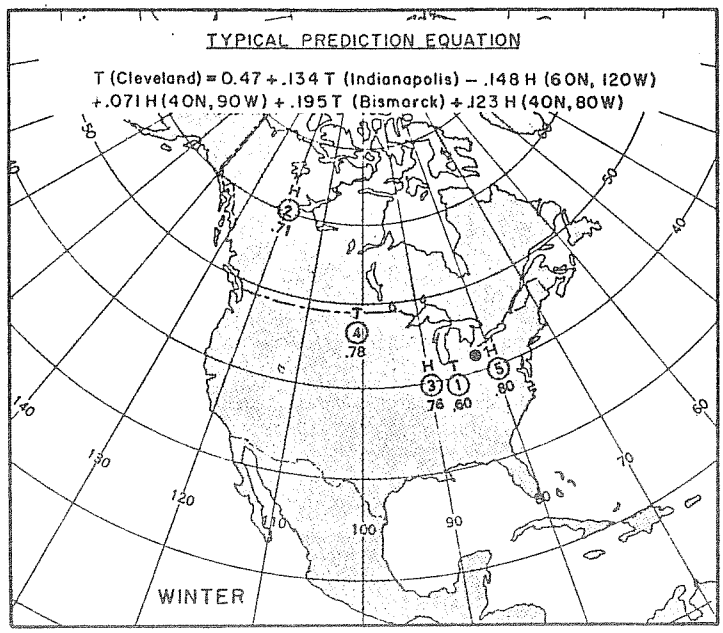

Figure 7. Multiple regression equation used in predicting 5-day mean temperature at Cleveland, Ohio (located by heavy solid dot) during the winter season, as a function of 5-day mean 700-mb height centered two days earlier (H) and 5-day mean surface temperature centered four days earlier (T) at points given in parentheses. The location of the predictors is given by the open circles, the order of selection by the number inside the circle, the type of variable by the letter above the circle, and the multiple correlation coefficient after inclusion of the given predictor by the decimal below the circle.

The second most important predictor is the 5-day mean 700-mb height centered two days before the forecast temperature and located in northwestern Canada at 60° N, 120° W. The combination of height at this point plus temperature at Indianapolis yields a multiple correlation of 0.71. The coefficient of this variable has a negative sign, as expected from the fact that high heights in a ridge of large amplitude in northwestern Canada produce strong northwesterly flow of colder polar air and hence low temperatures at Cleveland, while low heights lead to mild Pacific air in strong westerly flow.

Combination of the first two predictors with an additional one produces best results when the 700-mb height at 40° N, 90° W is used, raising the multiple correlation to 0.76. The positive sign of the coefficient before this variable suggests that high heights at this point lead to warm temperatures at Cleveland, while low heights are followed by cold weather.

The fourth predictor is the current temperature at Bismarck, N. Dak. which raises the multiple correlation to 0.78. The positive sign of its coefficient, like that of the temperature at Indianapolis, reflects the prevailing west to east drift of air masses.

The fifth predictor is the 700-mb height in the vicinity of Cleveland at 40° N, 80° W. Its positive coefficient, like that of the height selected earlier at 40° N, 90° W, reflects the fact that ridges are usually warm and troughs cold. At this point the screening process was stopped because no additional predictor produced any significant increase in the multiple correlation of 0.80 attained by these five variables.

Similar equations have been derived for each season of the year for 39 cities covering most of the United States. In all cases the equations appear to be physically reasonable. Consequently, they are quite stable; i.e., they gave nearly as good results in tests on independent samples as they did on the original developmental data. However, these tests were made using observed 5-day mean values of all predictors; in actual forecast practice, prognostic values must be used in the equations, with results discussed below.

3.3 Verification

Numerous experiments have been conducted to determine which height and temperature input to the prediction equations gives the best objective forecasts of 5-day mean temperature on an operational basis. In an earlier paper (Klein et al., 1960), it was concluded that the best forecast of upper-level heights then available as input for these equations was the 36-hr baroclinic 700-mb prognosis regularly prepared in the National Meteorological Center (NMC) from 1200 GMT data on forecast day. In a later paper (Klein et al., 1962), it was found that the most effective temperature input was the daily temperature prediction for the next day regularly prepared each morning at forecast offices throughout the country. Objective predictions of 5-day mean temperature were therefore prepared on a routine basis three times a week by using these short-range prognoses of 700-mb height and surface temperature as input in the prediction equations.

The official 5-day mean temperature forecasts prepared by NMC are expressed in five classes: much below, below, near, above, and much above normal. In order to express the objective temperature forecasts in terms of these classes, they are first multiplied by the reciprocal of the multiple correlation at each city in accordance with equation (11). The forecasts are then expressed as one of the five temperature classes according to the appropriate class limits at each city. After the forecast class for each city is plotted on a map, lines are drawn by interpolation delineating the distribution of the classes over the conterminous United States. The objective forecasts prepared in this way are routinely verified at a grid of 100 points fairly evenly spaced over the country in the same manner customarily used in extended forecasting.

Verification results for the 6-yr period from the fall of 1958 through the summer of 1964 were summarized on a seasonal, yearly, and overall basis by Klein (1965). The verification is given in terms of the standard skill score (equation (10)), computed from the percent of forecasts in exactly the correct class. It is also given in terms of the percent either exactly correct or correct within one class, in order to give credit for a near hit such as forecasting much below when below normal is observed.

The objective forecasts had positive skill in each season, with skill scores ranging from a maximum of 23 in the fall of 1960 to a minimum of 8 in the summer of 1961. The overall skill score for all cases was 16, but there was a consistent seasonal variation, with a difference of 3 to 4 points between average skill scores for fall and winter, on the one hand, and summer on the other. It is noteworthy that the skill score of the objective forecasts has been consistently higher than that of persistence, but slightly lower than that of the official forecaster, who has the objective forecast, the persistence control, and numerous additional tools at his disposal. In terms of percent within one class, however, the objective forecasts have been as good as the official ones, with an overall average of 80 percent exactly correct or one class off. As a result, the objective method is still being used to furnish guidance for 5-day and 10-day mean temperature forecasts produced at NMC.

4. Prediction of daily mean temperature

Since use of daily input in equations derived from mean data gives skillful 5-day mean temperature forecasts, can equally good results be obtained by applying these equations to make daily forecasts? Of course, equations for this purpose could be derived directly, but considerable success might be achieved by applying the already existing equations. If we neglect the difference in spatial scale between daily and 5-day mean phenomena, the objective 5-day forecasts can be converted into daily ones merely by increasing their numerical magnitude to compensate for the increased variability of daily compared to 5-day mean temperatures. For this purpose, it is necessary to multiply the predictions by the ratio of the standard deviation of daily temperatures to the standard deviation of 5-day mean temperature. After considerable experimentation, it was found that best results could be obtained by using a value of 1.4 for this ratio.

Figure 8 is based on forecasts prepared in this way by Klein et al. (1962) for 100 points, 5 temperature classes, and 37 cases during the fall of 1959, but this verification was obtained from temperatures observed on each individual day[2], rather than from the 5-day mean. The abscissa gives the number of days after forecast day (0) for which the forecast was verified, where days two to six constitute the customary 5-day forecast period. The ordinate shows the percent of the contiguous United States which was predicted in exactly the correct temperature class. The horizontal dashed line gives the amount that would be expected correct by chance, 22 percent. This is slightly lower than the score that would be expected by always forecasting the climatological normal, 25 percent.

The open circles represent the score of the objective predictions verified as daily forecasts. This score reaches a maximum of almost 40 somewhere between the 2d and 3d days and then drops off rapidly, although remaining above chance even on the 6th or final day of the period. If the objectives were perfect 5-day mean temperature forecasts, they would still fall short of 100 percent accuracy as daily

[2]The daily temperature is computed by taking the mean of the maximum and minimum.

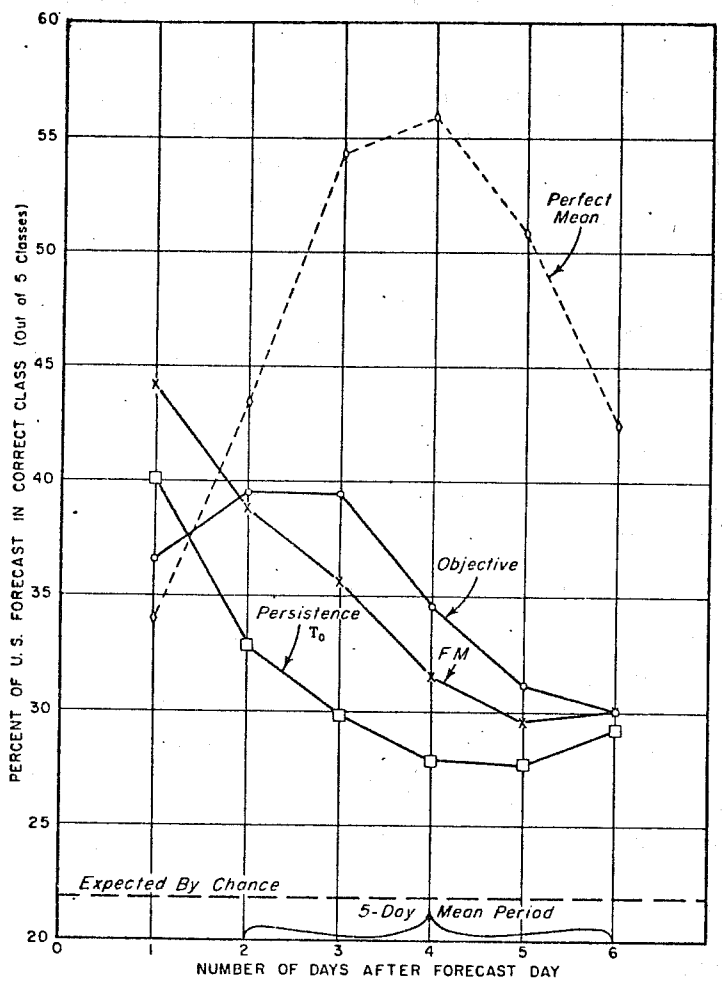

Figure 8. Percent correct for various forecasts verified in terms of temperatures observed on each individual day of the period (analyzed by means of daily class limits) for 37 cases from September 15 to December 8, 1959. The open circles were obtained from the objective forecasts, the open squares from the 5-day mean centered on forecast day (T_0), the x's from persistence of the local forecasts for day 2 (FM), and the open diamonds from the 5-day mean temperatures actually observed.

forecasts. This is indicated by the dashed curve in Fig. 8, which was obtained from 5-day mean temperatures actually observed during the fall of 1959. Although this curve reaches a peak as expected on the 4th or middle day of the 5-day period, even on this day it scores only about 56 percent correct. The differences between the curve for the objective and that for the perfect mean indicate first, that the objective method can still stand a lot of improvement, and second, that the objective forecast tends to be too slow; i.e., to be more accurate at the beginning than at the end of the 5-day forecast period.

The remaining two curves are persistence controls. The line of open squares was obtained from the estimated 5-day mean temperature centered on forecast day (T_0) and is similar to what would be obtained by use of the latest daily observed temperature. A later and

therefore more skillful measure of persistence is given by the line of x's obtained by using the local forecasts or FM's for day 2. Theoretically this curve should peak on the second day, but it actually scores highest on day 1, thereby indicating that the FM forecasts (like the objectives) tend to be too slow. The most important feature of Fig. 8 is the fact that the objective forecasts score higher than either T_O or FM on each day of the 5-day mean period (days 2 to 6), with maximum difference on day 3.

In order to test further the ability of the objective method to predict daily temperatures, all objective forecasts made during the cool season of 1960-61 were converted to daily forecasts and then compared to the mean daily temperature actually observed 2 and 3 days later. The objective predictions had smaller absolute errors than the FM's or the normal on both the 2d and 3d days, but the margin of superiority over the FM's was very small on day 2, in general agreement with the results shown in Fig. 8. As a result of this finding, facsimile transmission of 3-day temperature forecasts, along with a 72-hr prognostic surface map, was initiated by NMC on Sept. 18, 1961.

Verification of four years of the 3-day mean temperature forecasts (Klein, 1966) showed that the linear correlation between predicted and observed anomalies was 0.50 for the official, 0.47 for the objective, and 0.36 for persistence of the local 48-hr forecast (FM). By using observed instead of prognostic height and temperature input in the regression equations, I later produced 2-day mean temperature forecasts with an overall correlation of 0.53, compared to 0.51 for the FM's. The mean temperatures were then converted to maximum (max) and minimum (min) temperatures by simply adding the predicted anomaly to the normal max or min. Verification of these forecasts for a one-year period from Sept. 1964 through Aug. 1965 showed that the objective forecasts were superior to chance, climatology, and persistence, but not quite as good as subjective forecasts (Klein, 1966). A new project was therefore started in 1964 to acquire daily max/min data and derive forecast equations directly from those data. The results will be described in the next section.

5. Prediction of max/min temperatures

5.1 Derivation

An improved technique for objective prediction of max and min temperatures was derived by Klein et al. (1967) and modified by Klein and Lewis (1970). The revised method makes use of multiple regression equations derived for 131 first-order U.S. stations and 12 in southern Canada from 18 years of daily data (1948-65) stratified by 2-month periods (January-February, March-April, etc.). The predictors were selected by the computer by screening the following parameters: 1) 700-mb heights and 700-1000-mb thickness observed at 67 grid points in North America ∿12 hr before the valid time of the prognostic temperature; 2) maximum and minimum temperatures observed at the network of 143 stations from 12-24 hr before the prognostic valid time; and 3) the day of the year.

The screening process was stopped at the point where no additional pair of predictors could increase the explained variance by as much as 2 percent. This criterion was used as the automatic cutoff point for all equations derived in this project.

The derivation scheme first involves linear, multiple-regression equations computed by the method of least squares, in the form:

$$T_x^0 = a + \sum b_i Z7_i^{12} + \sum c_i H7_i^{12} + \sum d_i T_n^0 + \sum e_i T_x^{-1} + fD, \quad (12)$$

$$T_n^0 = g + \sum h_i Z7_i^{00} + \sum j_i H7_i^{00} + \sum k_i T_n^{-1} + \sum l_i T_x^{-1} + mD. \quad (13)$$

The first equation gives today's maximum temperature T_x^0 as the sum of a constant plus certain observed 700-mb heights at 1200 GMT ($Z7_i^{12}$), 700- to 1000-mb thickness at the same time ($H7_i^{12}$), minimum temperatures today (T_n^0), maximum temperatures yesterday (T_X^{-1}), and the day of the year (D), where each predictor is multiplied by its appropriate regression coefficient. Similarly, today's minimum temperature T_n^0 is given as a linear combination of selected 700-mb heights ($Z7_i^{00}$) and thickness ($H7_i^{00}$) at 0000 GMT, minimum (T_n^{-1}) and maximum (T_X^{-1}) temperatures observed yesterday, the date (D), and a constant.

A typical equation resulting from this derivation is illustrated in Fig. 9 for the max temperature at Bismarck during January and February. The first variable selected is the local min temperature on the same day, which, taken by itself, would explain 68.6 percent of the variance and produce a standard error of estimate of 8.6F. The second variable selected is the max temperature on the previous day to the northwest, at Calgary, Alberta, which, taken in conjunction with the Bismarck min temperature selected first, raises the reduction of variance to 77.7 percent and lowers the standard error to 7.3F. The third variable selected is the thickness just north of Bismarck at 50 N, 100 W, while the fourth is the 700-mb height at 55 N, 95 W. The negative sign before the latter term indicates that cold weather at Bismarck goes with easterly flow and blocking activity, while warm weather is accompanied by high index, westerly flow. The screening was stopped at this point with the multiple regression equation written in the lower left, a reduction of variance of 83.6 percent, and a standard error of 6.2F, because no other pair of predictors could increase the variance explained by even 2 percent.

Table 3 summarizes the characteristics of the multiple-regression equations derived in the first screening for the 131 cities in the United States and all months of the year. The first row gives the standard deviation of temperature and illustrates the well-known fact that temperatures are less variable in summer than in other seasons. This makes it difficult to explain a high percent of summer variability, so that the reduction of variance averages only 61 percent in July-August, compared to about 78 percent in the other months. In nearly all months, both the standard deviation and the reduction of variance are slightly higher for the maximum than for

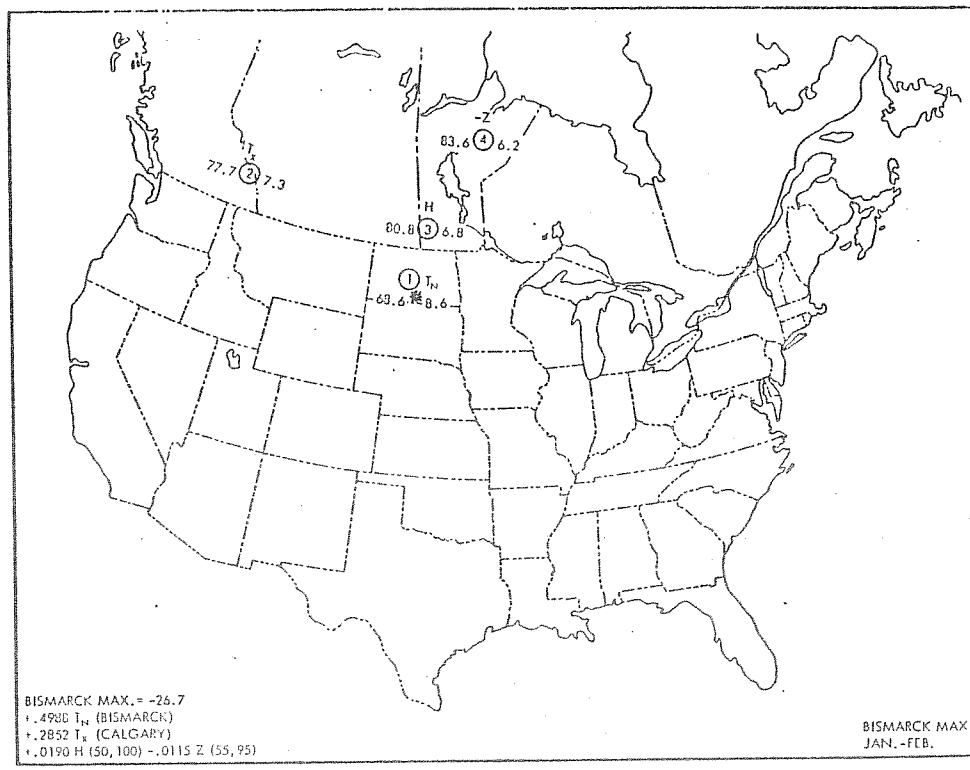

Figure 9. Multiple regression equation derived by screening for the maximum temperature at Bismarck, N.Dak., (star) during January-February. T_X is the maximum temperature (in ^{o}F plus 100), T_n the minimum temperature, H the 700-1000 mb thickness (ft), and Z the 700-mb height (ft). The location of the selected predictor is given by the open circle, the order of selection by the number inside, the reduction of variance (%) after inclusion of the given predictor by the number on the left, and the standard error of estimate (^{o}F) at each step by the number on the right.

the minimum. The standard error of estimate varies from 3.0F for the minimum in July-August to 5.6F for the minimum in January-February. For all months, it averages slightly larger for the maximum (4.5F) than for the minimum (4.3F).

The regression equations usually contain 4-5 variables, made up of at least one max temperature, one min temperature, one 700-mb height, and one 700- to 1000-mb thickness. Equations which forecast the min temperature contain on the average more than four times as many min (1.7) as max (0.4) temperature predictors, but equations for the max average only slightly more max (1.2) than min (0.8) predictors. The day of the year is quite unimportant, being selected only 15 percent of the time on the average. The local temperature, at the reference station itself, is selected about 70 percent of the time, but slightly more often for the min than the max.

5.2 Operational system

Table 4 illustrates the system used in preparing max and min temperature forecasts on an operational basis. The forecasts are prepared twice a day on the large NOAA computer in Suitland, Md., and the

Table 3. Characteristics of regression equations for predicting maximum and minimum temperatures from 700-mb heights, 700- to 1000-mb thicknesses, and surface temperatures, averaged for 131 cities in the United States.

	Jan. to Feb.	Mar. to Apr.	May to June	July to Aug.	Sept. to Oct.	Nov. to Dec.	Mean
For predicting maximum temperatures							
Standard deviation (°F)	11.3	12.0	9.2	6.2	10.2	11.5	10.1
Reduction of variance (%)	79.4	80.9	76.1	61.4	81.0	82.5	76.9
Standard error (°F)	5.0	5.1	4.3	3.8	4.3	4.7	4.5
No. of variables	4.3	4.3	4.5	5.3	4.2	4.2	4.5
No. of maximum temperatures	1.1	1.1	1.0	1.7	1.0	1.1	1.2
No. of minimum temperatures	0.8	0.6	0.7	1.4	0.6	0.7	0.8
No. of 700-mb heights	1.0	1.2	1.5	1.5	1.1	0.8	1.2
No. of thicknesses	1.4	1.2	1.3	0.7	1.3	1.3	1.2
No. of days of year	0.0	0.2	0.0	0.0	0.2	0.3	0.1
No. of local temperatures	0.6	0.4	0.6	0.6	0.5	0.6	0.6
For predicting minimum temperatures							
Standard deviation (°F)	11.1	9.9	7.4	5.0	8.9	10.6	8.8
Reduction of variance (%)	74.2	76.9	76.2	61.4	77.5	75.4	73.6
Standard error (°F)	5.6	4.5	3.6	3.0	4.1	5.2	4.3
No. of variables	4.4	4.5	4.3	5.5	4.3	4.4	4.6
No. of maximum temperatures	0.2	0.6	0.4	0.8	0.3	0.2	0.4
No. of minimum temperatures	1.5	1.5	1.4	2.2	1.6	1.8	1.7
No. of 700-mb heights	1.3	1.1	1.0	1.5	1.1	1.2	1.2
No. of thicknesses	1.4	1.2	1.2	0.8	0.9	1.1	1.1
No. of days of year	0.0	0.1	0.3	0.2	0.4	0.1	0.2
No. of local temperatures	0.7	0.7	0.8	0.9	0.8	0.8	0.8

Table 4. System for preparation of operational maximum and minimum temperature forecasts for 12-72 hr in advance.

Forecast	Output	Height and thickness input	Surface temperature input
		From 1200 GMT data	
12 hr	Max	Observed 1200 GMT today	Min observed today, max observed yesterday
24 hr	Min	12-hr numerical progs	Min observed today, 12-hr prog max
36 hr	Max	24-hr numerical progs	24-hr prog min, 12-hr prog max
48 hr	Min	36-hr numerical progs	24-hr prog min, 36-hr prog max
60 hr	Max	48-hr numerical progs	48-hr prog min, 36-hr prog max
72 hr	Min	60-hr numerical progs	48-hr prog min, 60-hr prog max
		From 0000 GMT data	
12 hr	Min	Observed 0000 GMT today	Max and min observed yesterday
24 hr	Max	12-hr numerical progs	Max observed yesterday, 12-hr prog min
36 hr	Min	24-hr numerical progs	24-hr prog max, 12-hr prog min
48 hr	Max	36-hr numerical progs	24-hr prog max, 36-hr prog min
60 hr	Min	48-hr numerical progs	48-hr prog max, 36-hr prog min
72 hr	Max	60-hr numerical progs	48-hr prog max, 60-hr prog min

same equations for the max and the min are used in 12-hr steps on an iterative basis. Here we assume that the min and max temperatures occur at their normal times of day; namely, in the early morning and late afternoon. For example, at 1200 GMT, the first forecast made is for the max that afternoon, and it is based on heights and thicknesses observed at forecast time, on the min temperature reported that morning, and on the max temperature observed on the previous day. The second forecast is for the min the following day and is based on 12-hr numerical forecasts of 700-mb height and 700- to 1000-mb thickness, on the same min used as input for the 12-hr

forecasts, and on the max for today generated in the first step. The third forecast is for the max tomorrow and is based on 24-hr numerical forecasts of height and thickness, on the 24-hr forecast of the min temperature made in step 2, and on the 12-hr forecast of the max made in step 1. The fourth forecast, for the min the day after tomorrow, is based on 36-hr numerical forecasts of upper air input and on the system's 24- and 36-hr surface temperature forecasts. The fifth forecast, for the max the day after tomorrow, uses as input 48-hr numerical prognoses and automated 48- and 36-hr temperature forecasts. The system is continued for another 12 hours and then stopped because no thickness forecasts are routinely available beyond 60 hr.

The numerical forecasts used as input to the prediction equations are obtained from operational numerical models run by NMC. Temperature input to the prediction equations now consists of observed max and min temperatures transmitted in the synoptic code at 0600 and 1800 GMT. These teletype reports are monitored by the NMC automatic data processing system. Unfortunately, on the average, about a dozen reports are missing each day. In these cases, the computer uses the objective forecast made 12 hr previously in place of the missing temperature, so that the prediction system is fully automated.

5.3 Verification

The forecasts are verified routinely at the end of each month for each of 131 stations in the 48 states. Fig. 10 (Klein, 1972) shows the mean absolute error for max and min (combined) for all stations and all months by calendar year. Two curves are presented, one for 24-hr projections and the other for 48-hr projections. The curves are generally parallel and both show a steady downward trend. From 1966 to 1971 the average error declined from 5.5F to 4.0F for 24 hr and from

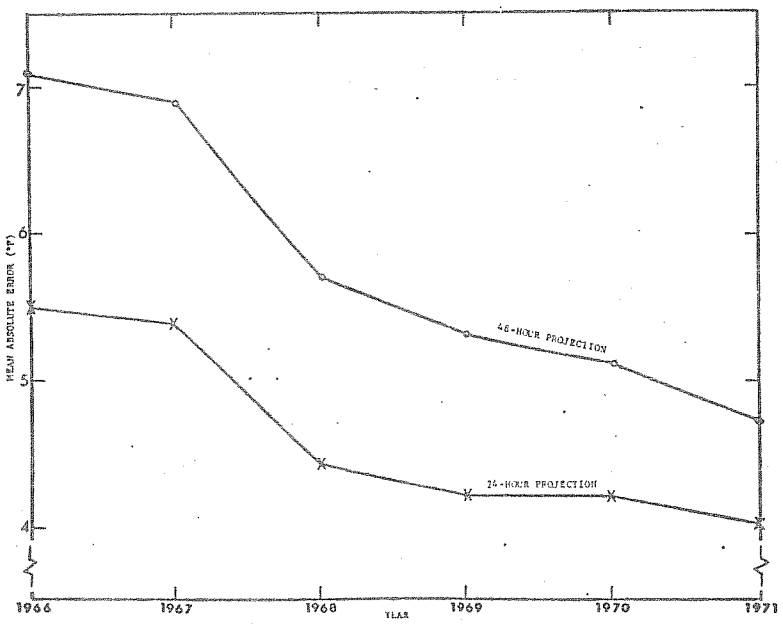

Figure 10. Mean absolute errors (^{o}F) of automated forecasts of maximum and minimum surface temperature for 24-hr projections (crosses) and 48-hr projections (circles). The errors are averages for maximum and minimum at 131 stations in the conterminous United States by calendar year from 1966 through 1971.

7.1F to 4.7F for 48 hr. Thus, the 48-hr forecasts in 1971 were more accurate than the 24-hr forecasts made four years earlier.

The long period improvement illustrated in Fig. 10 can be traced to a series of system changes. In March 1968 several program errors were corrected, and deliberate "inflation" of the forecasts was eliminated (Klein and Lewis, 1970). In March 1970 the barotropic 500-mb and Reed 1000-mb models were replaced by the baroclinic primitive equation model of Shuman and Hovermale (1968) as input to the forecast equations. In addition, surface reports 6 hr later than previously utilized were introduced into the system. In March 1971 climatologically-determined limits were incorporated to curb extreme temperature forecasts (Klein et al., 1971). Another factor contributing to the encouraging trend of Fig. 10 has been the general increase in accuracy of numerical products at the National Meteorological Center (NMC) during the years. As the numerical models improve, statistical equations which utilize them as input, derived by the "perfect prog" approach used here, are bound to improve also.

Figure 11 shows mean absolute errors at the same 131 cities as a function of forecast projection for one 12-month period. For periods from 12 to 60 hr in advance the curve is based on max and min combined. Because extended forecasts were prepared only once a day (from 0000 GMT data), the 72- and 96-hr forecasts represent the max only, while the 84-hr forecast pertains only to the min. This factor probably explains the peculiar dip in the curve at 84 hr, for the min normally exhibits less variability than the max and therefore is easier to forecast. Except for this irregularity, the error of the automated forecast increases almost linearly with time, from 3.6F at 12 hr to 6.4F at 96 hr, at a rate of about 0.4F per 12 hr.

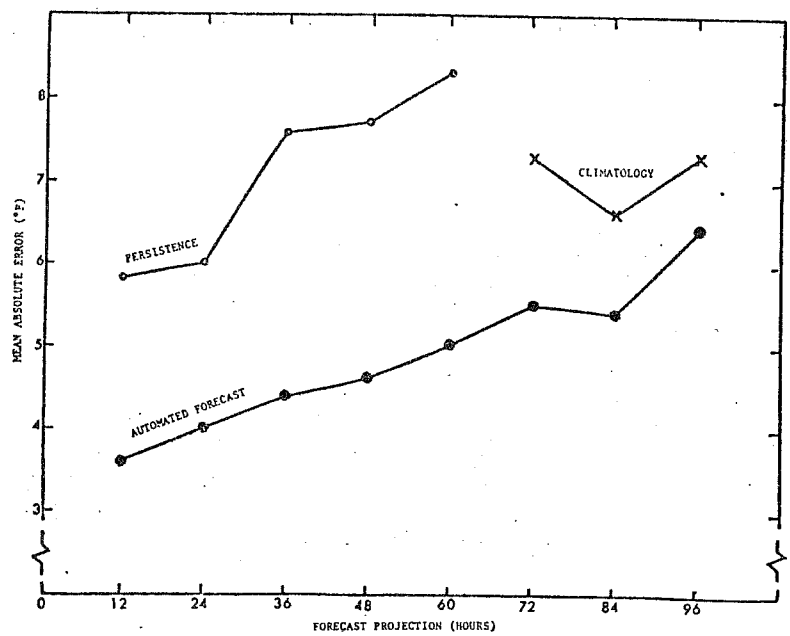

Figure 11. Mean absolute temperature errors (^{o}F) of persistence, climatology, and automated forecasts for 131 cities during the 12-month period from June 1971 through May 1972. Forecast projections are based on averages of maximum and minimum temperature for 12 through 60 hr, max only for 72 and 96 hr, and min only for 84 hr.

Two control curves are shown in Fig. 11. The first, labeled Persistence, gives the error that would be obtained by predicting that the latest observed max and min temperature will remain unchanged. It is illustrated only up to 60 hr because its error increases so rapidly with forecast projection. The second curve, labeled Climatology, indicates the error obtained by using the daily normal max and min temperature as a forecast. For convenience it is plotted only for the extended range, but it would have almost identical values for the shorter range. As such, it would provide a better forecast than persistence for projections of 36 hr and beyond, but persistence would be more accurate for less than 36 hr. It is noteworthy that the automated predictions are more skillful than either persistence or climatology for all forecast projections from 12 to 96 hr.

5.4 Improved predictors

An effort to improve the objective max/min forecast system described above (derived in 1968) was made by Klein and Marshall (1973) by screening additional predictors which were not available previously; namely, height at 850 and 500 mb, dewpoint depression at 850, 700, and 500 mb, and temperature at 1000, 850, 700, and 500 mb. Twice-daily analyzed values of these variables (plus 1000- and 700-mb height) for the period 1964 to 1969 were combined with surface data in a series of screening experiments to determine the predictability of temperature during the winter months (Dec., Jan., and Feb.).

First, max and min temperatures were screened as a function of upper air data observed at 1200 GMT on forecast day at 123 NMC grid points. Table 5 presents average values of RV for 5-term equations at 51 stations for five different groups of predictors. The first four lines compare the effectiveness of the field of geopotential height at four standard levels. For both max and min, the best height predictor of temperature is 500 mb and the second best is 1000 mb. Although these two levels have almost the same RV, their simple correlations with surface temperature are opposite in sign. Lows at 500 mb are usually cold at the ground, while highs are warm. However, at the 1000-mb level lows tend to be warm and highs cold. As a result, the intermediate level of 850 mb has the lowest RV of any height. The 700-mb level used in our 1968 system is also relatively poor, and it now seems, in retrospect, that better results might have been obtained with 500-mb height.

The second six lines of Table 5 show the effect of screening different thicknesses. All yield higher values of RV than height alone, but they differ considerably among themselves. The most effective thickness is 1000-500 mb for the max and 1000-700 mb for the min. These layers give better results than either 850-500 mb and 700-500 mb, which are apparently too far above the surface, or 1000-850 mb and 850-700 mb, which are probably too thin. On an overall basis, the 1000-700 mb thickness used in our 1968 system was probably as good a choice as any we could have made.

The next four lines give results of screening temperature at four mandatory levels. The two lowest levels (1000 and 850 mb) yield the highest RV's of Table 5, but the upper two levels (700 and 500 mb) are not as good as most of the thicknesses.

Table 5. Mean reductions of variance obtained by screening winter maximum and minimum temperatures as a function of single predictor fields observed at 1200 GMT on forecast day. Results are averaged for 5-term equations at 51 stations in the United States.

Predictors	RV (%) Today's Max	Tomorrow's Min
1000-mb height	49.3	47.6
850-mb height	43.8	41.3
700-mb height	45.6	44.1
500-mb height	49.8	48.2
1000-850 mb thickness	58.7	57.9
1000-700 mb thickness	65.6	63.4
1000-500 mb thickness	66.4	63.1
850-700 mb thickness	65.1	60.3
850-500 mb thickness	57.9	54.3
700-500 mb thickness	53.3	51.0
1000-mb temperature	74.5	73.9
850-mb temperature	72.1	68.9
700-mb temperature	59.8	55.7
500-mb temperature	49.5	49.4
850-mb dew point spread	34.7	33.4
700-mb dew point spread	26.4	24.3
500-mb dew point spread	19.9	19.8
850-700 mb mean spread	32.4	30.8
Mean relative humidity	27.2	--
Yesterday's maximum temp.	65.6	60.9
Yesterday's minimum temp.	57.4	66.7
Today's minimum temp	70.6	--

The next five lines are various measures of moisture which, hopefully, show the effect of cloud cover. These give the lowest values of RV in Table 5 since temperature and moisture are relatively independent. Best results are yielded by the lowest level tested (850 mb) and poorest by the highest level (500 mb). The two means (850-700 mb mean dewpoint spread and 1000-500 mb mean relative humidity) are not quite as good as the 850-mb spread.

The last three lines illustrate the importance of previous max and min. Note that the peak RV's in this group were exceeded in Table 5 only by the values for 1000- and 850-mb temperature.

Table 5 considers only single predictor fields; improved results can be obtained by screening several predictors concurrently. Ten-term multiple regression equations were therefore derived for both winter and spring by screening 12 predictor fields consisting of yesterday's max, today's min, and both 1200 and 2400 GMT values of the following five variables: 1000- and 850-mb temperatures, 1000- and 500-mb heights, and 850-700 mb mean dewpoint depression. These equations explained about 90 percent of the temperature variance, 10-15 percent more than the 1968 equations.

Although the above equations offer considerable potential for improvement over our 1968 system, this improvement can be realized only if the new equations are statistically stable, and the new grid point predictors can be forecast as accurately as the old ones. Thus, the real value of the 1973 system can be ascertained only through a test on independent data under operating conditions. The spring equations were therefore applied once a day (0000 GMT) to NMC prognostic data for the months of April, May, and June of 1972. Forecasts for 12 to 48 hours in advance were made in an iterative fashion from surface synoptic reports and grid point output of the PE model in accordance with the system outlined in Table 4. The forecasts were then verified in the customary way and compared to the current operational procedure.

Table 6 presents several verification statistics for forecasts of today's max, tonight's min, and tomorrow's max at 49 cities during the three-month test period. In terms of the simple correlation coefficient between forecast and observed temperature (line 1), our 1973 equations perform somewhat better than the 1968 equations. However, the mean absolute errors (line 2) of the 1973 system are larger than those of the 1968 system, particularly at the longer projections. One clue to the large errors of the 1973 equations is their sizeable bias or mean algebraic error (line 3), which is considerably greater than the bias of the 1968 system in the first two periods. Another explanation is their excessive variability, as illustrated by their large forecast standard deviations (line 4). These are not only greater than those of the 1968 system, but also larger than the standard deviations of the observed temperatures for both tonight's min (6.5) and tomorrow's max (7.8). This rather poor performance of the 1973 system can probably be attributed to defects in the PE prognostic data, particularly the low level temperatures and dewpoint depressions used as input to our equations.

Table 6. Verification of objective max/min temperature forecasts, averaged at 49 cities, made once a day (0000 GMT) for the period 1 April-30 June 1972 from NMC prognostic data by three different systems.

	1973	1968	MOS
a) Today's Max:			
Correlation between forecast and observed	.81	.78	.82
Mean absolute error (°F)	4.1	3.9	3.4
Mean algebraic error (°F)	1.5	0.8	0.3
Standard deviation of forecasts (°F)	7.7	6.5	6.9
b) Tonight's Min:			
Correlation between forecast and observed	.71	.71	.76
Mean absolute error (°F)	5.0	3.9	3.4
Mean algebraic error (°F)	2.7	-0.6	-0.3
Standard deviation of forecasts (°F)	6.8	5.1	5.2
c) Tomorrow's Max:			
Correlation between forecast and observed	.71	.70	.74
Mean absolute error (°F)	5.5	4.6	4.2
Mean algebraic error (°F)	-0.2	0.2	0.3
Standard deviation of forecasts (°F)	9.0	6.6	6.3

The last column of Table 6 presents verification statistics for experimental forecasts produced by applying Model Output Statistics (MOS) (Glahn and Lowry, 1972) to max/min temperatures. Multiple regression equations for this objective system were derived for the years 1970 and 1971 and then tested on the same 1972 cases and cities used to test our 1973 equations. The results show that the MOS forecasts had higher correlation coefficients, lower mean absolute errors, and smaller bias than the other two objective forecasts, while the MOS variability was about the same as that of the 1968 system. Details of the MOS system will be described in the next section.

6. Max/min forecasts based on MOS

6.1 MOS compared to perfect prog method

All temperature prediction techniques discussed up to this point have been applications of the "perfect prog" (PP) method. As illustrated in Fig. 12, this method utilizes observed historical data to specify local weather elements from concurrent (or nearly concurrent) weighted combinations of meteorological parameters. To use the derived equations for making a forecast, we apply them to the output of numerical prognostic models which simulate the observed circulation, as shown by the dashed arrow. Although errors in the numerical prediction will inevitably produce corresponding errors in the statistical forecast, the latter will improve each time the former is improved. An advantage of this method is that stable forecasting relations can be derived for individual locations and seasons from a long period of record. A disadvantage is that it takes no account of errors and uncertainties in the numerical model.

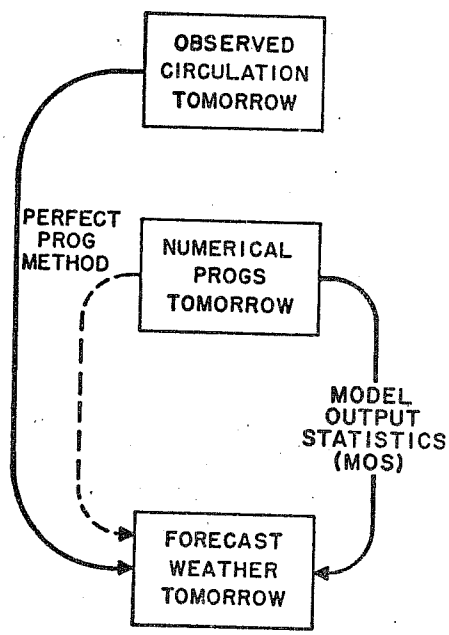

Figure 12. Two methods of combining numerical and statistical weather forecasting in schematic form.

In contrast, in Model Output Statistics (MOS), the predictor sample usually consists of a relatively short period of prognostic data produced by numerical models. In other words, the MOS method simply archives the output from numerical models and matches it with observations of local weather. In this way the bias and inaccuracy of the numerical model, as well as the local climatology, can be automatically built into the forecast system. Another characteristic of MOS is that it can include many predictors not readily available to the perfect prog method, such as vertical velocity, boundary layer wind and temperature, 3-dimensional air trajectories, etc.

6.2 Equations based on six-month seasons

Klein and Hammons (1975) described the application of the MOS

technique in deriving max/min temperature equations for a six-month season. Separate equations were developed for each of 228 stations in the conterminous United States, four forecast projections (approximately 24, 36, 48 and 60 hr after the initial model time), and both 0000 GMT and 1200 GMT cycles. Only data interpolated to a point over a station were used as potential predictors in that station's equations. The predictand was the station's calendar day maximum or minimum, depending on the particular projection.

For each projection an unique set of predictors was offered to the regression program in the derivation of the equations. In the six-month equations, we used 15 forecast fields from the trajectory model (Reap, 1972) and 31 forecast fields from the primitive equation (PE) model (Shuman and Hovermale, 1968). Not all of these fields were of different types; we often tried model output predictors that were valid at one or two times near the max (or min) valid time. Moreover, some predictors were filtered by a 5-point or 9-point space smoother, according to the particular parameter, level and projection. As potential predictors, we also included the sine and cosine of the day of the year to aid in capturing seasonal trends. Finally, station observations of surface conditions reported 6 hr after the initial cycle time (including the max/min for the previous day) were offered as predictors in the first projection (today's maximum from 0000 GMT or tonight's minimum from 1200 GMT).

Using this procedure, we derived equations for both the warm (April-September) and cool (October-March) seasons. From August 1973 to August 1975, NWS based its objective temperature guidance on these six-month equations. On the basis of mean absolute error, the MOS forecasts were equal or superior to the perfect prog forecasts (run as a control) for all projections, types and seasons, except for the 60 hr min during the cool season. The overall improvement with the MOS forecasts was 0.2-0.5°F at the 126 stations common to both systems.

Although the MOS verifications showed improvement over the perfect prog, the forecasts at some stations tended to deteriorate near the beginning and end of the six-month season. After some investigation, we concluded that the terms with sine and cosine of the day of the year could cause the forecasts at certain locations to vary by 20-30°F over the six-month season under similar atmospheric conditions.

6.3 Development of equations for three-month seasons

Because of the above problems, Hammons et al. (1976) divided the year into four three-month seasons; namely, spring (March to May), summer (June to August), fall (September to November) and winter (December to February). By 1975 we had archived sufficient data for a stable dependent sample. For the fall and winter equations we had six years of data (1969-75), providing over 410 cases per season. For spring and summer, five years of developmental data were available (1970-74); this generally meant that over 370 cases were used in deriving these seasonal equations. As with the six-month seasons, we developed MOS 10-term equations for both 0000 and 1200 GMT cycles, for each of the 228 max/min stations, and for four projections, valid approximately 24, 36, 48, and 60 hr after the initial cycle time. The

predictand in the development was the calendar day maximum or minimum, depending on the particular projection.

When we changed to three-month seasons, we also increased the number of predictors offered to the screening regression program for any one projection. Table 7 lists the potential predictors used in the 0000 GMT cycle. The list for the 1200 GMT cycle is identical except that 1) the projections are tomorrow's min, tomorrow's max, day after tomorrow's min, and day after tomorrow's max; 2) the surface synoptic observations are obtained from the 1800 (instead of the 0600) GMT reports; and 3) the valid times of the predictors are in hours after 1200 (instead of 0000) GMT.

Table 7 incorporates several changes made to our former six-month predictor list. For example, in the first and second projections we increased the number of valid times of the PE temperature and moisture forecasts. Many of the model predictors underwent additional space smoothing at all projections, while in the last projection some data were filtered by a 25-point space smoother. We also added surface observations as predictors in the second projection. Though we had earlier found that the effect of surface observations was relatively small in this particular forecast period, we later discovered that their importance depended strongly on the season of the year. Finally, we selected a number of new predictors, such as vorticity, stability and divergence, and dropped some of the older ones that were either unimportant in temperature prediction or redundant.

Figure 13 indicates the average standard error of estimate at 0000 GMT for all 228 stations for the four sets of seasonal equations. The smallest standard errors occurred during the summer, while winter had the largest errors. The errors during spring and fall were about the same in magnitude and intermediate between those of summer and winter. These results were expected since the standard errors tend to increase as the variability of the predictand increases from summer to winter.

Because of operational requirements, we developed backup equations for the first two projections that use model output but no surface observations. A plot of the standard errors in the first and second projections, both with and without surface observations (Fig. 13), indicates the importance of these observations as predictors. In the first projection during the winter season, the use of surface data reduced the average standard error by 0.5°F. This decrease was greater than any that we had previously obtained. For the second projection in winter, the inclusion of surface reports improved the standard error by 0.2°F. This change is about the same as that shown for the other three seasons during the first period. Use of surface observations in the second projection during spring, summer and fall reduced the standard error by only small amounts.

The shapes of the curves indicate a general decrease in skill with increasing projection, as one would expect; however, the standard errors for the 60 hr min were actually less than those for the 48 hr max in the spring and summer seasons. This reflects the difficulty of forecasting the max during the warm part of the year. In the warmer months, because of the presence of small-scale convective clouds, the

Table 7. Potential predictors of maximum and minimum surface temperatures for three-month MOS screening regression. Numbers indicate valid time of predictors in hours after 0000 GMT. Stars indicate the predictor was smoothed by 5 points (*), 9 points (**), or 25 points (***).

Predictor	Today's Max	Tomorrow's Min	Tomorrow's Max	Day After Tomorrow's Min
		a) PE Model		
850-mb height	12,24	24,36	36,48	48,48*
500-mb height	12,24	24,36	36,48	36*,48,48*
1000-500-mb thickness	12,24	24,36	36,48	48,48*
1000-850 mb thickness	12,24	24,36	36,48	48,48*
850-500-mb thickness	12,24	24,36	36,48	48,48*
1000-mb temperature	12,24,24*,36*	24*,36,36*,48*	36*,48,48*,48**	48*,48**,48***
850-mb temperature	12,24,24*,36*	24*,36,36*,48*	36*,48,48*,48**	48*,48**,48***
700-mb temperature	24	24*	24**	-
Boundary layer potential temp	12,24,24*,36*	24*,36*,48*	36*,48*,48**	48*,48**,48***
Boundary layer U wind	12,24*	24*,36*	36*,48*	48*,48**,48***
Boundary layer V wind	12,24*	24*,36*	36*,48*	48*,48**,48***
Boundary layer wind speed	24	36	48	48*,48**
850-mb U wind	24	24*	24**	24***
850-mb V wind	24	24*	24**	24***
700-mb U wind	24	24*	24**	24***
700-mb V wind	24	24*	24**	24***
1000-mb relative vorticity	24*	36*	48**	48***
850-mb relative vorticity	24*	36*	48*	48**
500-mb relative vorticity	24*	36*	48*	48**,48***
850-mb vertical velocity	24	24*	24**	-
650-mb vertical velocity	24	24*	24**	-
Stability (1000-700-mb temp)	24	24*	24**	-
Stability (850-500-mb temp)	24	24*	24**	-
400-1000 mean rel hum	12*,24*,36*	24*,36*,48*	36**,48**	48**,48***
Precipitable water	18*,30*	30*,42*	42*,42**	42**,42***
Boundary layer wind divergence	24*	36*	48*	48**,48***
		b) Trajectory Model		
Surface temperature	24,24*	24*,24**	24*,24**	24**,24***
850-mb temperature	24,24*	24*,24**	24*,24**	24**,24***
700-mb temperature	24	24*	24**	24**,24***
Surface dew point	24,24*	24*,24**	24*,24**	24***
850-mb dew point	24*	24*	24**	24***
700-mb dew point	24	24*	24**	24***
700 mb-surface mean rel hum	24	24*	24**	24***
850-mb 12-hr net vert displ	24	24*	24**	24***
850-mb 24-hr net vert displ	24	24*	24**	24***
700-mb 12-hr net vert displ	24	24*	24**	24***
700-mb 24-hr net vert displ	24	24*	24**	24***
Surface 12-hr horiz conv	24,24*	24*,24**	24*,24**	24***
850-mb 12-hr horiz conv	24	24*	24**	24***
George's K index	24	24*	24**	24***
		c) Other Variables		
Sine day of year	00	00	00	00
Cosine day of year	00	00	00	00
Sine of twice day	00	00	00	00
Cosine of twice day	00	00	00	00
Latest surface temperature	06	06	-	-
Latest surface dew point	06	06	-	-
Latest cloud cover	06	06	-	-
Latest surface U wind	06	06	-	-
Latest surface V wind	06	06	-	-
Latest surface wind speed	06	06	-	-
Latest ceiling	06	06	-	-
Previous maximum	06	-	-	-
Previous minimum	-	06	-	-

max is usually more variable than the min. In contrast, during the winter months the minimum seems harder to forecast because local effects, such as drainage winds and low-level cloudiness, tend to dominate night-

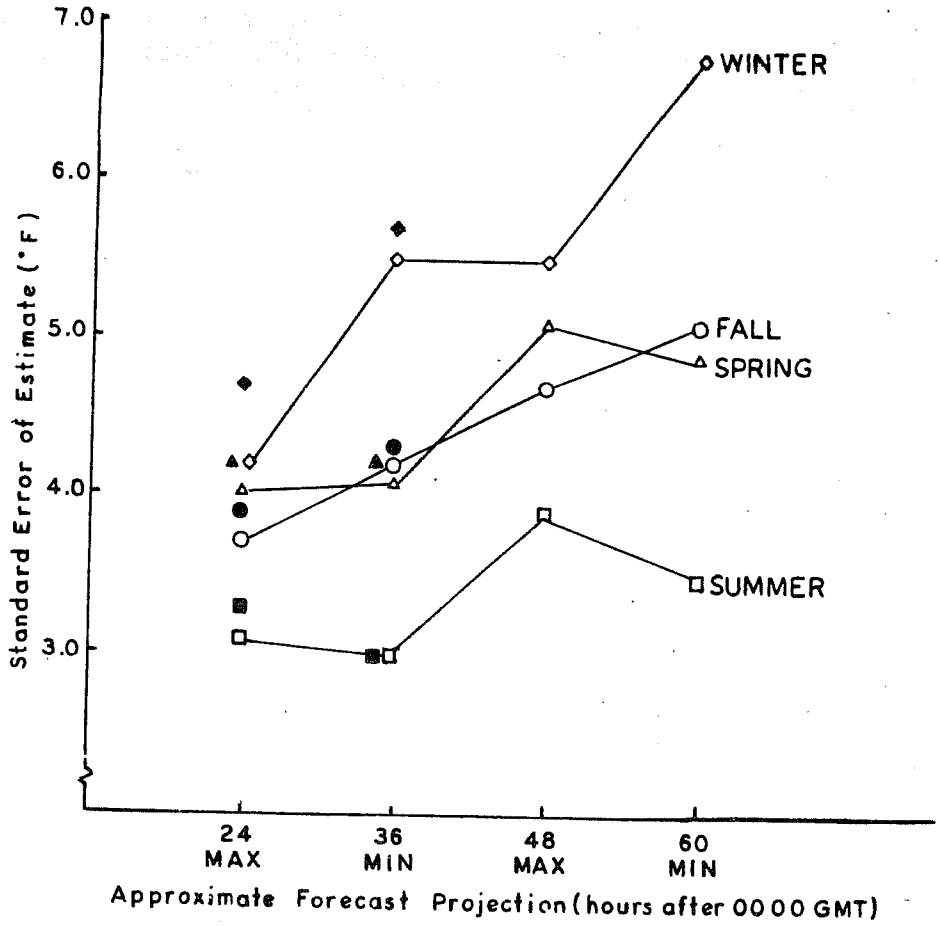

Figure 13. Standard errors of estimate averaged at 228 stations in the conterminous United States for the three-month MOS equations from the 0000 GMT cycle for spring, summer, fall and winter. The 24-and 36-hr backup equations do not use observations, but only model output as predictors. The standard errors for these equations are shown by the appropriate, solid geometrical figure.

time cooling, while the daytime max is more subject to large-scale synoptic conditions.

Table 8 presents the relative importance of the predictors in the 10-term equations for spring (0000 GMT cycle). This tabulation is based on both the frequency and order of selection. If a predictor was picked as the first term in an equation, it was assigned a value of ten points. If it was the second predictor in that equation, it was given nine points, and so on. Summing these points for the predictors in all the equations for one projection illustrates the most important types of predictors chosen in the MOS equations. The relative order varied, of course, from season to season, but the majority of the predictors shown in Table 8 were important in all seasons.

Generally, for both max and min forecasts, the model forecasts of low-level temperatures, boundary layer winds and mean relative humidity, as well as the sine and cosine terms, were frequently selected

Table 8. Importance of primitive equation (PE) and trajectory model (TM) predictors on basis of frequency and order of selection in 10-term equations for maximum and minimum spring (March-May) temperatures at 228 stations (0000 GMT data). Surface synoptic (SS) reports at 0600 GMT were included as predictors for today's maximum and tomorrow's minimum.

Rank	Today's maximum	Tomorrow's minimum	Tomorrow's maximum	Day after tomorrow's min
1.	SS previous max	PE 850 mb temperature	PE boundary layer pot temp	PE 850 mb temperature
2.	PE boundary layer pot temp	Cosine day of year	PE 850 mb temp	Cosine day of year
3.	PE 1000 mb temperature	PE precipitable water	Sine twice day of year	PE boundary layer *V* wind
4.	TM surface temperature	PE 1000 mb temperature	PE boundary layer *U* wind	PE 1000 mb relative vorticity
5.	PE mean relative humidity	PE boundary layer pot temp	PE mean relative humidity	PE boundary layer pot temp
6.	Sine twice day of year	SS latest temperature	Cosine day of year	PE 500-1000 mb thickness
7.	SS latest temperature	PE 500-1000 mb thickness	PE 500 mb height	PE 1000 mb temperature
8.	PE 850-1000 mb thickness	PE mean relative humidity	PE 1000 mb temperature	Sine twice day of year
9.	PE 850 mb temperature	SS previous min	TM surface temperature	PE mean relative humidity
10.	PE 500 mb height	TM surface dew point	PE boundary layer *V* wind	TM 850 mb convergence

predictors. For today's max, the previous day's max, the latest (0600 GMT) surface temperature, and some measure of the station cloudiness (either ceiling or cloud cover at 0600 GMT) were usually important predictors in all seasons. For tomorrow's min, surface observations were somewhat less important than in the first projection, although the latest observed temperature and the previous min were often picked in the spring and fall for the 36-hr forecast. Table 8 indicates that of the new predictors added to the three-month screenings, the sine twice day of the year, surface observations for the second projection, and the PE 1000 mb geostrophic relative vorticity were important factors in the spring equations.

For each season and projection, we summarized the three predictors that occurred most frequently in the leading three terms of the 0000 GMT equations (Table 9). In the forecast equations, the first three terms normally account for most of the reduction in variance of the temperature. For the 24 hr max, the previous max and the PE boundary layer potential temperature were extremely important in all four seasons. Similarly, the PE 850 mb temperature was a very important predictor in nearly all seasons for the 36 hr minimum, the 48 hr maximum and the 60 hr minimum. In addition, the PE precipitable water was frequently selected for forecasting the min, but not the max. The table reaffirms the importance of the latest surface temperature (0600 GMT) in the second projection during both winter and summer. This shows the 36-hr minimum forecast tends toward persistence in those seasons. That tendency was not as strong in the transitional (spring or fall) seasons. The sine and cosine terms were most important in spring and fall when they could account for the more pronounced seasonal temperature trend. Of the new predictors used in the three-month season development, the sine twice day of the year, the latest observed surface temperature (second projection), the PE 500-850 mb thickness, and the PE 1000 mb geostrophic relative vorticity were important in the first three terms of the equations.

Before we implemented our new three-month equations, we tested them on independent data for two weeks in July 1975 to have a comparison with the then-operational six-month warm season equations.

Table 9. List of predictors that are used most often in the first three terms of the 0000 GMT equations for the three-month seasons. The frequency is computed by summing over all 228 stations. The first three terms of the equations explain most of the forecast variance. (PE, primitive equation model output; TM, trajectory model output; SS, surface synoptic reports.)

Spring	Summer	Fall	Winter
24 h maximum	24 h maximum	24 h maximum	24 h maximum
SS obs max temp	SS obs max temp	SS obs max temp	SS obs max temp
PE boundary layer pot temp	PE 850 mb temp	PE boundary layer pot temp	PE 850–1000 mb thickness
PE 1000 mb temp	PE boundary layer pot temp	Cos day of year	PE boundary layer pot temp
36 h minimum	36 h minimum	36 h minimum	36 h minimum
Cos day of year	PE precip water	PE precip water	PE 850 mb temp
PE 850 mb temp	PE 850 mb temp	PE 850 mb temp	SS latest obs temp
PE precip water	SS latest obs temp	Cos day of year	PE boundary layer pot temp
48 h maximum	48 h maximum	48 h maximum	48 h maximum
Sin twice day of year	PE 850 mb temp	Cos day of year	PE 850 mb temp
PE boundary layer pot temp	PE boundary layer U wind	PE boundary layer pot temp	PE boundary layer pot temp
Cos day of year	PE boundary layer pot temp	PE 850 mb temp	TM surface temp
60 h minimum	60 h minimum	60 h minimum	60 h minimum
PE 850 mb temp	PE 500–1000 mb thickness	Cos day of year	PE 850 mb temp
Cos day of year	PE 500–850 mb thickness	PE 850 mb temp	PE 1000 mb rel vort
PE boundary layer pot temp	PE boundary layer V wind	PE precip water	PE boundary layer pot temp

The verification results for 126 stations for both sets of equations showed that the forecasts based on the three-month sample were better than the operational forecasts by about 0.2°F in mean absolute error. We began using the three-month equations to produce the operational temperature forecasts on 30 July 1975.

7. Verification statistics for max/min forecasts

In this section, I shall compare the accuracy of MOS and PP methods, MOS three-month and six-month equations, and MOS early and final guidance (Dallavalle et al., 1977). We shall also examine the distribution of forecast errors, the decline of skill with forecast projection, seasonal differences, and the long period trend in forecast accuracy.

7.1 Overall trends

Figure 14 shows the annual mean absolute errors of the objective temperature forecasts for the 24- and 48-hr projections, max and min combined for 126 cities, from 1968 to 1976. Note that the overall trend is for a decrease in the mean absolute errors. When PP forecasts were made, this improvement was probably caused by increased accuracy of the numerical models used as input to our regression equations and by improved operational procedures. After August 1973, the forecasts improved with the advent of MOS. By 1975 the 48-hr forecasts were roughly as good as the 24-hr forecasts made during 1970-72.

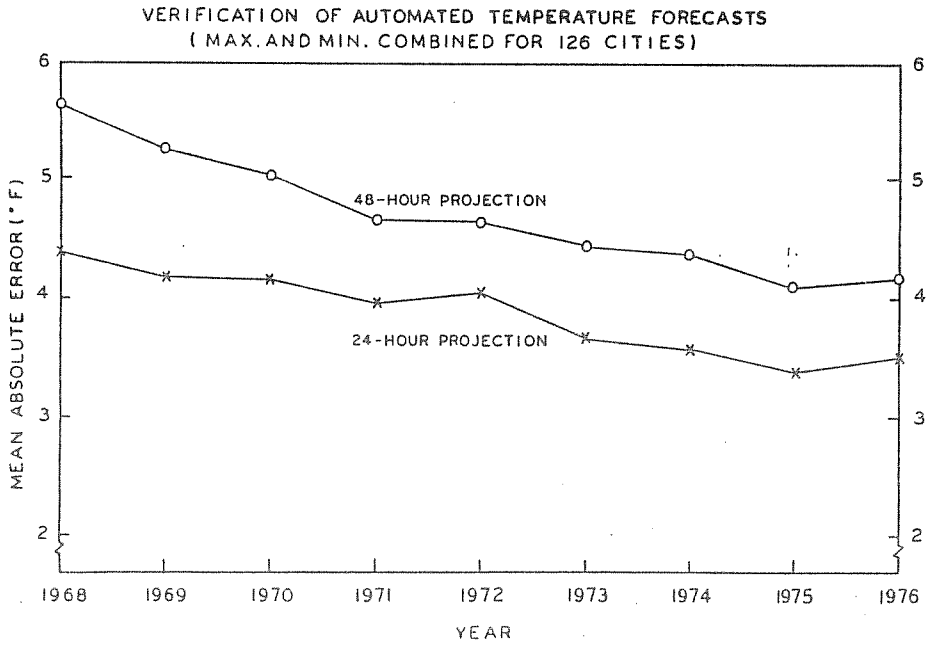

Figure 14. Mean absolute errors (^{o}F) of the automated max/min temperature forecasts for the period 1968-76. Errors are averaged for max and min combined at 126 cities over 12 months of each year.

7.2 Improvement of MOS over perfect prog forecasts

Another indication of improvement in temperature forecasting can be found by comparing mean absolute errors of the MOS and PP forecasts. We, therefore, routinely verify the forecasts for 126 stations common to both systems. The results are shown in Fig. 15 for the period from September 1973 through May 1977 for the 24- to 60-hr projections. It is clear that, for both the max and min, the MOS guidance has smaller mean absolute errors than the PP forecasts. For the min, the improvement with MOS is virtually the same at all projections except 60 hr, where the decrease in mean absolute error is only 0.2^{o}F. For the max, improvement of MOS over PP generally decreases from the first to the fourth projection. The mean error over all forecasts from 24 to 60 hr averages 0.4^{o} F less for MOS than for PP. In the graph for the max, note the dashed line for the 72-hr projection. MOS forecasts of the 72-hr max from 0000 GMT have been made since mid-December 1976 and verified since April 1977. The verification here represents only April and May 1977. Still the MOS improvement at 72 hr over PP continues the trend established in the earlier projections. Overall, the forecast accuracy has increased since we went to MOS equations in 1973.

7.3 Three-month versus six-month MOS equations

Table 10 and Fig. 16 compare the accuracy of the three-month and six-month MOS equations.

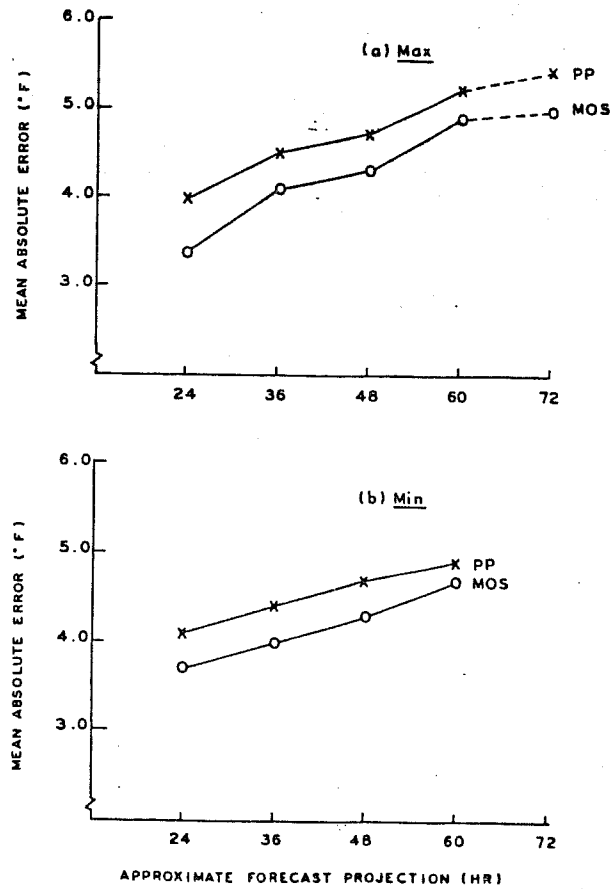

Figure 15. Mean absolute errors (^{o}F) of objective temperature forecasts of the max (a) and min (b) as a function of projection from September 1973 through May 1977. x indicates perfect prog forecasts, o indicates MOS forecasts. The 72-hr max forecasts were verified for April and May 1977 only.

Comparative verification (mean absolute errors) by season for both the max and min forecasts and the 24- through 60-hr projections is shown in Table 10. Both MOS and PP forecast verifications are divided according to whether MOS forecasts were based on six-month or three-month season equations. The six-month equations are verified from September 1973 through May 1975. The three-month equation verification covers the period September 1975 through May 1977. As expected, the mean absolute errors for winter are the largest, and those for summer the smallest, of all seasons. In any one season, the PP error for the same type and projection stays approximately constant from one two-year period to the next. In this respect, PP provides a good benchmark. Note that in the winter, when six-month MOS equations were used, the MOS forecasts of the minimum were less accurate at 48- and 60-hr than the PP forecasts. However, when we switched to three-month winter equations, the MOS forecasts of both min and max were better at all projections than PP. During the other three seasons, MOS was better than PP in all items, and the three-month MOS equations improved the forecast accuracy relative to PP more than the six-month equations did at all projections for the min.

In Fig. 16, we combined the max and min mean absolute errors for all four seasons as a function of the projection. Figure 16-a refers to six-month MOS equations and 16-b is for three-month MOS equations. Note again that the improvement in mean absolute error of the MOS forecasts over the PP guidance increased when we went to equations based on three-month seasons. Though some of this growth in skill was due to the additional predictors we screened for the three-month equations, the greater part was caused by the shorter seasonal stratification we used (Dallavalle and Hammons, 1977).

7.4 Other statistical properties of forecast errors

Figure 17 shows the error distribution by percentage for the PP and MOS forecasts from September 1975 (when three-month MOS equations began) through May 1977. Here all forecast projections and types have been combined. Both curves appear to be nearly normally distributed. Note that the MOS curve has a greater percentage of cases with errors close to 0.0°F than does PP, and a smaller percentage of large errors. In fact, only 1.4 percent of all MOS forecasts had errors in excess of ± 15°F. Both systems had mean algebraic errors of approximately 0.5°F during this period. The distributions for the Sept. 1973 through May 1975 period are not shown, but they are almost identical. The approximate normality of temperature forecast errors indicates that the mean absolute error is equal to about four-fifths the root mean square error and is a good statistic to use in comparing the relative accuracy of two different forecast systems.

Table 10. Mean absolute errors by season for automated max and min temperature forecasts for the 24- through 60-hr projections. Both perfect prog (PP) and MOS forecasts are shown. The verifications are for the period from Sept. 1973 through May 1975, when six-month MOS equations were used, and from Sept. 1975 through May 1977, when three-month MOS equations were operational.

Projection	Type	MOS	PP	MOS	PP
		a) Winter			
		6 Month		3 Month	
24 h	Min	4.4	4.8	4.5	5.0
36 h	Min	5.2	5.3	4.8	5.4
48 h	Min	5.4	5.3	5.3	5.6
60 h	Min	6.0	5.8	5.7	5.9
24 h	Max	3.8	4.5	3.9	4.5
36 h	Max	4.5	4.9	4.6	5.1
48 h	Max	4.9	5.4	5.0	5.4
60 h	Max	5.4	5.8	5.6	6.0
		b) Spring			
		6 Month		3 Month	
24 h	Min	3.7	4.1	3.6	4.2
36 h	Min	4.1	4.4	3.9	4.4
48 h	Min	4.3	4.8	4.3	4.9
60 h	Min	4.8	5.0	4.5	5.0
24 h	Max	3.7	4.2	3.5	4.2
36 h	Max	4.5	4.8	4.1	4.9
48 h	Max	4.6	5.0	4.4	4.9
60 h	Max	5.4	5.6	5.1	5.8
		c) Summer			
		6 Month		3 Month	
24 h	Min	2.8	3.0	2.6	3.1
36 h	Min	2.9	3.2	2.8	3.4
48 h	Min	3.2	3.4	3.2	3.5
60 h	Min	3.3	3.5	3.3	3.6
24 h	Max	2.7	3.2	2.8	3.3
36 h	Max	3.3	3.7	3.4	3.8
48 h	Max	3.4	3.8	3.6	3.9
60 h	Max	3.8	4.2	4.1	4.3
		d) Fall			
		6 Month		3 Month	
24 h	Min	3.9	4.1	3.5	4.1
36 h	Min	4.3	4.5	3.8	4.6
48 h	Min	4.6	4.7	4.2	4.7
60 h	Min	4.9	5.0	4.6	5.0
24 h	Max	3.3	3.9	3.3	3.8
36 h	Max	3.9	4.2	3.9	4.3
48 h	Max	4.2	4.5	4.2	4.5
60 h	Max	4.6	4.8	4.8	5.0

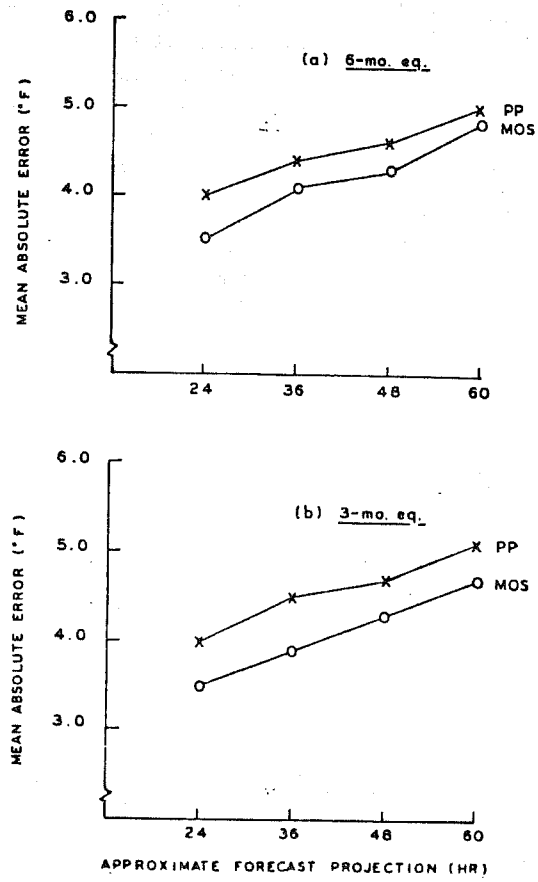

Figure 16. Mean absolute errors (°F) of the objective temperature forecasts of max and min combined as a function of projection. (a) is for the period from Sept. 1973 through May 1975 when six-month MOS equations were used. (b) is for Sept. 1975 through May 1977 when three-month MOS equations were operational. x indicates perfect prog forecasts; o indicates MOS forecasts.

The deterioration of forecast accuracy with projection is illustrated in Fig. 15. Note that the mean absolute errors of the MOS temperature forecasts, especially the min, increase approximately linearly with time. For the max, the 36-hr error (from the 1200 GMT forecast cycle) seems to be greater than expected by interpolating between the 24- and 48-hr forecasts (0000 GMT cycle), but no explanation of this diurnal difference is apparent.

On a seasonal basis (Table 10) the mean absolute errors of the forecasts increase least with time in summer and most during winter. Note that in summer for the same projection, the error for the max is always larger than that for the min. As discussed earlier, in the warmer months the max is harder to forecast, but during the colder months, the minimum is more difficult to predict.

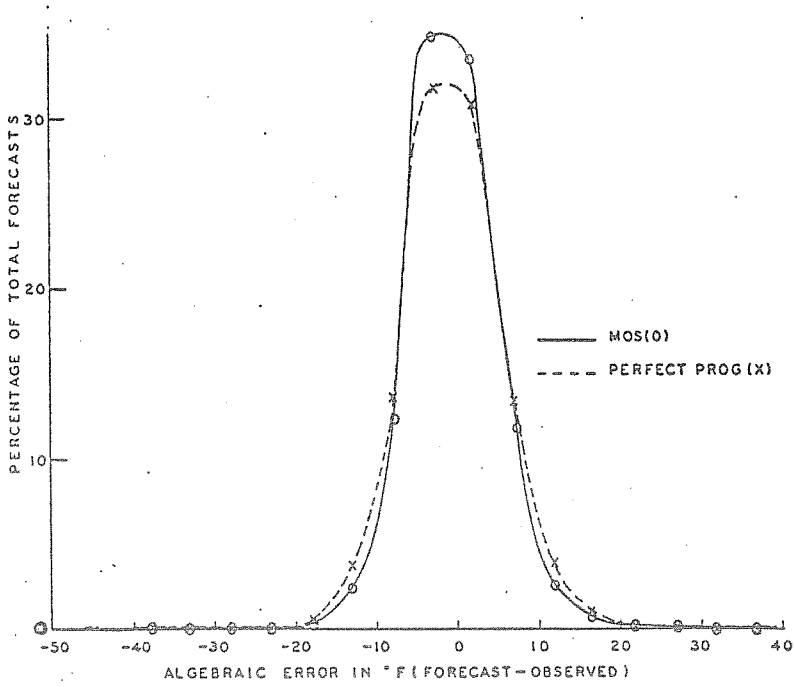

Figure 17. Error distribution, by percentage, of the MOS (solid line: O) and perfect prog (dashed line: X) temperature forecasts for the period Sept. 1975 through May 1977. For this distribution the class size is 5° F (0 to 4, 5 to 9, etc.), and the percentages are plotted at the midpoint of the class interval.

7.5 Early versus final MOS guidance

Until 1978, all our MOS max/min temperature forecast equations were derived from the output of the coarse-mesh, hemispheric, primitive equation (PE) model. However, it has recently become apparent that the PE model is not as accurate as the Limited-Area Fine Mesh (LFM) model (Howcroft and Desmaris, 1971), which is run about three hours earlier. Until recently we had not archived enough LFM data to derive max/min equations directly. We therefore produced temperature forecasts from the LFM indirectly--by simply substituting LFM fields in the PE-derived equations. No surface observations were used as predictors. The forecasts were transmitted from Suitland, Md., as "early" guidance during both 0000 and 1200 GMT cycles. Guidance was produced in this fashion from Sept. 1976 to May 1978. Comparative verification enables us to determine the effect of applying different numerical models as input to MOS equations.

Figure 18-a shows the mean absolute errors of both the early and final guidance as a function of projection for the max and min combined. These figures are verifications at 228 stations for the period of Oct. 1976 through May 1977. At every projection the final guidance is more accurate than the early guidance. At the last projection, in fact, the PE forecasts average 0.6° F lower in mean absolute error than the LFM forecasts. This is not surprising since the regression equations were developed to account for systematic PE model errors.

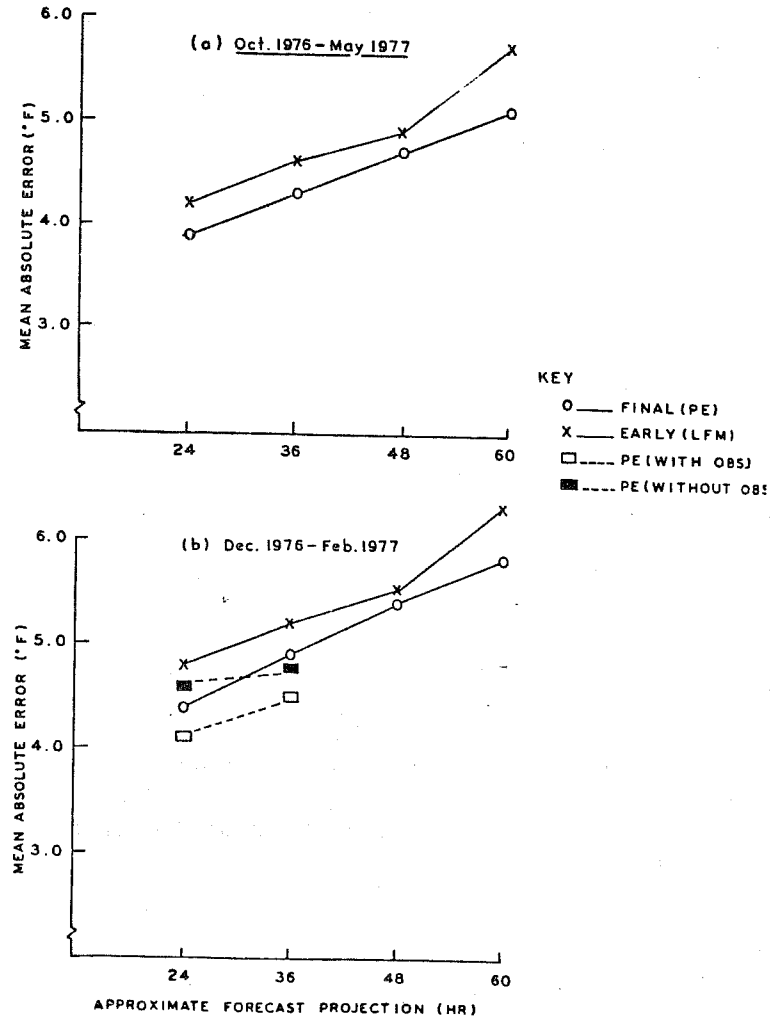

Figure 18. Mean absolute errors (°F) of the early (X) and final operational guidance (O) as a function of projection with max and min combined for: (a) Oct. 1976-May 1977; (b) Dec. 1976-Feb. 1977. The dashed lines give results of a test run for the first two projections where □ indicates PE forecasts with surface observations and ■ denotes PE forecasts without observations.

A comparison between the early and final guidance in the first two projections is complicated by the fact that we use surface observations as predictors in the 24- and 36-hr final forecasts. In a test (dashed) we re-ran the temperature guidance, both with and without observations, from the PE model for the Dec. 1976 through Feb. 1977 period. Thus, the only difference between the two types of forecasts was the inclusion of observations in the first and second projections. As Fig. 18-b indicates, the use of observations improved the temperature guidance by 0.5° F in the first projection and 0.3° F in the second. This is similar to what we had seen in the winter developmental sample (Hammons, et al., 1976). The actual operational early and final guidance, which was for a slightly different sample than the test, is also verified in Fig. 18-b for the winter season. Note that the difference between the early and final guidance in the first two projections was approximately the same as the difference in the test PE guidance run with and without observations. Thus, it appears that

in the first two projections for the winter season, the early LFM guidance, averaged over the 48 states, was about as accurate as the PE guidance would be without observations. However, in the final two projections, LFM input to the PE-based equations caused some deterioration in the temperature forecasts. Furthermore, the deterioration appears to increase with increasing projection.

In some of our earlier experiments (Dallavalle and Hammons, 1976) we found that LFM fields applied to the PE equations during late spring and early summer increased the mean absolute error of the temperature forecasts by about 0.2^{o} F in the first two projections. Since Fig. 18-b is valid only for winter, it may be that differences between the two sets of forecasts, particularly at the initial projections, are seasonally dependent. In the same report we also noted a definite geographical variation in forecast skill between the LFM and PE temperature guidance. Particularly in the southwestern United States, the early guidance seemed to deteriorate more than in other areas. On the other hand, in the northeast the early guidance was more accurate than the final in some instances.

7.6 Summary

The principal findings of this section may be summarized as follows:

1) The accuracy of the National Weather Service's automated max/min temperature guidance has gradually increased over the past decade.

2) The MOS temperature forecasts are generally more skillful than corresponding forecasts made by the perfect prog method.

3) The MOS temperature forecasts based on three-month seasons are more accurate than those based on six-month seasons.

4) Max/min temperature forecast errors are distributed in a nearly normal fashion.

5) Temperature forecast errors increase almost linearly with projection and are greater in winter than in summer.

6) The max temperature is harder to forecast than the min in summer; the opposite is true in winter.

7) Use of surface observations as predictors improves the MOS temperature guidance during the first two periods.

8) Introduction of LFM forecast fields into the PE temperature equations causes some deterioration in forecast accuracy. This loss of skill seems to increase with increasing projection and to vary both geographically and seasonally.

8. Max/min and three-hourly temperatures from LFM model

Carter et al. (1978) recently derived a new set of multiple regression equations for the early guidance temperature forecast system. The major improvements associated with this development include the use of archived output from the LFM model and observed weather elements from surface reports. Also, for the first time, equations to predict three-hourly surface temperatures (Grayson and Dallavalle, 1977) were derived simultaneously with those for max/min temperatures. This procedure should provide for greater consistency among the objective forecasts of these variables.

This particular application of the MOS technique involves matching surface temperature observations (predictand data) with various combinations of numerical model forecasts, observed weather elements, and climatic factors (predictor data). We use a forward selection procedure to derive linear regression equations to predict the surface temperature, $\hat{T}$. Each equation is of the following form:

$$\hat{T} = a_o + a_1 X_1 + a_2 X_2 + \ldots + a_k X_k \qquad (14)$$

where the caret indicates an estimate, the a_i's are multiple regression coefficients, and the X_i's are predictors selected by the screening procedure. Since we simultaneously derived equations to predict a particular max or min and certain specific three-hourly temperatures, the X_i's for several sets of equations are identical, but the a_i's are unique to each individual equation. The screening technique selects the predictor which yields the highest reduction of variance for any one of the predictands when combined with the other terms in a multiple regression equation. The same procedure is followed until 10 predictors have been selected.

8.1 Details of the derivation

The predictands for the max/min equations are maxima and minima reported for local calendar days. In contrast, the three-hourly temperature predictand data consist of surface temperature observations for specific times throughout each day (i.e., 0000, 0300, 0600, ..., 2100 GMT). Both max/min and three-hourly temperature observations are available in TDL's developmental data archive from October of 1972 through the present.

The potential predictors consist of various archived forecast fields from the LFM model, meteorological parameters derived from LFM output, weather elements observed as much as three hours after the LFM's normal input data times (i.e., reports taken at 0300 GMT and 1500 GMT), and several climatological factors. LFM forecasts for projections of 6, 12, 18, and 24 hours are available in the MOS data archive from October of 1972 through the present. Forecasts for longer projections of 30 and 36 hours are available from April 1975 to the present; 42- and 48-hr fields are available from February 1976 to the present.

Table 11 shows the LFM predictors and projections we actually screened for each forecast period. These included: heights,

Table 11. Potential numerical model predictors used to derive the MOS early guidance (LFM-based) temperature prediction equations for the 0000 GMT forecast cycle. The stars indicate the field was smoothed by 5 points (*), 9 points (**), or 25 points (***).

LFM OUTPUT FIELD	TODAY's MAX 3-HOURLY SET #1	TOMRW's MIN 3-HOURLY SET #2	TOMRW's MAX 3-HOURLY SET #3	DAY AFTER TOMRW's MIN
1000-MB HEIGHT	12*,24*	24*,30*,36*	36**,42**,48**	48**,48***
850-MB HEIGHT	12,24	24,30,36	36,42,48	48*,48**
500-MB HEIGHT	12,24	24,30,36	36,42,48	36*,48*
500-1000 MB THICKNESS	0,6,12,18,24	24,30,36	36,42,48	48*
850-1000 MB THICKNESS	0,6,12,18,24	24,30,36	36,42,48*	48*,48**
500-850 MB THICKNESS	0,6,12,18,24	24,30,36	36,42,48*	48*
1000-MB TEMP	0,12*,24*	0,24*,36*	36**,48**	48**,48***
850-MB TEMP	0,6,12,18,24	0,24*,30*,36*	36*,42*,48*	48*,48**
700-MB TEMP	0,12,24	24,30,36	36*,42*,48*	48*,48**
BND LYR POT TEMP	6,12,18,24	24*,30*,36*	36*,42*,48*,48**	48*,48**
BND LYR U	6,12,18*,24*	24*,30*,36*	36*,42*,48*	48*,48**
BND LYR V	6,12,18*,24*	24*,30*,36*	36*,42*,48*	48*,48**
BND LYR WND SPD	6,12,18*,24*	24*,30*,36*	36*,42*,48*	48*,48**
850-MB U	6,12,18*,24*	24*,30*,36*	36*,42*,48*	48**
850-MB V	6,12,18*,24*	24*,30*,36*	36*,42*,48*	48**
700-MB U	12,24*	24*,36*	36*,48*	48**
700-MB V	12,24*	24*,36*	36*,48*	48**
850-MB REL VORT	6*,12*,18*,24*	30**,36**	42**,48**	48**
500-MB REL VORT	12*,24*	30**,36**	42**,48**	48**
850-MB VERT VEL	12*,24*	36*	48**	48***
700-MB VERT VEL	12*,24*	30*,36*	42**,48**	48***
700-1000 MB TEMP DIF	12,24	36*	48*	48**
500-850 MB TEMP DIF	12,24	30*,36*	42*,48*	48**
BND LYR REL HUM	0*,6*,12*,18*,24*	24*,30*,36*	36**,42**,48**	48***
MEAN REL HUM	6*,12*,18*,24*	24*,30*,36*	36**,42**,48**	48***
PRECIP WATER	6*,12*,18*,24*	30*,36*	42**,48**	48***
1000-MB DEW PT	6*,12*,18*,24*	30*,36*	42*,48*	48**,48***
850-MB DEW PT	12*,24*	30*,36*	42*,48*	48**
700-MB DEW PT	12*,24*	30*,36*	42*,48*	48**
BND LYR WND DIVG	6*,12*,18*,24*	30*,36*	42**,48**	48***
850-MB TEMP ADV	12*,24*	30*,36*	42**,48**	48***
500-MB VORT ADV	12*,24*	30*,36*	42**,48***	48***

temperatures, potential temperatures, horizontal wind components, vertical wind velocities, relative humidities, dew points, and precipitable water at various projections and levels throughout the lower troposphere. In addition, we computed horizontal wind speed and divergence, relative vorticity, temperature and vorticity advection, thickness, and stability (i.e., the temperature difference between two levels). As indicated in Table 11, some of these fields were space-smoothed over 5, 9, or 25 model grid points in order to reduce the amount of small scale noise inherent in the numerical output. The LFM forecasts were then interpolated from grid-points to the location of each of the stations in the predictand data sample.

As shown in Table 12, we also screened surface temperature, dew point, wind, cloud amount, ceiling height, snow cover (during the cool season), and max and min temperatures for the previous calendar day from synoptic and hourly observations. We used only the previous min for the 0000 GMT cycle equations because, in day-to-day operations, the max for the previous calendar day is not available from the 0000 GMT synoptic reports. Analogously, since the calendar day min is not available operationally at 1200 GMT, we screened only the previous max for the 1200 GMT cycle equations.

Table 12. Potential observed predictors used to derive the early guidance temperature prediction equations.

ELEMENT	0000 GMT CYCLE	1200 GMT CYCLE
SFC TEMPERATURE	0300 0000 2100 (YESTERDAY)	1500 1200
SFC DEW POINT TEMP	0300	1500
CLOUD COVER	0300	1500
SFC U WIND	0300	1500
SFC V WIND	0300	1500
SFC WIND SPEED	0300	1500
CEILING HEIGHT	0300	1500
PREVIOUS MAXIMUM TEMP		1200
PREVIOUS MINIMUM TEMP	0000	
SNOW COVER	1200 (YESTERDAY)	1200

Additionally, we screened the first and second harmonics of the day of the year as potential predictors for all the equations. This was done in an attempt to capture the normal seasonal trend of temperature.

We derived the max/min and three-hourly temperature prediction equations for a number of different forecast projections. The 0000 GMT cycle max/min equations are for projections of approximately 24 (today's max), 36 (tomorrow's min), 48 (tomorrow's max), and 60 (day after tomorrow's min) hours from 0000 GMT. Analogously, the 1200 GMT cycle max/min projections are approximately 24 (tomorrow's min), 36 (tomorrow's max), 48 (day after tomorrow's min), and 60 (day after tomorrow's max) hours from 1200 GMT. The three-hourly temperature projections for both forecast cycles are valid every three hours from 6 through 51-h inclusive.

Figure 19 shows the combinations of forecast projections we used to develop the max/min and three-hourly equations. The three basic sets of equations are:

1) max/min equations for the first (24-hr) period and three-hourly equations for projections of 6, 9, 12, 15, 18, 21, 24, and 27 hours.

2) max/min equations for the second (36-hr) period and three-hourly equations for projections of 27, 30, 33, 36, and 39 hours.

3) max/min equations for the third (48-hr) period and three-hourly equations for projections of 39, 42, 45, 48, and 51 hours.

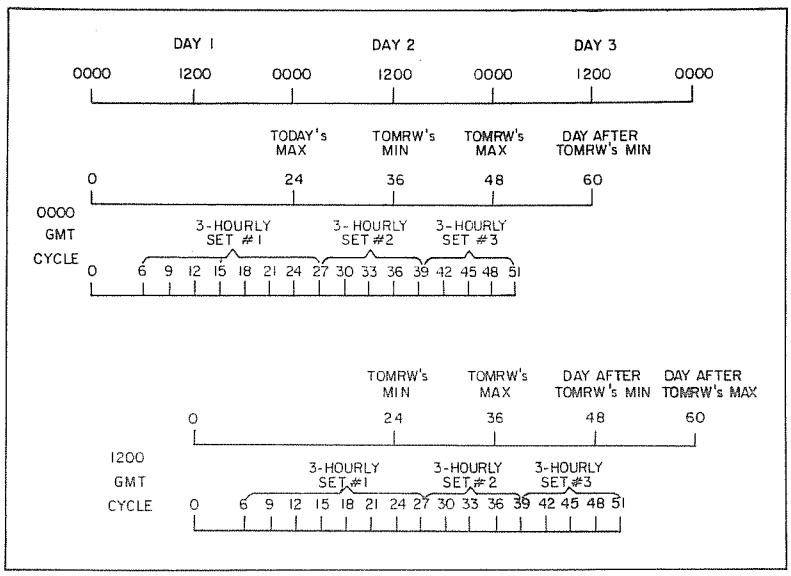

Figure 19. Forecast periods associated with the early guidance temperature prediction equations.

We also derived an additional set of 60-hr max/min equations separately.

We derived the first set of 24-hr max/min and 6- to 27-hr three-hourly equations by stratifying the data into three-month seasons, similar to the approach used by Hammons et al. (1976). However, due to the scarcity of developmental data from the LFM model for projections beyond 24 hours, we used a six-month stratification for the other sets of prediction equations (longer projections).

8.2 Screening regression results

We derived separate sets of 10-term regression equations for each station and forecast cycle. We screened observed weather elements as potential predictors--in addition to the forecast output from the LFM model and the first and second harmonics of the day of the year--when we developed the first two sets of max/min and three-hourly temperature equations shown in Fig. 19. We also derived "backup" equations for these two sets by not including surface observations as potential predictors. The backup equations are used to generate forecasts in day-to-day operations when surface observations are missing and the "primary" equations cannot be applied.

Table 13 shows the predictors selected for the 0000 GMT cycle 24-hr max and 6- to 27-hr three-hourly temperature prediction equations valid during April through June for Washington, D.C. All the equations contain the same 10 predictors, but the individual regression coefficients and constants differ. For Washington, the surface temperature observed at 0300 GMT and the 24-hr LFM forecast of the 1000-mb temperature are the most important predictors. Together, these two variables account for 83 percent to 96 percent of the reduction in variance of the various surface temperatures. Seven other fields from the LFM model and the cosine of the day of the

Table 13. Predictors used in the equations to forecast today's calendar maximum and three-hourly temperatures valid 6 to 27 hours from 0000 GMT for Washington, D.C. (DCA). Cumulative reductions of variance and standard errors of estimate are also shown for each equation. The developmental data were from April, May, and June of 1973-77 (394 days).

PREDICTOR	PROJECTION (HOURS)	SMOOTHING (POINTS)	CUMULATIVE REDUCTION OF VARIANCE								
			MAX	3-HOURLY TEMPERATURES							
			24	6	9	12	15	18	21	24	27
Obs 0300 GMT sfc temp	--	--	0.633	0.946	0.896	0.888	0.754	0.638	0.551	0.592	0.636
LFM 1000-mb temp	24	5	0.865	0.946	0.896	0.903	0.885	0.861	0.833	0.839	0.829
LFM 850-mb dew point	12	5	0.865	0.962	0.945	0.935	0.885	0.861	0.834	0.840	0.843
LFM 700-1000 mb temp	24	0	0.866	0.962	0.945	0.935	0.886	0.862	0.835	0.847	0.869
LFM mean rel humidity	18	5	0.884	0.963	0.945	0.937	0.895	0.879	0.860	0.865	0.871
LFM bound layer U component	18	5	0.899	0.963	0.945	0.937	0.899	0.888	0.868	0.870	0.874
LFM 1000-mb dew point	18	5	0.904	0.963	0.945	0.942	0.909	0.895	0.870	0.871	0.874
LFM bound layer pot temp	6	0	0.905	0.964	0.946	0.942	0.909	0.895	0.871	0.876	0.884
Cosine day of the year	--	--	0.906	0.965	0.948	0.952	0.912	0.895	0.873	0.876	0.884
LFM 500-mb vorticity adv	24	5	0.908	0.965	0.948	0.952	0.913	0.897	0.878	0.880	0.889
STANDARD ERROR OF ESTIMATE (°F)			3.30	1.93	2.38	2.33	3.23	3.73	4.01	3.75	3.40

year add from 2 percent to 6 percent more to the cumulative reductions of variance. Equations for the other stations have similar characteristics.

Nearly all the potential predictors we offered were selected by the screening regression procedure for one station or another. However, these selections were not uniformly distributed and a few predictors predominated. Table 14 lists the five most important predictors for all cities combined. This tabulation is based on both the frequency and order of selection. For the primary equations, observed temperatures and low-level temperature, dew point, and thickness forecasts from the LFM model predominated. For the backup equations, the LFM surface temperature analyzed from initial data was also an important predictor. In essence, this field served as a substitute for the station's surface temperature observation. The cosine of the day of the year was also frequently selected for the longer projections.

Figure 20 shows the average standard errors of estimate for all stations combined for the warm season max/min equations. The standard errors for the 24-hr max from 0000 GMT and the 24-hr min from 1200 GMT were obtained by averaging the values for the spring (April-June) and summer (July-September) seasons. The errors were smaller in summer than in spring and increased almost linearly with projection.

The comparative magnitudes of the errors indicate that the max is more difficult to predict than the min during the warmer part of the year. Fig. 20 also shows that the use of observed predictors

Table 14. Important predictors for the new early guidance surface temperature prediction equations for the warm season (0000 GMT cycle, 230 stations).

24-HR MAX (PRIMARY)	36-HR MIN (PRIMARY)	48-HR MAX
Obs sfc temp	LFM 850-mb temp	LFM 850-1000 mb thickness
LFM 850-mb dew point	LFM 500-1000 mb thickness	LFM mean rel humidity
LFM 850-mb temp	Cosine day of year	Cosine day of year
LFM 850-1000 mb thickness	LFM 1000-mb temp	LFM 850-mb temp
LFM 1000-mb temp	LFM 850-1000 mb thickness	LFM 500-1000 mb thickness
24-HR MAX (BACKUP)	**36-HR MIN (BACKUP)**	**60-HR MIN**
LFM 850-1000 mb thickness	LFM 850-mb temp	LFM 500-1000 mb thickness
LFM 850-mb temp	Cosine day of year	Cosine day of year
LFM 850-mb dew point	LFM 850-1000 mb thickness	LFM 1000-mb dew point
LFM 1000-mb temp	LFM 1000-mb temp	LFM 700-mb temp
LFM 500-1000 mb thickness	LFM 850-mb dew point	LFM 850-mb temp

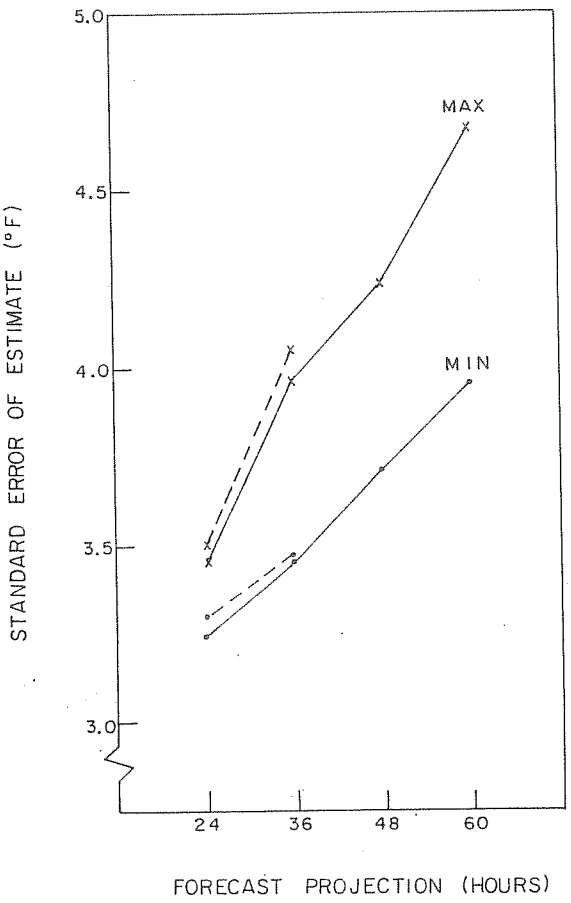

Figure 20. Developmental standard errors of estimate averaged at approximately 230 stations for the early guidance max and min temperature equations valid during the warm season. The standard errors for the primary equations are indicated by a solid line; those for the backup equations are indicated by a dashed line.

by the primary equations reduced the standard errors by about 0.1° F on the average. These findings are consistent with those of Hammons et al. (1976).

The comparable developmental standard errors of estimate for the three-hourly temperature equations are given in Fig. 21. Once again, we averaged the spring and summer values for the first set of equations (i.e., projections of 6 to 27 hours). Unlike the previous figure, these results apply to the 0000 GMT cycle equations only. We also averaged the developmental standard errors for the two sets of equations valid for projections of 27 and 39 hours from 0000 GMT.

The standard errors in Fig. 21 generally increase with increasing forecast projection, although not monotonically. The curves peak at 2100 GMT each day. As was noted for the max, this characteristic appears to be related to the presence of small-scale convective cloudiness in the mid to late afternoon. Also of interest is the manner in which the observed predictors substantially improve the standard errors of the primary equations for projections out to about 15 hours.

The standard errors for the 1200 GMT cycle three-hourly temperature equations (not shown) are generally quite similar to those in Fig. 21, except the patterns are off-set by 12 hours.

9. Operational aspects

Both max/min and three-hourly temperature forecasts generated from the new equations have been available for use as guidance by NWS forecasters since June 1, 1978. The forecasts for approximately 230 stations are distributed twice daily through the Federal Aviation

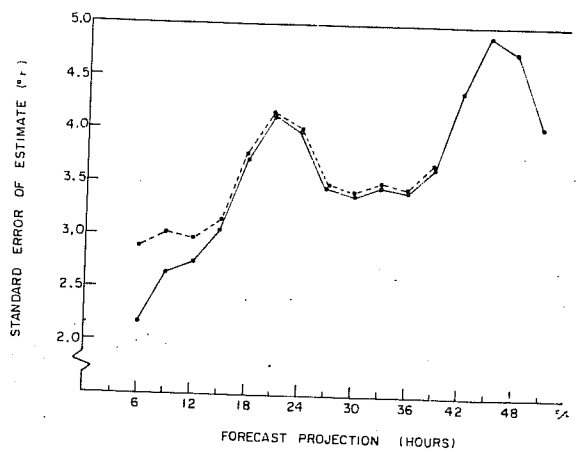

Figure 21. Developmental standard errors of estimate averaged at approximately 230 stations for the 0000 GMT cycle early guidance three-hourly temperature equations valid during the warm season. The standard errors for the primary equations are indicated by a solid line; those for the backup equations are indicated by a dashed line.

Administration's Weather Message Switching Center at Kansas City. Calendar day max/min forecasts for the first four periods (projections of approximately 24, 36, 48, and 60 hours) are now presented in early teletypewriter messages along with three-hourly forecasts for projections of 6 through 51 hours. A final bulletin gives max/min forecasts for projections of 24 to 60 hours at 228 MOS stations and 72-hr perfect prog forecasts at 126 stations. The early messages are transmitted about 4 hours after model run time, the final bulletin about 8 hours later.

The automated guidance max/min temperature forecasts for 24- to 60-hr projections also appear on a computer-produced four-panel chart available on several facsimile circuits. An example of the chart is shown in Fig. 22. The computer-drawn isotherms for 10-degree Fahrenheit intervals are based on MOS forecasts at 228 stations and perfect prog forecasts at 16 stations. The 32 degree isotherm is indicated by a dashed line. Due to a lack of space, forecasts are plotted on the chart for only 151 stations (135 MOS stations and 16 perfect prog stations). Occasionally, because the isotherms are smoothed by an objective analysis scheme, the contour values may not agree with the forecasts plotted on the map. In such cases, the individual station values represent the correct forecasts.

The regression equations are strongly dependent on the accuracy of the numerical model predictions used as input. When the field forecaster has good reason to believe the model prediction is in error, he should modify the automated temperature forecast accordingly. For example, if a trough or front has intensified or accelerated, corresponding changes to the forecast temperature patterns should be considered. Specific localized conditions and mesoscale features should also be taken into account by the local forecaster.

The early and final guidance temperature forecasts are currently being generated in day-to-day operations from new finer-scale versions of the LFM and PE models, the LFM-II and 7LPE, respectively. The LFM model was changed by NMC in Sept. 1977; the PE model in Jan. 1978. Preliminary tests indicate that the adverse impact of the model changes is minimal, and the early guidance is now of the same quality, or occasionally better than, the final guidance. However, field forecasters occasionally will need to adjust for any irregularities or unusual characteristics in the automated temperature guidance.

Acknowledgment

I would like to thank the many dedicated employees of NWS who have collaborated with me on various phases of temperature research during the past twenty years. In particular, the following individuals have been responsible for much of the material presented in this lecture: Billy Lewis, Isadore Enger, Curtis Crockett, Frank Lewis, George Casely, Frederick Marshall, Gordon Hammons, Paul Dallavalle, Gary Carter, and Albert Forst.

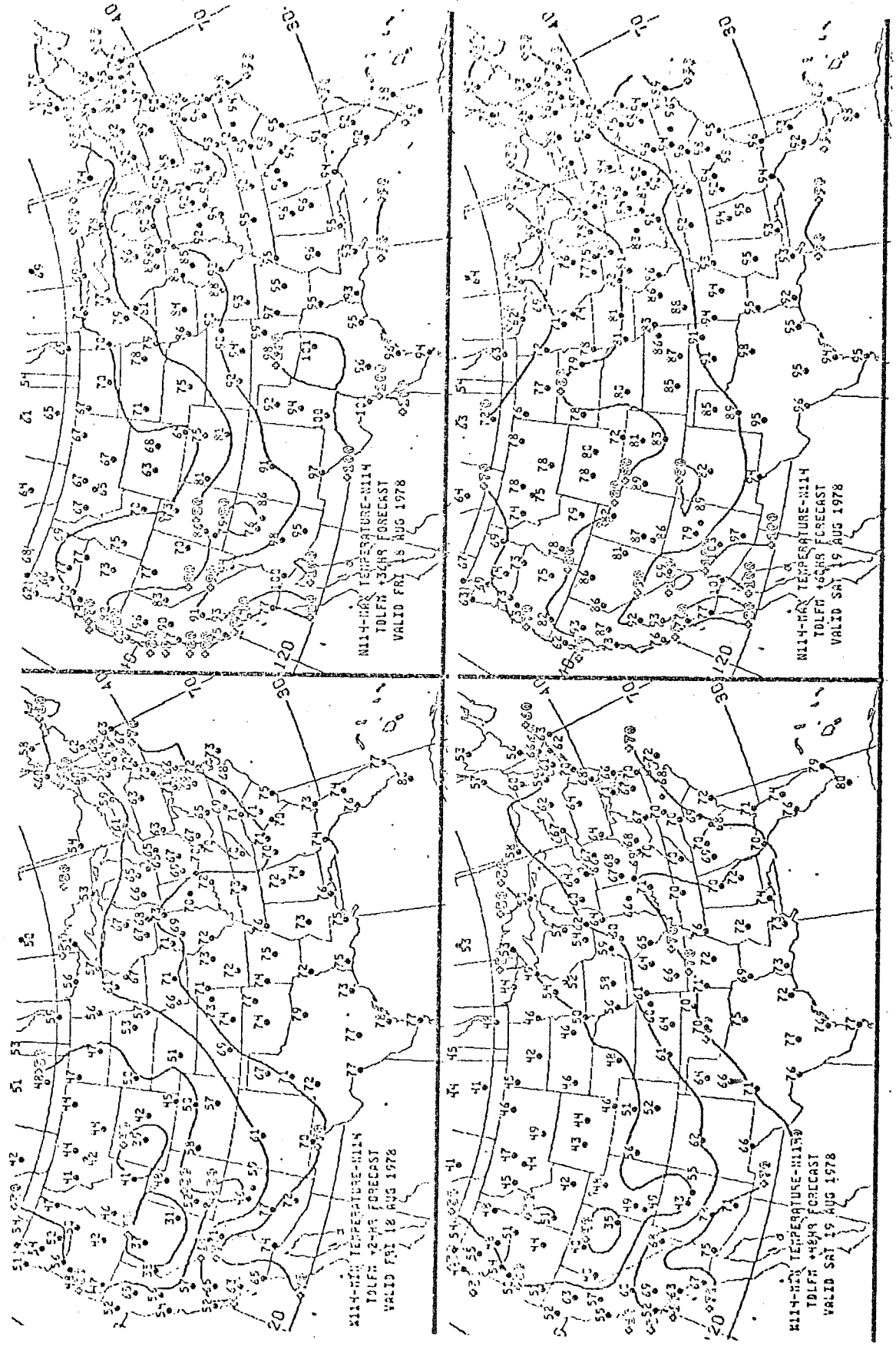

Figure 22. An example of the computer-produced four-panel max/min temperature chart.

References

Carter, G. M., A. L. Forst, W. H. Klein, and J. P. Dallavalle, 1978: Improved automated forecasts of maximum/minimum and 3-hourly temperatures. Preprints, Conference on Weather Forecasting and Analysis and Aviation Meteorology, Silver Spring, Md., Oct. 1978, 8 pp.

Dallavalle, J. P. and G. A. Hammons, 1976: Use of LFM data in PE-based max/min forecast equations. TDL Office Note 76-14, 10 pp.

_____, and _____, 1977: Tests of various predictor lists used in development of maximum/minimum forecast equations. TDL Office Note 77-19, 16 pp.

_____, W. H. Klein, and G. A. Hammons, 1977: Verification of the National Weather Service's objective maximum/minimum temperature guidance. Preprints, Fifth Conference on Probability and Statistics, November 1977, Las Vegas, Nev., Amer. Meteor. Soc., 347-352.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11, 1203-1211.

Grayson, T. H., and J. P. Dallavalle, 1977: Development and operational use of 3-hr objective temperature forecasts. Preprints, Fifth Conference on Probability and Statistics, November 1977, Las Vegas, Nev., Amer. Meteor. Soc., 53-56.

Hammons, G. A., J. P. Dallavalle, and W. H. Klein, 1976: Automated temperature guidance based on three-month seasons. Mon. Wea. Rev., 104, 1557-1564.

Howcroft, J., and A. Desmaris, 1971: The limited-area fine mesh (LFM) model. Tech. Procedures Bull. No. 67, National Weather Service, 11 pp.

Klein, W. H., 1962: Specification of monthly mean surface temperature from 700-mb heights. J. Appl. Meteor., 1, 154-156.

_____, 1965: Application of synoptic climatology and short-range numerical prediction to five-day forecasting. U.S. Weather Bureau, Research Paper No. 46, Washington, D.C. 109 pp.

_____, 1966: Objective forecasts of surface temperature from one to three days in advance. J. Appl. Meteor., 5, 137-147.

_____, 1972: On the accuracy of automated max/min temperature forecasts. J. Appl. Meteor., 11, 1381-1384.

_____, and F. Lewis, 1970: Computer forecasts of maximum and minimum temperatures. J. Appl. Meteor., 9, 350-359.

_____, and F. Marshall, 1973: Screening improved predictors for automated max/min temperature forecasting. Preprints, 3rd Conf. on Probability and Statistics in the Atmospheric Sciences, Amer. Meteor. Soc., Boston, Mass., 36-43.

Klein, W. H., and G. A. Hammons, 1975: Maximum/minimum temperature forecasts based on model output statistics. Mon. Wea. Rev., 103, 796-806.

_____, B. M. Lewis, and I. Enger, 1959: Objective prediction of five-day mean temperature during winter. J. Meteor., 16, 672-682.

_____, _____, C. W. Crockett, and I. Enger, 1960: Application of numerical prognostic heights to surface temperature forecasts. Tellus, 12, 378-392.

_____, _____, and _____, 1962: Objective forecasts of daily and mean surface temperature. Mon. Wea. Rev., 90, 11-17.

_____, F. Lewis, and G. P. Casely, 1967: Automated nationwide forecasts of maximum and minimum temperature. J. Appl. Meteor., 6, 216-228.

_____, _____, and G. A. Hammons, 1971: Recent developments in automated max/min temperature forecasting. J. Appl. Meteor., 10, 916-920.

Miller, R. G., 1958: A statistical procedure for screening predictors in multiple regression. Final Report, Contract No. AF19(604)-1590, The Travelers Weather Research Center, Inc., Hartford, Conn., 238 pp.

Reap, R. M., 1972: An operational three-dimensional trajectory model. J. Appl. Meteor., 11, 1193-1202.

Shuman, F. G., and J. B. Hovermale, 1968: An operational six-layer primitive equation model. J. Appl. Meteor., 7, 525-547.

Stidd, C. K., 1954: The use of correlation fields in relating precipitation to circulation. J. Meteor., 11, 202-213.