U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
OFFICE OF SYSTEMS DEVELOPMENT
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 85-7

DEVELOPMENT OF MEDIUM RANGE FORECAST EQUATIONS FOR MAXIMUM/ MINIMUM TEMPERATURE AND PROBABILITY OF PRECIPITATION

Joseph M. Palko, John S. Jensenius, Jr., and J. Paul Dallavalle

April 1985

DEVELOPMENT OF MEDIUM RANGE FORECAST EQUATIONS FOR MAXIMUM/ MINIMUM TEMPERATURE AND PROBABILITY OF PRECIPITATION

Joseph M. Palko, John S. Jensenius, Jr., and J. Paul Dallavalle

1. INTRODUCTION

Since November 1982, the Techniques Development Laboratory has been generating forecasts of the calendar day maximum/minimum temperature (max/min) for projections of approximately 84 through 192 hours after 0000 GMT. The guidance is based on an application of the perfect prog approach (Klein and Lewis, 1970) and is dependent on the 0000 GMT run of the Global Spectral Model (Sela, 1980). The forecasts are available for 131 stations in the conterminous United States. This guidance is not issued directly to National Weather Service field forecasters; rather, the max/min predictions are used by forecasters in the Medium Range Forecast Group (MRFG) of the National Meteorological Center to assist in preparing the 3- to 5-day forecast package. As Hughes (1984) indicated, the MRFG is able to improve consistently over the objective perfect prog forecasts.

Three factors have now convinced us to change our approach to the medium range forecasting task. First, we have an adequate sample (over 4 years) of Global Spectral Model (GSM) output to derive Model Output Statistics (MOS) equations. Secondly, we think that we can produce more accurate temperature guidance. Finally, we want to provide the MRFG with some form of precipitation guidance. Earlier efforts (Lu, 1983) in producing temperature guidance indicated that the MOS approach could significantly improve the medium-range forecasts. Moreover, though a previous attempt (Bermowitz, 1978) to derive medium range probability of precipitation (PoP) guidance had failed, the deficiency was likely because the available barotropic extension model contained no moisture forecasts. In contrast, the GSM predicts mean relative humidity and precipitation amount at each forecast projection. Consequently, we've now developed MOS equations to predict the max/min and PoP from 0000 GMT GSM output for projections of 2 to 6 days in advance. Forecasts are available for over 200 stations in the conterminous United States. In deriving the equations, we were guided by the realization that the radiation physics and convective parameterization in the current GSM are about to undergo significant revision. Because of this, we developed the MOS equations so as to minimize the adverse impact of model changes. In this report, we will describe the development of the medium range max/min and PoP forecast equations and explain what we did to mitigate the effects of model changes on MOS.

2. APPROACH

Forecast equations developed through the MOS technique (Glahn and Lowry, 1972) relate surface weather elements (predictands) to forecast fields (predictors) generated by dynamical models. Various climatic terms and observations can also be used as possible predictors. The forecast equations are produced by a forward selection linear regression approach. The resultant linear equations account for physical relationships between the model-generated data and the surface weather elements, as well as certain types of systematic errors in the dynamical model. If the original systematic biases in the dynamical model disappear, or if the model itself is replaced by a different version, the quality of guidance produced by the original MOS equations will likely deteriorate. However, it is possible that improvement in the model forecasts could more than compensate for deterioration in the MOS guidance caused by elimination of the original biases. In this scenario, the MOS guidance could actually improve in skill.

Since we wanted to develop equations from the current GSM archive that could be applied to a modified GSM, we tried to limit potential predictors to those that would be least affected by the proposed model changes. For instance, we used many geostrophic variables and avoided predictors like vertical velocity. For model projections of less than 48 hours, the predictor fields were smoothed by a 5-point arithmetic average. Predictors with projections between 48 and 84 hours were smoothed by a 9-point average, while predictors with projections of 96 hours or greater were filtered with a 25-point average. These spatial averages were intended to smooth out the differences in the "noise" and the small-scale features of the two models. For four variables, namely, 850-mb temperature, 1000-500 mb thickness, 1000-850 mb thickness, and 850-700 mb thickness, we averaged forecasts valid at 0000 and 1200 GMT and used the means as predictors. By computing the averages over time, we hoped to reduce further the impact of changes related to the diurnal cycle of the model. For the PoP equations, we also included the relative frequency of precipitation (.01 inch or more during a 24-h period from 0000 GMT to 0000 GMT) as a potential predictor. Similarly, the 1951-80 max/min normals were used as climatic predictors for the max/min temperature equations. For developmental purposes, the calendar year was divided into two seasons: cool (October through March) and warm (April through September). GSM data were available from October 1980 through September 1984.

As discussed in Dallavalle et al. (1980) and Dallavalle and Murphy (1983), consistency in the max/min forecasts is enhanced when the equations are developed simultaneously. In other words, this type of derivation reduces the chance of the forecast min exceeding the forecast max for the same day. Accordingly, we developed simultaneously the equations that predict the max and min for the same calendar day. With this method, the predictors entered in both equations are the same, but the coefficients differ. The first predictor chosen by the regression process is the variable that produces the greatest reduction of variance for either of the predictands. Subsequent equation terms are selected according to the predictor that produces the greatest reduction in variance for either predictand (max or min) when combined with other predictors already in the equations.

Through experience and testing, we decided to limit the regression equations to 10 terms and require each predictor to reduce the variance by 0.75% for temperature and 0.50% for precipitation. In other words, the regression procedure was stopped when no remaining predictor could contribute the minimum additional reduction of variance or when 10 terms were included in the equations.

3. DEVELOPMENT OF MAXIMUM/MINIMUM TEMPERATURE EQUATIONS

Max/min forecast equations were derived for each station in our sample (single station approach). Potential predictors were GSM forecasts of mean relative humidity, geostrophic wind components, heights, thicknesses, temperature, temperature advection, relative vorticity, and relative vorticity advection (see Table 1 for a detailed list). The sine and cosine of the day of the year were also used as potential climatic predictors along with the normal max and min. All GSM predictors were taken from the 0000 GMT run of the model. The predictand data consisted of the calendar day (midnight to midnight, local time) max and min temperature observations. If Day 1 denotes the day on which the GSM is run, then we derived equations to predict the min and max for Days 2, 3, 4, 5, and 6.

Because of the limited data, we did not test any of the new medium-range forecast equations on an independent sample. During development of the temperature equations, however, we experimented with several methods for deriving the equations. The purpose was to produce equations that, in our view, contained terms that were meteorologically meaningful, while not being too model sensitive. During the testing, we produced equations for each projection by varying the amount of variance that each added predictor needed to contribute in order to be included in the equation. Values of 1% and 0.5% were both used. The equations with a 1% cutoff contained, in our opinion, too few meteorological predictors. In general, when using the 1% value as the reduction of variance cutoff, the number of meteorological predictors ranged from an average of 3.9 on Day 2 to 3.1 on Day 6. The equations derived with 0.5% reduction of variance cutoff contained anywhere from an average of 5.9 meteorological predictors on Day 2 to 5.4 on Day 6. Unfortunately, some of the predictors appeared to be selected because of statistical (model) relationships, rather than meteorological relationships. In our opinion, 0.75% was a reasonable compromise for the cutoff in reduction of variance. We think that using this value in developing the final equations gave us meaningful relationships that were not excessively tuned to the current GSM.

In the past, the upper limit on the number of terms permitted in temperature equations was normally set at 12. In this case, however, we limited our equations to 10 terms to reduce the possibility that the equations might be unstable. Actually, the number of terms per equation rarely exceeded eight because of the limit on the reduction of variance.

Table 2 shows both the reductions of variance and standard errors of estimate for the medium range forecast equations. During the cool season, the equations explain less of the variance of the min temperature than of the max. During the warm season, the opposite occurs.

For all the equations, the predictor used most often was the 1000-850 mb time-averaged thickness. For the shorter projections (Day 4 and earlier), the mean relative humidity was also important. As Table 3 shows, however, with increasing forecast projections (Day 5 and Day 6), the climatic terms (in this case, the cosine day of the year) were selected most frequently.

4. DEVELOPMENT OF PROBABILITY OF PRECIPITATION EQUATIONS

For the PoP equations, we divided the country into 10 regions for the cool season (Fig. 1) and 8 regions for the warm season (Fig. 2). The stations comprising a region were chosen on the basis of geographical proximity and on an analysis of the relative frequencies of precipitation during both seasons. In this development, stations with similar frequencies and geography were grouped together to form regions; equations were then derived for these regions (regionalized equations) rather than for individual stations. Potential predictors (Table 4) used in the development of the equations included GSM forecasts of relative humidity, geostrophic wind components, heights, temperature, temperature advection, relative vorticity, and vorticity advection. We also included the relative frequencies of precipitation and the sine and cosine of the day of the year as climatic predictors. In addition, because we developed the PoP equations for specific regions, we screened the station elevation to account for local geography. The PoP predictand is the occurrence of .01 inch or more of precipitation in a 24-h period ending at 0000 GMT. PoP equations were derived for Days 2, 3, 4, 5, and 6, that is, for periods of 24-48, 48-72, 72-96, 96-120, and 120-144 hours, respectively, after 0000 GMT.

As with the temperature derivation, the PoP equations were not tested on independent data. During development, however, we did conduct several experiments on the dependent data. These tests consisted primarily of adjusting the reduction of variance cutoff, varying the limits of the binary relative humidity predictors, varying the space smoothing of the humidity predictors, and forcing predictors.

The first developmental test was to produce equations by using either a 1% or 0.5% cutoff for the reduction of variance. Our results indicated that with a 1% cutoff, too few meteorological predictors were selected for the equations. On the other hand, the 0.5% cutoff seemed to produce equations containing an adequate number of meteorological fields. Moreover, with this value, equations for adjacent regions appeared to be more consistent. On this basis, we selected 0.5% as the cutoff value throughout the rest of our experimentation and in our final development.

Because of the importance of relative humidity, we chose to use it as both a continuous and a binary predictor. A binary predictor is given a value of one if the variable from which it is derived is less than or equal to the limit and is set to zero, otherwise. We developed three experimental sets of equations using binary limits of relative humidity of 10% to 70% in increments of 10%, 30% to 90% in increments of 10%, and 30% to 70% in increments of 20%. None of the variations significantly improved the reductions of variance compared to the other two. We found that the occurrences of relative humidity of 20% or less were rare. Consequently, we eliminated the possibility of 10% and 20% as binary predictor limits. When the binary limits were in increments of 20%, the equation coefficients were often undesirably large; therefore, in our final equations, we used the limits of 30 to 90%, in increments of 10%.

The final step in our testing was to determine which space smoothings of the relative humidity gave us the best results for the longer range projections (Days 4, 5, and 6). We made available to the screening process the relative humidity predictors filtered through both a 9-point and a 25-point space smoother. We found that the use of a 25-point smoother on the predictors with projections of 96 hours or greater yielded the greatest reduction of variance. Moreover, we think that the 25-point filter may eliminate some small-scale differences between the original and modified GSM.

After examining the PoP equations, we made a final adjustment to the regression procedure. For the Day 2, 3, and 4 equations, the mean relative humidity predictor (continuous) valid at the mid-point of the period (for example, at 36 hours for the 24 to 48 hour period) was chosen most of the time as the first predictor for nearly all regions. To enhance both spatial and temporal consistency, we decided to force the mid-point mean relative humidity into the equations for each region and projection. We think this modification will also enable the forecasters to interpret and modify the PoP guidance more effectively.

The reductions of variance and the NWS Brier scores for the cool and warm season PoP equations are given in Table 5. Note that the NWS Brier score is half the score defined by Brier (1950). The reductions of variance (RV) decrease with increasing projection, as expected. Note that the warm season RV is less than the cool season RV for the same projection. The convective nature of precipitation in the warm season makes the PoP guidance less skillful. In fact, the RV's are such that the warm season PoP equations are approximately one day less skillful than the cool season equations.

The three most important predictors for the PoP equations are given in Table 6. Remember that the mean relative humidity was forced into all the equations as the first predictor. In the cool season, for Days 2, 3, and 4, the model forecasts of the v wind component and the geostrophic vorticity are important in indicating vertical motion. By Day 5, however, as the dynamical model deteriorates in skill, the climatic predictor (precipitation relative frequency) becomes increasingly important. In contrast, during the warm season when significant meteorological features are not as well-organized on the synoptic scale, the climatic predictor is important for all projections.

5. IMPLEMENTATION OF THE MEDIUM-RANGE FORECAST EQUATIONS

The medium-range forecast equations for both max/min temperature and PoP were not tested on independent data. We intend to implement the equations operationally during the spring of 1985 and, thus, use real-time GSM forecasts in evaluating the MOS guidance. Comparisons with climatic forecasts and max/min guidance based on the perfect prog approach will be made. In addition, because NMC has been running two versions of the GSM since December 1984, we may have the opportunity to test the new equations on both the GSM used for development and the revised version. This should allow us to determine if our approach to deriving less model-sensitive MOS equations was successful.

At this point, neither the max/min nor the PoP guidance will be disseminated to the NWS forecast offices. Rather, the MOS prognoses will be available to NMC's Medium Range Forecast Group. The forecasters will be able to interpret and modify the guidance according to their interpretation of the numerical models.

Both the Navy and Air Force expressed an interest in having the medium range PoP forecasts at U.S. military stations for which TDL provides guidance. Since the PoP equations are regionalized, the appropriate equation can be applied to any location within the region. Consequently, PoP guidance will be generated and disseminated for over 100 military locations shown in Fig. 1.

6. CONCLUSIONS AND FUTURE PLANS

We've derived MOS equations from the 0000 GMT version of the GSM to predict calendar day max/min temperature and 24-h PoP for projections of 2 to 6 days. Our goal was to develop relationships that were meteorologically sound and, yet, were not tuned specifically to one particular dynamical model. The equations are to be tested on a modified version of the GSM which is expected to become fully operational in April 1985. Evaluation of our results will indicate how successful our approach was. If the MOS guidance deteriorates little or not at all when applied to the revised GSM, we may be able to use the same technique in responding to future model changes.

7. ACKNOWLEDGMENTS

Many people contribute to the successful operation of the MOS system. Whether the task involves archiving of data, implementation of software, design of the system, or preparation of the manuscript, we could not manage without the able efforts of our fellow workers. We thank all of you.

REFERENCES

Bermowitz, R. J., 1978: The use of MOS for extended-range PoP forecasts. Techniques Development Laboratory Seminar.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. Mon. Wea. Rev., 78, 1-3.

Dallavalle, J. P., J. S. Jensenius, Jr., and W. H. Klein, 1980: Improved surface temperature guidance from the limited-area fine mesh model. Preprints Eighth Conference on Weather Forecasting and Analysis, Denver, Amer. Meteor. Soc., 1-8.

_____, and M. C. Murphy, 1983: Development of LFM-based MOS temperature forecast equations for Alaska. TDL Office Note 83-14, National Weather Service, NOAA, U.S. Department of Commerce, 27 pp.

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.

Hughes, F. D., 1984: Operational medium range forecasting at the National Meteorological Center. Preprints Tenth Conference Weather Forecasting and Analysis, Clearwater Beach, Amer. Meteor. Soc., 333-340.

Klein, W. H., and F. Lewis, 1970: Computer forecasts of maximum and minimum temperature. J. Appl. Meteor., 9, 350-359.

Lu, R., 1983: Development of MOS medium-range max/min temperature guidance for the cool season from the spectral model. Unpublished report, 8 pp. (Available on request from Techniques Development Laboratory, Silver Spring, Md.)

Sela, J. G., 1980: Spectral modeling at the National Meteorological Center. Mon. Wea. Rev., 108, 1279-1292.

Table 1. GSM and climatic predictors used in the derivation of the medium range max/min temperature forecast equations.

Mean relative humidity (surface to ∿ 500 mb)
Geostrophic u, v wind components - 850 mb, 500 mb
Geostrophic wind speed - 850 mb, 500 mb
Height - 500 mb
Temperature - 500 mb
Geostrophic temperature advection - 850 mb, 500 mb
Geostrophic relative vorticity advection - 850 mb, 500 mb
Geostrophic relative vorticity - 850 mb, 500 mb
12-h average temperature - 850 mb
12-h average thickness - 1000-850 mb, 1000-500 mb, 850-700 mb
Normal max temperature
Normal min temperature
Sine of day of the year
Cosine of day of the year

Table 2. Reductions of variance (RV) in percent and standard errors of estimate (SEE) in °F for both the cool season and warm season medium range temperature forecast equations. Results are averages for over 200 stations in the contiguous United States.

Projection	Cool Season		Warm Season	
	RV	SEE	RV	SEE
Day 2 Min	78.9	5.6	85.3	3.8
Day 2 Max	81.3	5.7	81.2	4.6
Day 3 Min	74.0	6.2	81.7	4.2
Day 3 Max	76.1	6.5	75.9	5.3
Day 4 Min	68.0	6.8	76.7	4.8
Day 4 Max	70.1	7.2	70.4	5.8
Day 5 Min	61.2	7.5	71.9	5.2
Day 5 Max	63.6	8.0	64.9	6.3
Day 6 Min	54.7	8.1	68.2	5.5
Day 6 Max	57.6	8.6	60.6	6.7

Table 3. The three most important predictors for the medium-range temperature forecast equations. The selection was based on both the frequency and the order in which the predictor was chosen for the regression equations.

Projection	Cool Season	Warm Season
Day 2 Min/Max	1000-850 mb thickness (12-h average)	Mean relative humidity
	Mean relative humidity	850-mb temperature (12-h average)
	Cosine day of the year	1000-850 mb thickness (12-h average)
Day 3 Min/Max	1000-850 mb thickness (12-h average)	1000-850 mb thickness (12-h average)
	Cosine day of the year	Mean relative humidity
	Mean relative humidity	Cosine day of the year
Day 4 Min/Max	1000-850 mb thickness (12-h average)	1000-850 mb thickness (12-h average)
	Cosine day of the year	Cosine day of the year
	Mean relative humidity	Mean relative humidity
Day 5 Min/Max	Cosine day of the year	Cosine day of the year
	1000-850 mb thickness (12-h average)	1000-850 mb thickness (12-h average)
	850-mb temperature (12-h average)	1000-500 mb thickness (12-h average)
Day 6 Min/Max	Cosine day of the year	Cosine day of the year
	1000-850 mb thickness (12-h average)	1000-850 mb thickness (12-h average)
	850-mb temperature (12-h average)	1000-500 mb thickness (12-h average)

Table 4. GSM and climatic predictors used in the derivation of medium-range PoP forecast equations.

Mean relative humidity (surface to $\sim$ 500 mb)
Precipitation relative frequencies
Geostrophic u, v wind components - 850 mb, 500 mb
Geostrophic wind speed - 850 mb, 500 mb
Height - 500 mb
Temperature - 500 mb
Geostrophic temperature advection - 850 mb, 500 mb
Geostrophic relative vorticity advection - 850 mb, 500 mb
Geostrophic relative vorticity - 850 mb, 500 mb
Station elevation
Sine of day of the year
Cosine of day of the year

Table 5. The average reductions of variance (RV) in percent and the NWS Brier scores (BS) are given for the medium range cool and warm season PoP equations. Note that the Brier score presented here is one-half of that defined by Brier (1950).

Projection	Cool Season		Warm Season	
	RV	BS	RV	BS
Day 2	33.8	.118	23.5	.135
Day 3	24.4	.135	16.9	.147
Day 4	17.4	.148	12.1	.156
Day 5	13.1	.156	8.9	.162
Day 6	9.1	.164	6.3	.165

Table 6. Same as Table 3 except for the PoP equations.

Projection	Cool Season	Warm Season
Day 2	Mean relative humidity 500-mb geostrophic v wind 500-mb geostrophic vorticity	Mean relative humidity Precip. relative frequency 850-mb geostrophic vorticity
Day 3	Mean relative humidity 500-mb geostrophic v wind 500-mb geostrophic vorticity	Mean relative humidity Precip. relative frequency 500-mb geostrophic u wind
Day 4	Mean relative humidity 500-mb geostrophic v wind 500-mb geostrophic vorticity	Mean relative humidity Precip. relative frequency 500-mb geostrophic v wind
Day 5	Mean relative humidity Precip. relative frequency 500-mb geostrophic v wind	Mean relative humidity Precip. relative freqency 500-mb geostrophic v wind
Day 6	Mean relative humidity Precip. relative frequency 500-mb geostrophic vorticity	Mean relative humidity Precip. relative frequency 500-mb geostrophic u wind

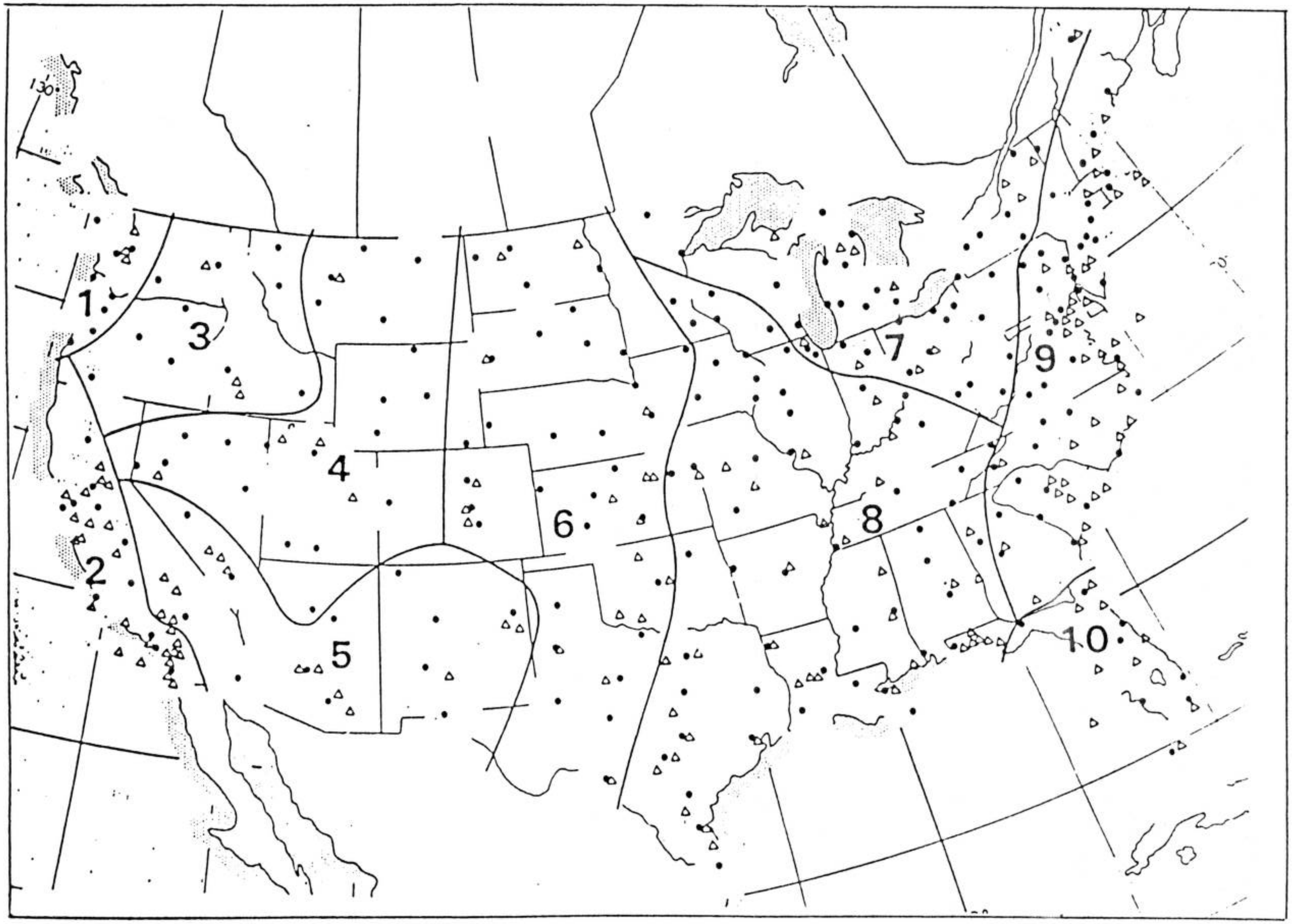

Figure 1. Regions used in the development of the cool season PoP equations. Civilian stations are denoted by the solid circles while the military stations are indicated by the open triangles. In development, only the data for the civilian sites were used; in operations, however, the regionalized MOS equations can be applied to all the stations.

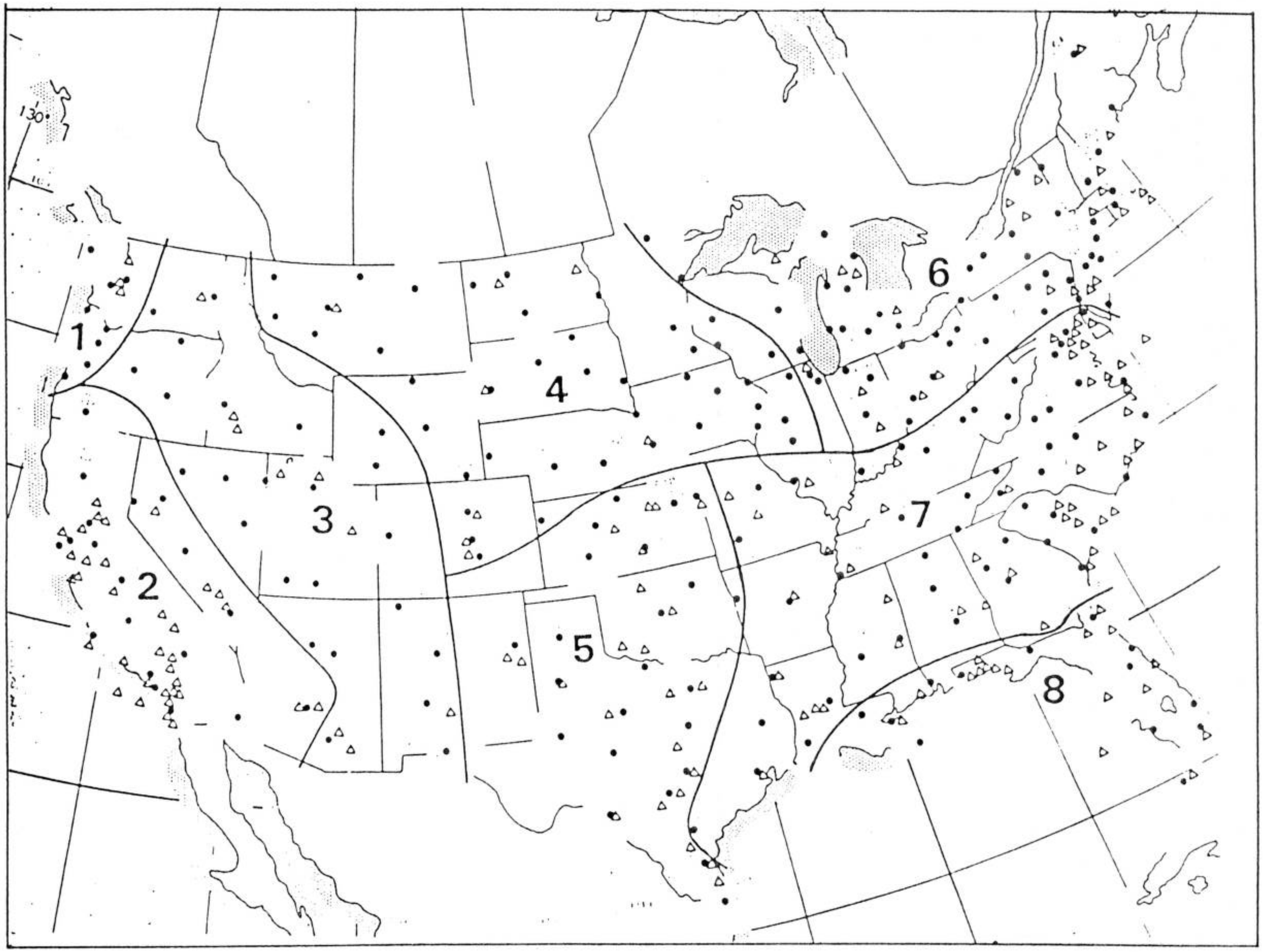

Figure 2. Same as Fig. 1, but for the warm season.