Atmósfera (1988), **1**, pp. 113-129

Statistical prediction of monthly mean temperature anomalies in the United States during winter months

WILLIAM H. KLEIN and RUNHUA YANG

Department of Meteorology, University of Maryland, College Park, MD 20742-2425, U.S.A.

(Manuscript received April 21, 1988; accepted June 10, 1988)

RESUMEN

Anomalías de la temperatura media mensual del aire en 50 estaciones de superficie en el área contigua de los Estados Unidos, durante los meses invernales de enero, febrero y marzo, de 1951 a 1980, son estadísticamente cribados como funciones de datos pasados, presentes y futuros, promediados en diferentes intervalos de tiempo. Entre los predictores potenciales se incluyen los campos de alturas en 700 mb, la temperatura del aire y de la superficie del mar, la cubierta de nieve y los índices de ENSO. Las anomalías de altura futuras son amortiguadas de acuerdo con la precisión de los pronósticos numéricos diarios de hasta 10 días, producidos operacionalmente en el Centro Nacional de Meteorología (NMC). Toda la significación estadística es evaluada por medio de simulaciones de Monte Carlo. El principal resultado es que el uso juicioso de predicciones de la altura para 1-10 días, ofrece la promesa más grande para el mejoramiento inmediato de las predicciones de la temperatura media mensual.

ABSTRACT

Anomalies of mean monthly air temperature at 50 surface stations in the contiguous United States during the winter months of January, February and March from 1951 to 1980 are statistically screened as functions of earlier, centered or future time means of different length. Potential predictions include fields of 700 mb height, air and sea surface temperature, snow cover and ENSO indices. Future height anomalies are damped in accordance with the accuracy of daily numerical prognoses out to 10 days produced operationally at the National Meteorological Center. All statistical significance is evaluated by means of Monte Carlo simulations. The principal result is that judicious use of 1-10 day numerical height predictions offers the greatest promise for immediate improvement of monthly mean temperature forecasts.

1. Introduction

In the Climate Analysis Center (CAC) of the United States, forecasts of monthly mean anomalies of surface air temperature (SAT) have long been produced by a two-step process developed by Namias (1953). The first step is to predict the coming month's distribution of mean 700 mb height. The second step is to translate the prognostic 700 mb height field produced in step 1 into a field of concurrent SAT. During the last six years, step 2 has been largely performed in an objective fashion by means of specification equations derived for the United States by Klein (1983, 1985a), for Canada and Alaska by Klein (1985b), and for Europe and Asia by Klein and Yang (1986). However, step 1 is still largely subjective in nature (Gilman, 1983), and its accuracy leaves much to be desired (Walsh, 1984). For example, SAT forecasts produced by applying objective specification equations to official monthly mean height predictions for the 5-year period 1982-86 had an average skill score of approximately 13. If the height forecasts had been perfect, the skill score of the temperature specifications would have averaged about 40, three times as much (Klein, 1985).

In view of the above, the main object of the present paper is to test the feasibility of bypassing step 1 and producing a completely objective prediction of SAT, without the necessity of first forecasting next month's heights. A series of statistical experiments will be described to see how well this goal can be achieved by using either purely antecedent predictors (section 3) or damped short range numerical predictions (section 5).

Similar studies have recently been published for the first technique by Barnett and Preisendorfer (1987) and for the second by Harnack *et al.* (1986). An attempt to combine specification with prediction will also be discussed (section 4). But first our basic methodology and data will be explained in the next section.

2. Basic data and methods

The basic data for this study were taken from the three winter months of January, February and March during the standard WMO (World Meteorological Organization) 30-year period from 1951 through 1980. In order to increase the sample size to 90, data for the three months were combined or pooled for regression analysis. Because of this pooling, all data were expressed as standardized anomalies (computed separately for each month as departure from 30-year mean divided by standard deviation) to equalize the contribution of Jan., Feb. and Mar. to the overall result.

The predictand (dependent variable) in this paper was always the standardized anomaly of monthly mean surface air temperature (SAT) during the winter months of Jan., Feb. and March for the 30 years: 1951-80 at a network of 50 surface stations in the contiguous United States (Fig. 1). The predictors to be discussed in section 3 consisted of four types of independent variables expressed as

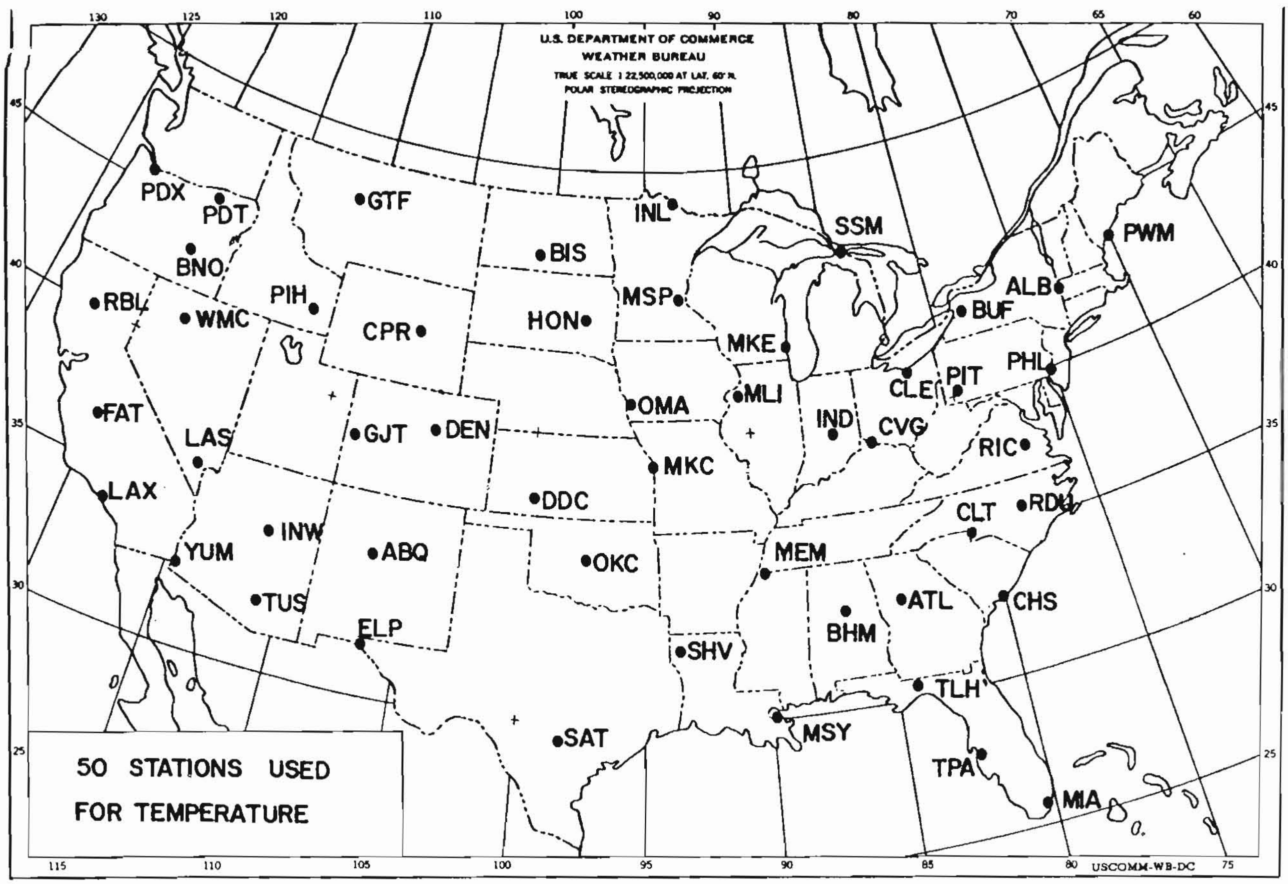

Fig. 1. Locations and call letters of 50 surface stations used for temperature experiments.

standardized anomalies of the mean observed during the month before the target month; namely 700 mb height at 56 widely spaced grid points over North America and surrounding oceans (Fig. 2), the field of SAT (Fig. 1), the latitudinal extent of snow cover in the United States at each 5 degrees of longitude from 70^{o}W to 120^{o}W, and seven indices that express the effects of El Niño and the Southern Oscillation (ENSO). Because of their great persistence, the ENSO-related predictors were also tested at five additional lags: 2, 3, 4 months, plus 1 and 2 seasons.

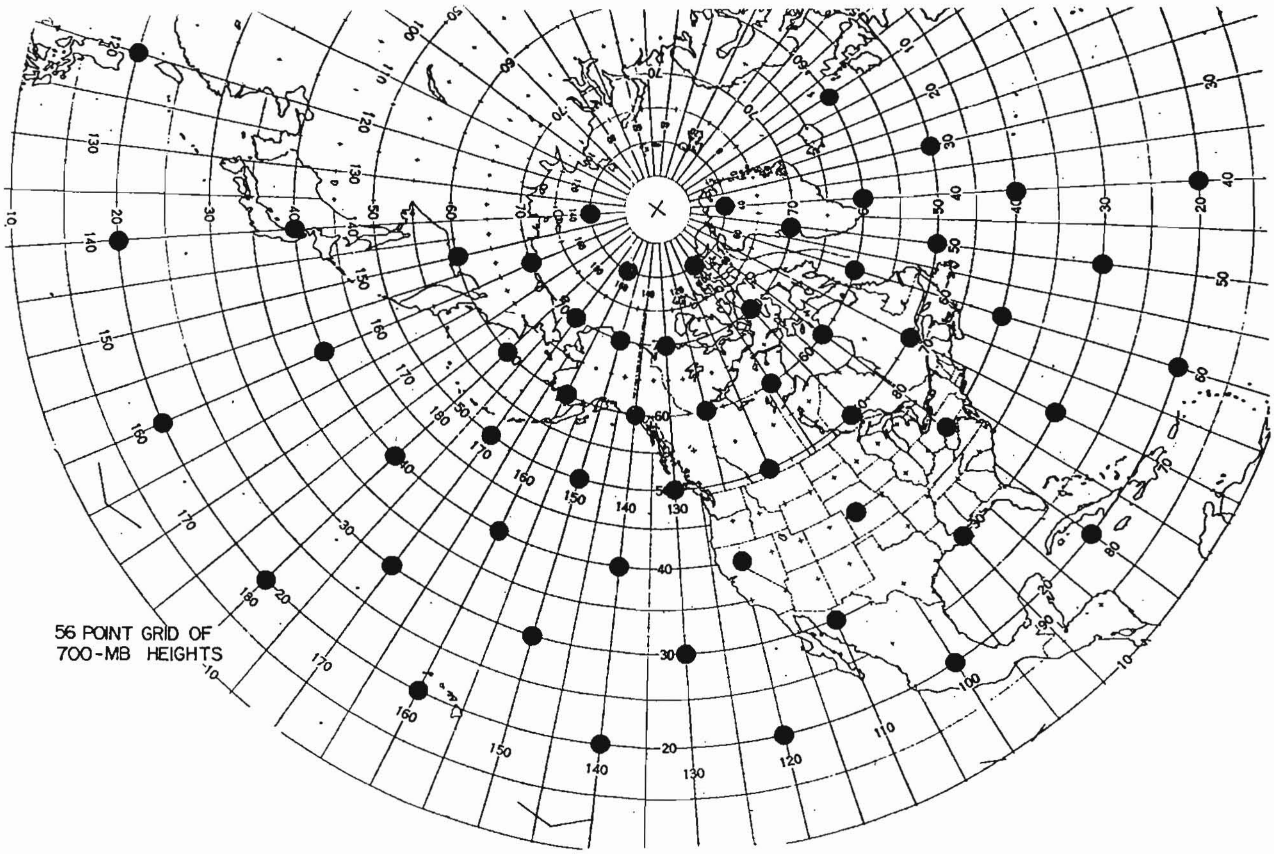

Fig. 2. Locations of 56 grid points used to delineate the field of 700 mb heights.

The predictors to be discussed in section 4 are identical to those used in section 3 but with one addition. Consideration will be given to not only antecedent variables (prediction), but also to the concurrent field of 700 mb height (h) taken at the same grid of 56 points plotted in Fig. 2 (specification). The object is to examine how much specification can be improved by including previous observed variables.

The potential predictors to be discussed in section 5 will be standardized anomalies of 700 mb height at the same grid of 56 points plotted in Fig. 2 but for shorter periods than a month. These independent variables were tested in order to capitalize on the skill inherent in daily numerical prognoses routinely produced by the Medium Range Forecast model of the National Meteorological Center (NMC) (Sela, 1982 and many subsequent improvements). Both observed and forecast values of 700 mb height were obtained from data archived at CAC and then processed in various ways to be discussed in section 5.

Our basic method of determining the effect of the various predictors on SAT will be a stepwise forward selection procedure of linear multiple regression called screening (Klein, 1983). Although screening has been used successfully to derive specification equations for SAT, it has suffered from lack of a satisfactory cut-off criterion to terminate the process of adding terms to the multiple regression equation. For the present paper, we therefore determine the cut-off criterion in a completely objective fashion by means of a Monte Carlo technique similar to one used by Walsh (1984). We randomly shuffled the year of our temperature predictand 100 times, kept the months and all predictor data intact, applied our screening regression program to each shuffled time series, computed classical F values each time a variable was added to the regression equation, and ranked the resulting F's to determine the 95th percentile. We repeated this procedure for ten different stations around the country and averaged the results to obtain Fig. 3, showing the critical F value as a function of the number of variables in the regression equation (step) and the number of potential predictors (M). These critical F's are greater than classical values (crosses) and increase with increasing M because

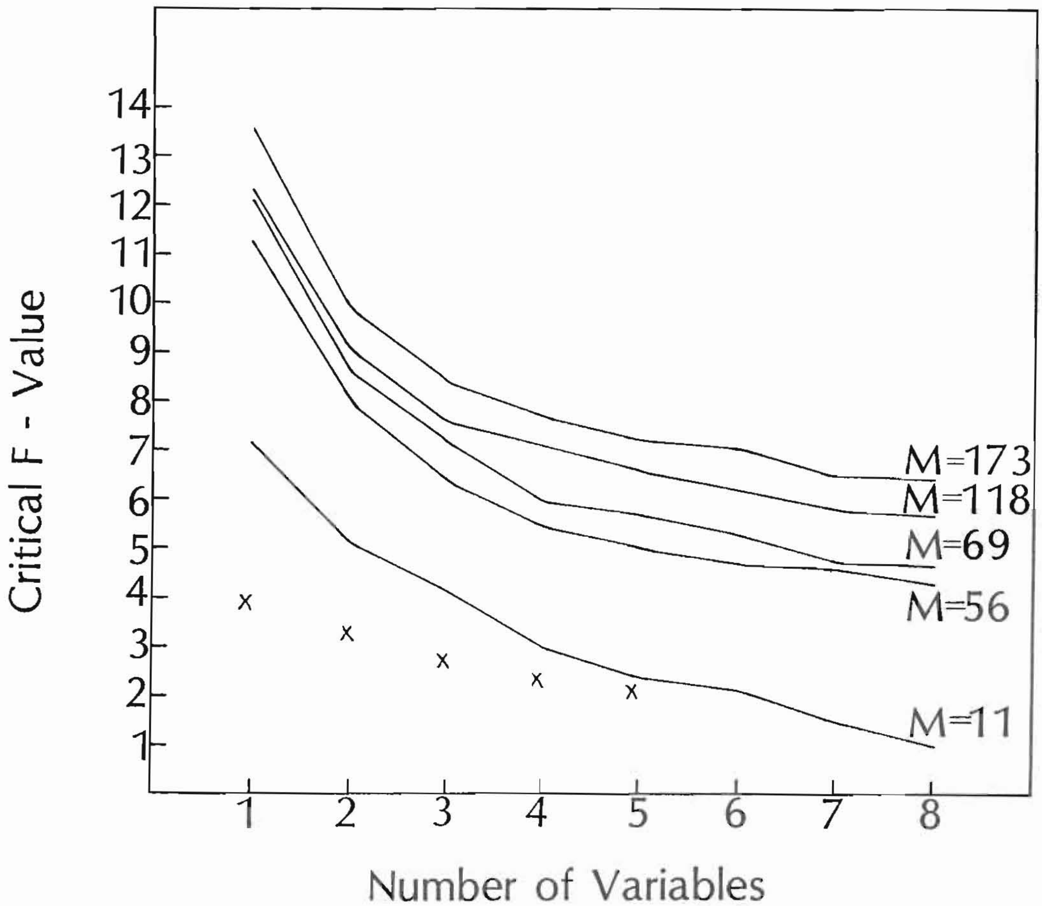

Fig. 3. Mean F-values at the 95% significance level obtained by screening data with the year of the predictand randomly shuffled 100 times at each of 10 stations and the predictors unchanged. The critical F is shown as a function of the number of potential predictors (M) and variables in the regression equation. The classical case (without screening) is given by the crosses.

of the *a posteriori* nature of screening in which the variable giving the highest correlation is selected after the fact. Therefore the critical F's plotted in Fig. 3 must be exceeded for our equations to be truly significant at the 95% confidence level.

3. Use of antecedent predictors

Table 1 summarizes results of screening our winter temperature predictand as a function of monthly mean antecedent potential predictors, all observed one month before SAT and expressed as standardized anomalies, with the cut-off point determined from Figure 3. On average, the field of 56 heights (H) explains 6.5% of the 50-station temperature variance (RV or square of the multiple correlation coefficient x 100) by means of 0.5 variables, with significant contribution at 19 stations. Adding the previous local temperature produces a significant improvement at only 1 station (Pocatello).

Predictors				Sig. contrib. of added variables when first selected	
Type	No.	RV(%)	No. terms	No. stations	Added RV(%)
H	56	6.5	0.5	19	17.0
H, L	57	6.6	0.5	1	17.8
H, T	106	7.0	0.6	5	11.1
H, S	67	8.5	0.7	9	11.0
H, T, S	117	8.9	0.7	12	12.1
T	50	3.6	0.3	12	15.0
S	11	4.2	0.4	20	10.4

Table 1. Results of screening monthly mean temperature standardized anomalies at 50 surface stations in the United States during Jan., Feb. and March, 1951-1980, as a function of 700 mb height at 56 grid points over North America and adjacent oceans (H), local temperature (L), temperatures at network of 50 stations (T), and snow extent at 11 longitudes in the United States (S). All predictors are taken one month before the target month and expressed as standardized anomalies. All significance is determined by a Monte Carlo technique.

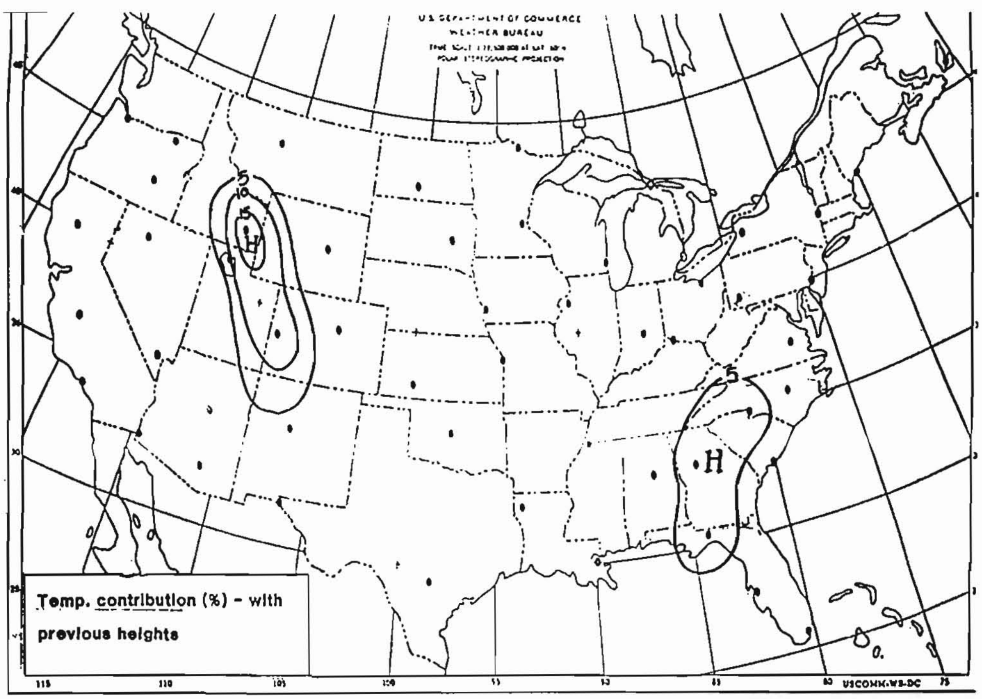

Fig. 4. Percent of variance of monthly mean standardized temperature anomaly contributed by surface temperature upon first being selected in screening previous month's standardized anomalies of temperature at 50 cities of Fig. 1 and height at 56 points of Fig. 2. All results based on data for Jan., Feb. and Mar., 1951-1980.

Considering temperatures at the entire network of 50 stations (T), in conjunction with H, raises the RV to 7% and the number of terms to 0.6. The temperature now contributes significantly at 5 stations, largely because of northwest to southeast movement of temperature anomalies in the Southeast and persistence of local effects in the Great Basin, as implied by Figs. 4 and 5.

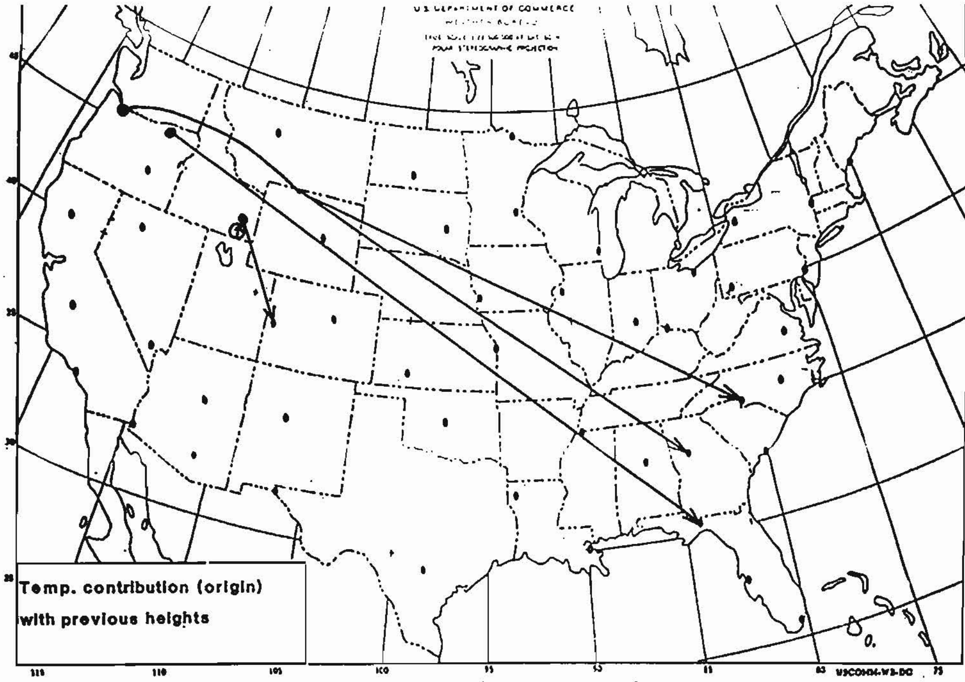

Fig. 5. Same as Fig. 4 but for location of first station selected (origin of arrow) to predict temperature of station at tip of arrow. For example, last month's temperature at Pendleton helps predict this month's temperature at Tallahassee. Stations without arrows did not select previous temperature as a predictor.

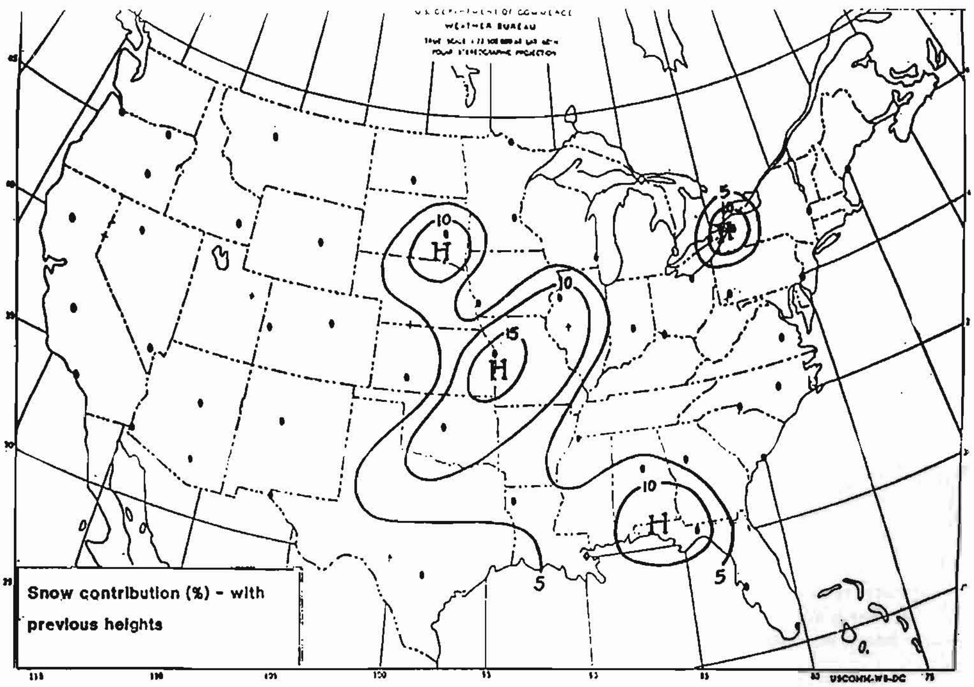

Fig. 6. Same as Fig. 4 but for screening the extent of snow cover at 11 longitudes from 70°W to 120°W instead of temperature at 50 stations.

Line 4 shows the effect of screening H and the latitudinal extent of snow cover (S) at 11 longitudes in the United States at each 5 degrees from 70°W to 120°W, as digitized by Walsh *et al.* (1982). This procedure raises the RV to 8.5% with 0.7 terms, and S contributes significantly at 9 stations. Fig. 6 shows that these stations are limited to the eastern half of the country where S adds from 0 to 16% to the RV. Fig. 7 indicates that S at longitude 100°W is more important than the local S, probably because of prevailing southeasterly movement of cold air masses in winter. Line 5 of Table 1 shows that screening H, T and S together (M=117) raises the average RV to 8.9% by means of only 0.7 terms. If equations based upon this combination of predictors hold up well on independent data, they might produce a useful objective monthly temperature forecast.

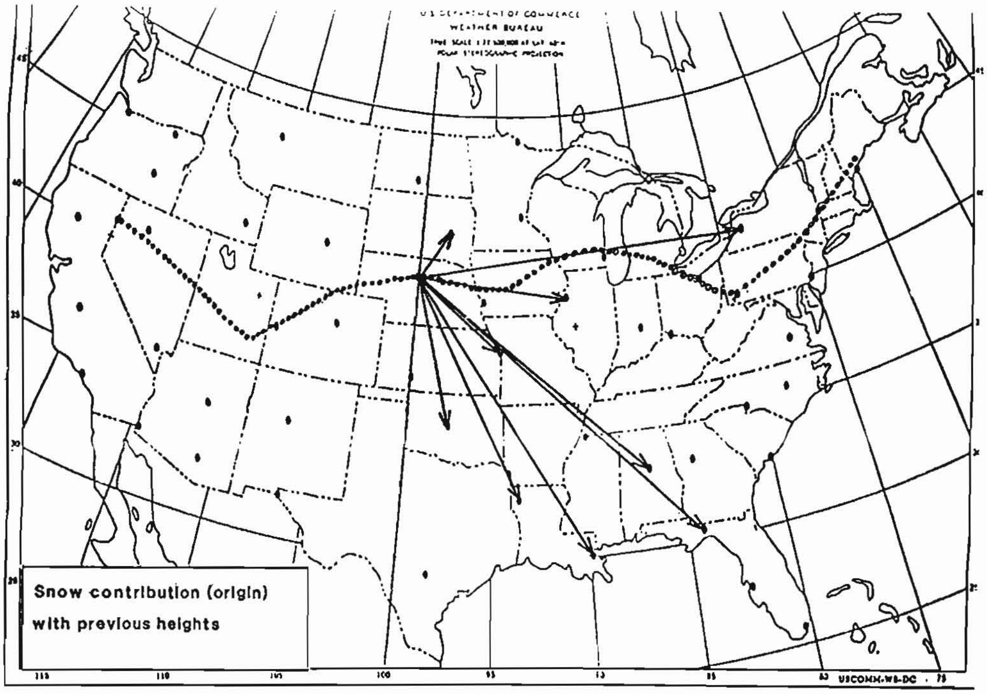

Fig. 7. Longitude of snow cover first selected in screening monthly mean surface temperature at 50 cities of Fig. 1 as a function of previous month's 700 mb height at 56 points of Fig. 2 and snow extent (Walsh *et al.*, 1982) at 11 longitudes from 70°W to 120°W. The arrows originate where the selected longitude intersects the southernmost extent of the mean snow cover for Jan., Feb. and Mar., 1951-1980 (line of open circles) and terminate at the station affected.

The last two lines of Table 1 show that the fields of T and S by themselves explain only about 4% of the variance of SAT. Comparison with line 1 shows that antecedent heights are somewhat more important than previous temperatures or snow extent in predicting next month's surface temperature. Locations of the first grid point selected in screening SAT as a function of H are illustrated in Fig. 8. No consistent patterns can be found in the location of the points in this figure.

Because the last few years have seen a revival of interest in the effects of El Niño and the Southern Oscillation on long-range weather forecasting (Rasmusson and Carpenter, 1982), we screened SAT as a function of H plus each of seven different ENSO-related predictors, all taken at six lags (1, 2, 3 and 4 months plus 1 and 2 seasons. The data for the 30 years: 1951-1980 were obtained from

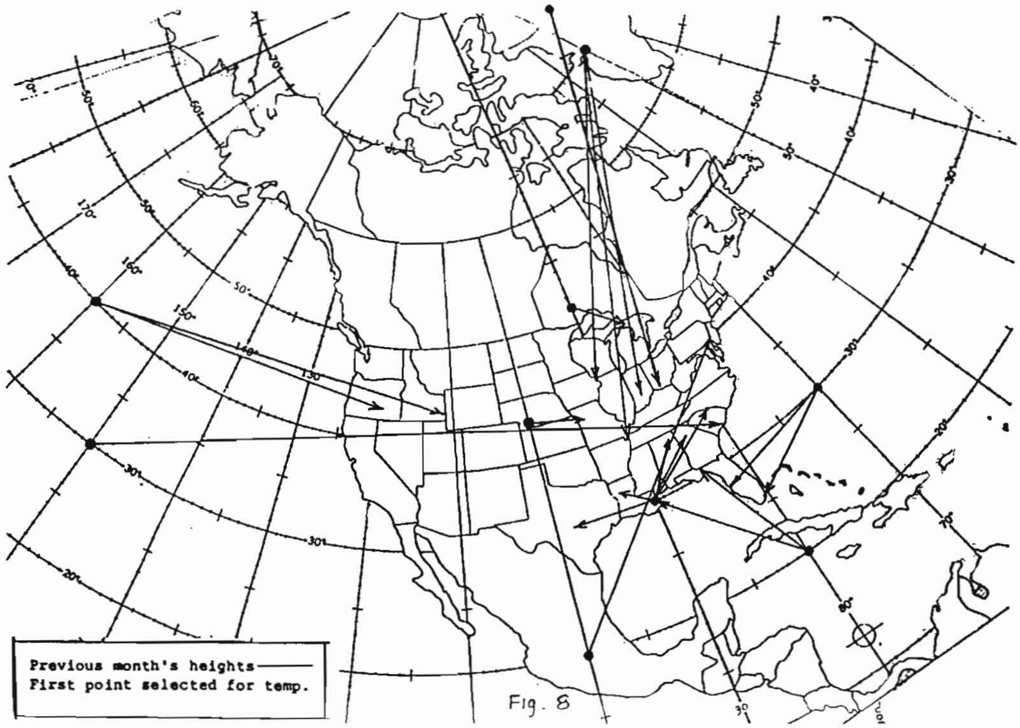

Fig. 8. Location of first grid point with height selected (origin of arrow) in predicting temperature at tip of arrow.

CAC, Weare (1986) and Angell and Korshover (1983). The ENSO variables are located in Fig. 9, and the screening results are summarized in Table 2. Note that all predictors are taken in the equatorial Pacific basin, where four involve sea surface temperature (SST), one (Tahiti-Darwin) sea level pressure, one precipitation (Line Islands: Fanning and Christmas), and one (Balboa) the quasi-

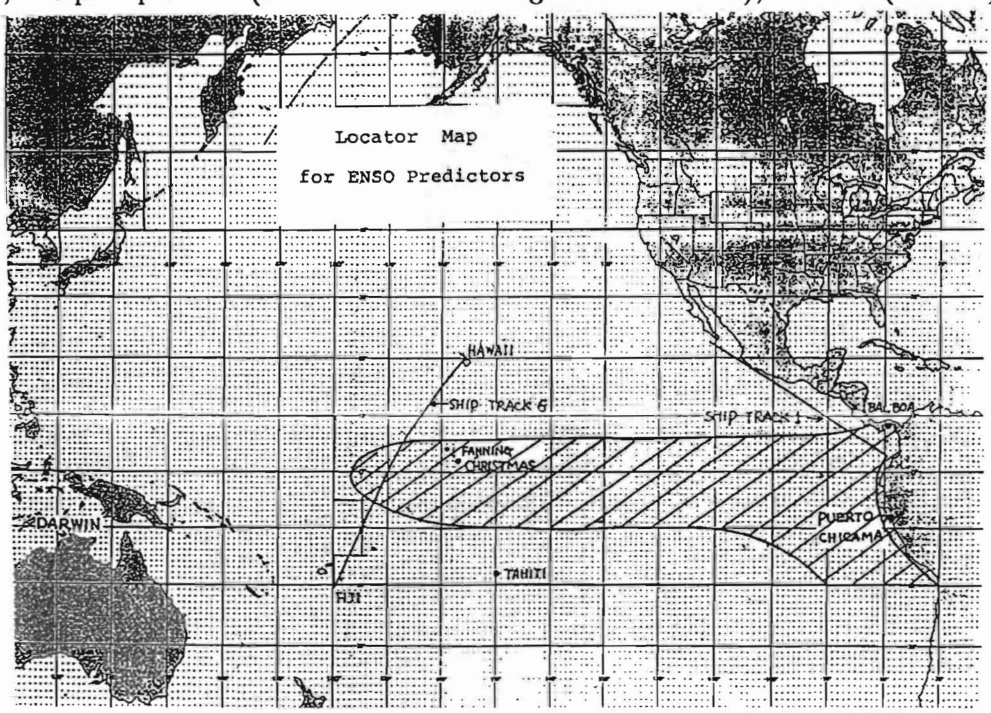

Fig. 9. Locator map for ENSO predictors. Shading delineates the area with loadings ≥ 0.1 in the empirical orthogonal function of sea surface temperature (SST) used for the El Niño index of Weare (1986). Ship tracks 1 and 6 locate regions with SST defined by Rasmusson and Carpenter (1982).

biennial oscillation. None of the ENSO predictors was able to increase the RV yielded by H alone. It is possible that stratification of the cases into warm and cold events might yield better results (Ropelewski and Halpert, 1986; Weare, 1987), but we did not have enough data in our sample to justify such a procedure.

Predictors	50-station mean		Significant contribution	
	RV(%)	No. terms	No. stations	Added RV(%)
Heights at 56 points	6.5	0.5	19	17.0
Tahiti - Darwin pressure	6.4	0.5	2	10.5
Line Islands precipitation	5.9	0.4	0	0
Puerto Chicamo SST	5.9	0.4	0	0
Ship track no. 1 (SST)	5.9	0.4	0	0
Ship track no. 6 (SST)	6.4	0.5	2	7.7
El Niño index (SST)	6.4	0.5	2	12.9
Balboa 50 mb zonal wind	5.9	0.4	0	0

Table 2. Results of screening monthly mean standardized temperature anomalies at 50 surface stations in the United States during Jan., Feb. and March, 1951-1980, as a function of various ENSO predictors in conjunction with 700 mb heights at 56 grid pints over North America and adjacent oceans during previous month. All predictors are expressed as standardized anomalies and taken at 6 different lags (1, 2, 3 and 4 months plus 1 and 2 seasons). All significance is assessed by a Monte Carlo technique.

4. Mixed prediction and specification

We next performed a series of screening experiments to see whether the antecedent predictors listed in Tables 1 and 2 could improve the specification of monthly mean temperature for the same 50 cities plotted in Fig. 1, same period of record and same 56 monthly mean 700 mb heights (Fig. 2), but for concurrent values (h) rather than antecedent ones (H). Table 3, line 1, shows that specification from concurrent heights explains 66% of the variance of SAT, compared to only 6.5% from antecedent heights (Table 1, line 1). Inclusion of previous local temperature, as in specification equations currently operational at CAC, increases the RV by an average of 2.5%, with significant contributions at 24 stations (line 2). Line 3 shows that inclusion of the entire network of antecedent temperatures, rather than only L, raises the RV of specification to almost 71%, and T is now significant at 43

Predictors		50-station means		Sig. contrib. when first selected	
Type	No.	RV(%)	No. terms	No. stations	Added RV(%)
h	56	66.2	3.9	50	66.2
h, L	57	68.7	4.6	24	5.0
h, T	106	70.6	4.8	43	5.2
h, S	67	68.2	4.6	29	3.5
h, L, S	68	69.3	4.7		
h, H	112	65.7	3.6	21	5.7
h, T, S	117	69.6	4.3		
h, H, T, S	173	70.6	4.6		

Table 3. Results of screening monthly mean standardized temperature anomalies at 50 surface stations in the United States during Jan., Feb. and March, 1951-1980, as a function of simultaneous 700 mb height at 56 grid points over North America and adjacent oceans (h), in conjunction with various predictors one month previous including 56 heights (H), local temperature (L), network of 50 temperatures (T), and snow extent at 11 longitudes (S). All predictors are expressed as standardized anomalies and all significance is determined by a Monte Carlo technique.

stations. The contribution of T varies from 0 in several regions to 17% in the upper Plateau (Fig. 10). As before, must of this increment comes about through west-to-east motion, except for some persistence in the western third of the country (Fig. 11). The last five lines of Table 3 show that

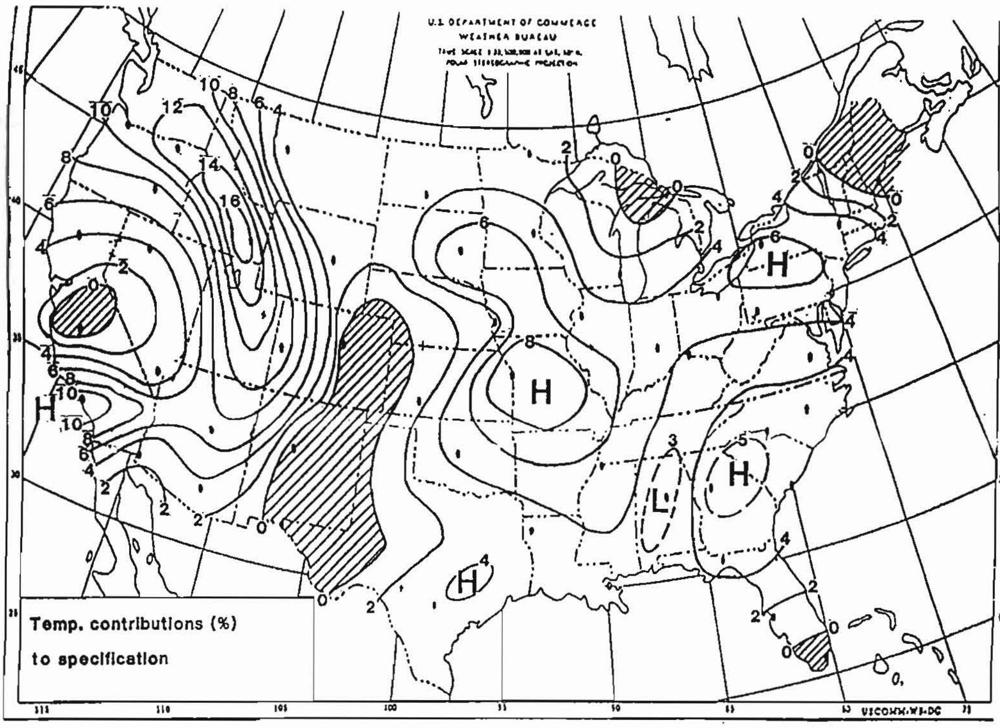

Fig. 10. Same as Fig. 4 except concurrent heights were screened instead of previous ones.

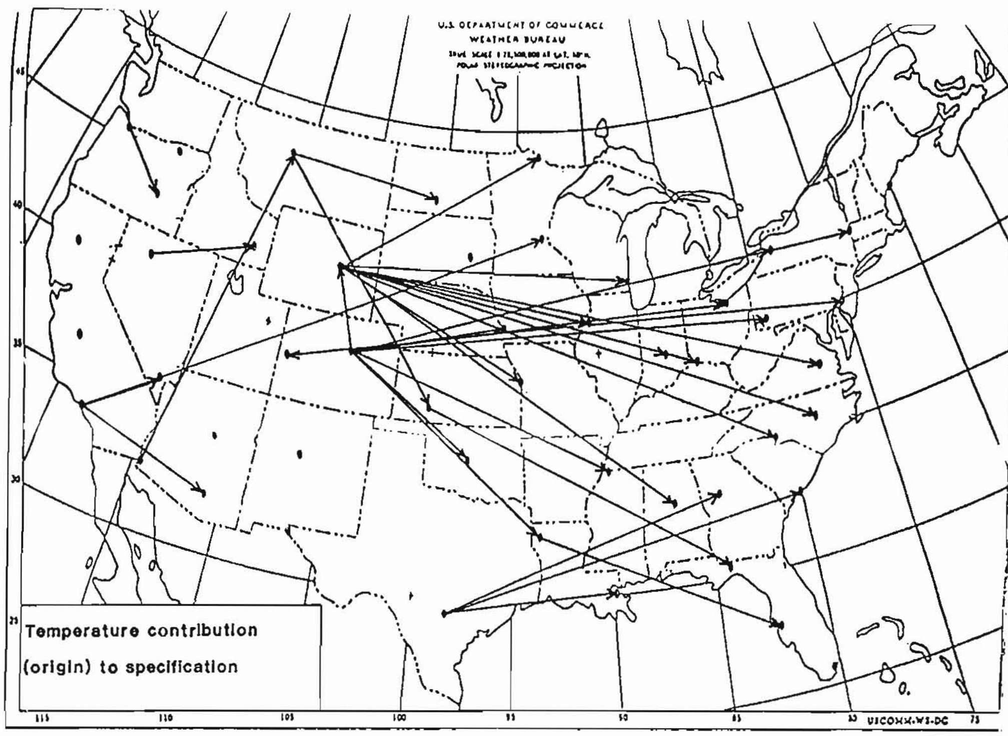

Fig. 11. Same as Fig. 5 except concurrent, instead of previous, heights were screened.

addition of previous snow (S) and heights (H), alone or in combination, produces no improvement. We therefore conclude that the accuracy of temperature specification might be increased by a small but significant amount by screening the field of previous monthly temperatures, rather than only the local values, together with the field of concurrent heights.

For our next series of experiments, we screened the same ENSO predictors and lags used for Table 2, but this time in conjunction with the simultaneous height field (h), rather than the antecedent heights (H). Table 4 shows, as before, that none of these variables was able to significantly increase the RV yielded by heights alone. The best ENSO index was precipitation in the Line Island (line 3), which produced the highest RV and largest number of significant stations of any ENSO variable, but the overall increase in RV was only 0.2%. It seems that the effects of ENSO on United States weather are produced primarily through the circulation pattern, as expressed in the height field, rather than by any direct effect on temperature.

Predictors	50-station mean		Significant contribution	
	RV(%)	No. terms	No. stations	Added RV(%)
Heights at 56 points	66.2	3.9	50	66.2
Tahiti - Darwin pressure	65.8	3.7	5	3.2
Line Islands precipitation	66.4	3.9	15	3.8
Puerto Chicamo SST	65.4	3.6	1	8.1
Ship track no. 1 (SST)	65.5	3.6	2	2.1
Ship track no. 6 (SST)	65.8	3.7	3	3.0
El Niño index (SST)	65.7	3.7	4	3.6
Balboa 50 mb zonal wind	65.6	3.7	1	0.8

Table 4. Results of screening monthly mean standardized temperature anomalies at 50 surface stations in the United States during Jan., Feb. and March, 1951-1980, as a function of various ENSO predictors in conjunction with simultaneous 700 mb height at 56 grid points over North America and adjacent oceans. All predictors are expressed as standardized anomalies and taken at 6 different lags (1, 2, 3 and 4 months plus 1 and 2 seasons). All significance is assessed by a Monte Carlo technique.

5. Use of other periods

In this section we use the same SAT predictand as before, but we limit our potential predictors to the heights of Fig. 2 for periods other than a month. The left side of Fig. 12 plots the RV obtained by screening SAT as a function of standardized anomalies of mean heights observed for periods from 5 to 28 days *before* the target month begins. The decrease of RV with length of averaging period indicates that recent data are more important than older data in forecasting SAT during winter. This result is consistent with papers by Roads and Barnett (1984) on forecasting 500 mb height and Donn *et al.* (1986) on monthly temperature prediction. However, a study by Harnack *et al.* (1986) suggests that an opposite conclusion might apply to the warm season.

The right side of Fig. 12 is similar to the left side but for heights during 5-28 days period *after* the beginning of the target month. The RV now increases with length of averaging period as we gain more information about the target month, but no period is as good as the observed monthly mean (RV of 66%). On the other hand, all shorter periods are better than the previous month's mean (RV of only 6.5%).

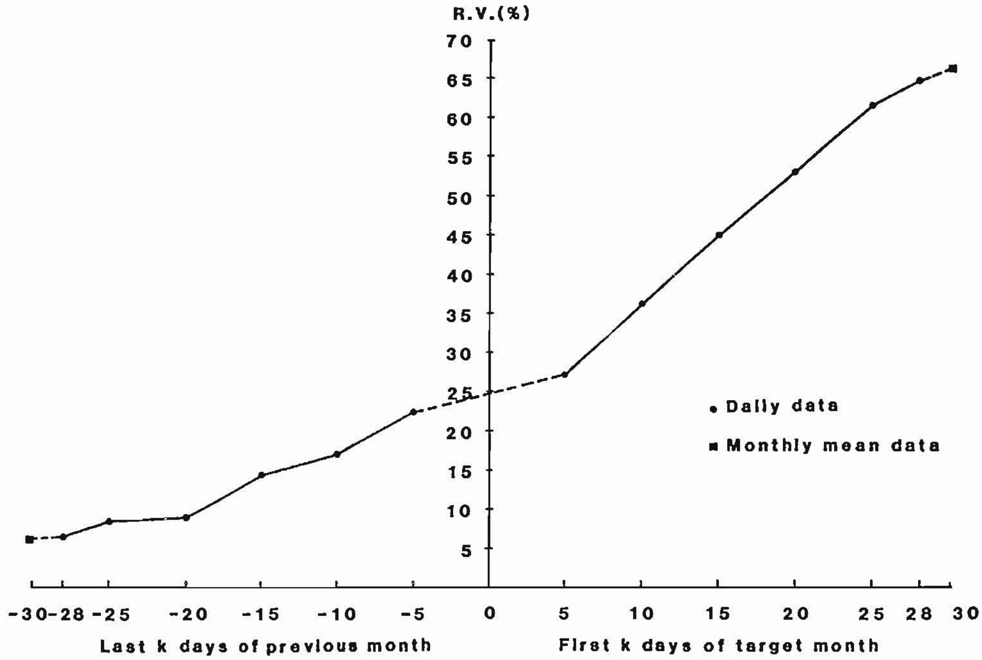

Fig. 12. Percent of monthly mean temperature variance of the 50 stations of Fig. 1 explained by mean 700 mb height at 56 points of Fig. 2 for periods from 5 to 28 days before (left) and after (right) beginning of target month.

Figure 13 shows the results of screening SAT as a function of standardized anomalies of mean heights of varying length but always *centered* on the first day of the target month. For example, the 20-day mean consists of heights observed during the last 10 days of the previous month plus the first 10 days of the new month, the 30-day mean is made up of the 15 days before and 15 days after the 1st of the month, etc. Values of RV in this figure rise slowly from 32% at 10 days to 48% at 60 days

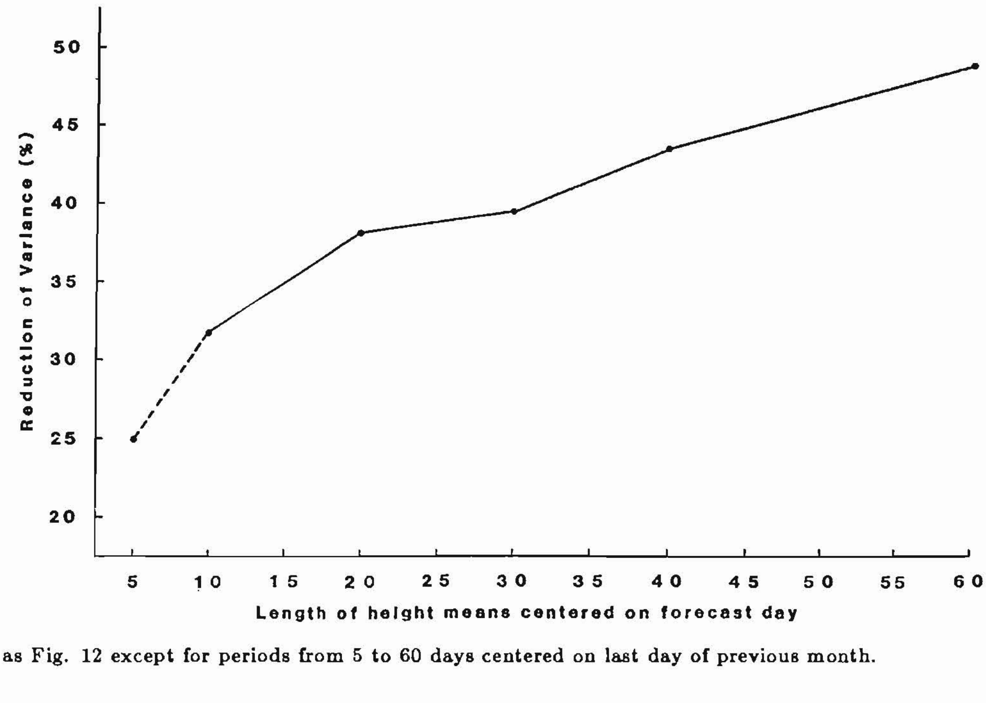

Fig. 13. Same as Fig. 12 except for periods from 5 to 60 days centered on last day of previous month.

as more and more information about the target month is included in the centered mean. However, only one of these centered means is a better predictor than its second half alone, as can be seen by comparing the RV's in Fig. 13 with the RV's for periods of half its length in Fig. 12, right side. For example, the centered 30-day mean in Fig. 13 has an RV of 39%, but its corresponding 15-day mean in Fig. 12 (right) has an RV of 45%. Similar statistics for the centered 60-day mean versus second half 30-day mean are 48% and 66%, respectively. Thus, it appears that previous heights, whether in the form of antecedent values (Fig. 12, left) or centered means (Fig. 13), make little contribution to SAT not already contained in future heights.

Figure 12 (right) shows that 36% of the SAT variance could be explained by a perfect height forecast for the first ten days of the target month. Since medium range numerical predictions are routinely prepared at NMC for this period, we estimated their usefulness in forecasting SAT by simulating daily prognostic heights and then screening their means. We used two types of simulation. The first multiplied standardized anomalies of observed 700 mb height for days 1 to 10 by the exponential damping functions of Harnack *et al.* (1986), reproduced in Fig. 14. The second weighted

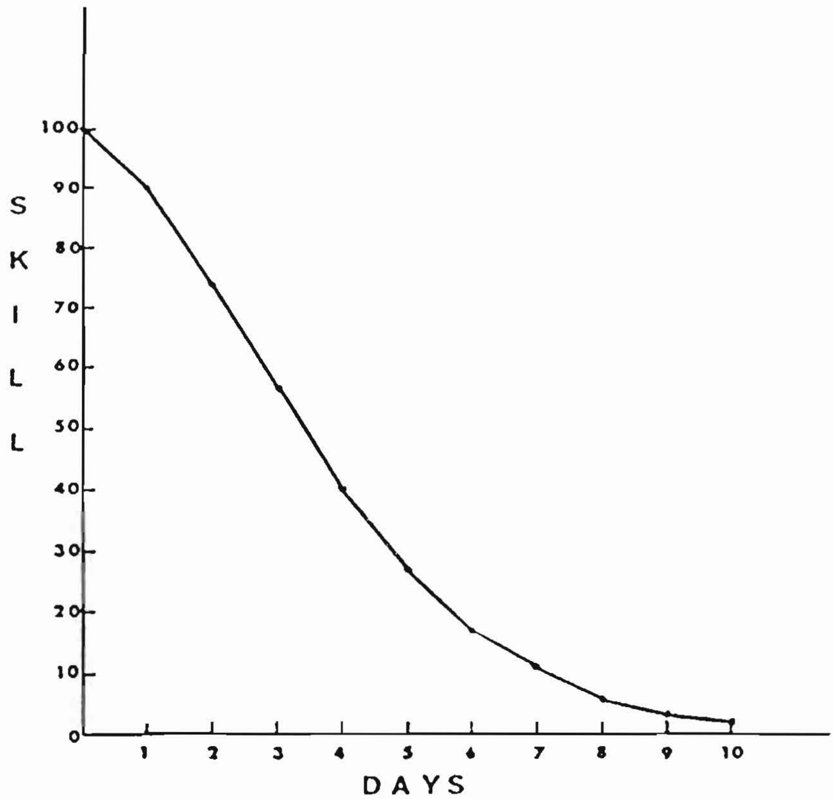

Fig. 14. Plot of skill (%) vs time (days) of the numerical 700 mb height anomaly forecasts simulated by the exponential damping function of Harnack *et al.* (1986).

daily numerical forecasts at NMC by their mean correlations with observed heights in the Northern Hemisphere (Fig. 15 with correlation assumed to be 0 at 15 days). The results (Fig. 16) show that simulation by either method for 5, 10 or 15 days should produce an RV at least 25%, greater than any value for antecedent heights plotted in Fig. 12. Thus if systematic errors (Epstein, 1987) could be eliminated from numerical models, they would be very valuable in making monthly mean forecasts.

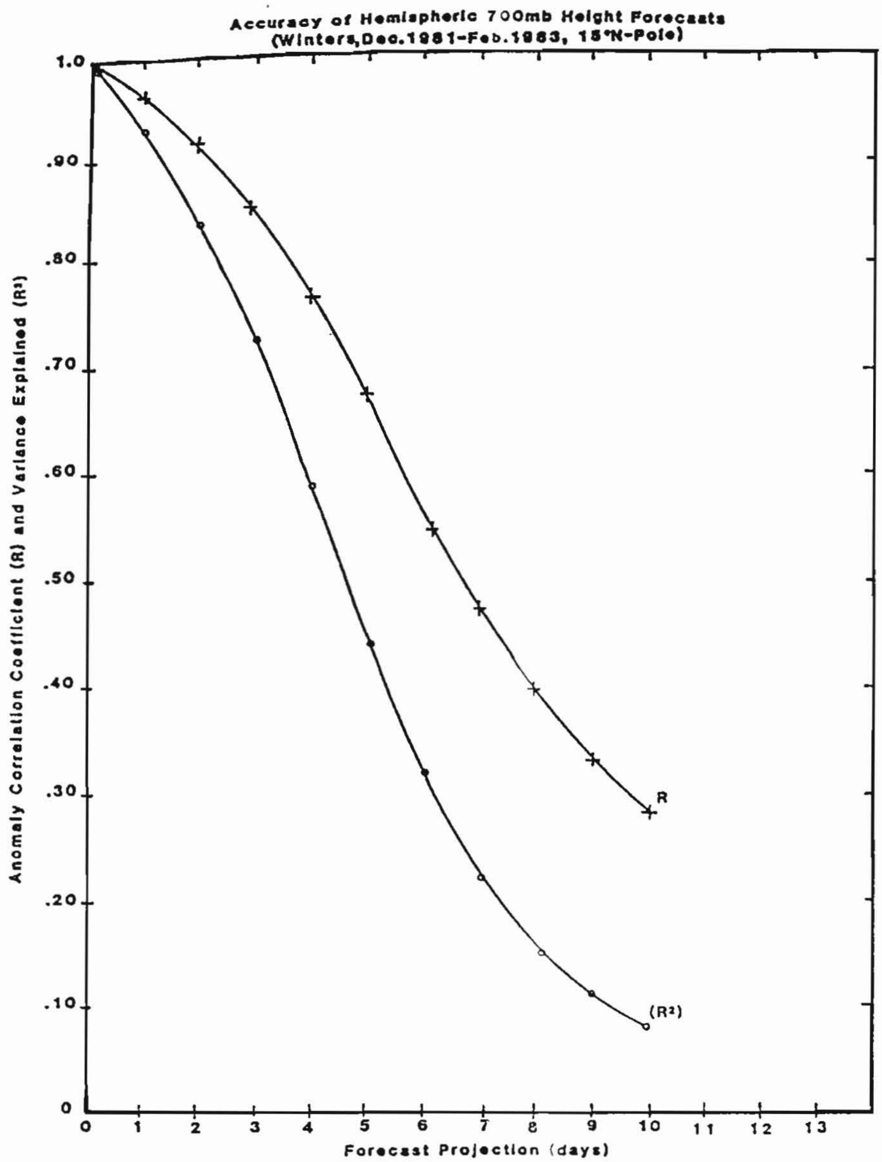

Fig. 15. Mean values of linear correlation coefficient (R) and variance explained (R^2), average over 512 grid points in the Northern Hemisphere at intervals of 5° latitude and 10° longitude, between forecast and observed anomalies of daily 700 mb height fcr 1-10 day predictions made by the NMC spectral model on 180 winter days from Dec. 1, 1981 to Feb. 18, 1983.

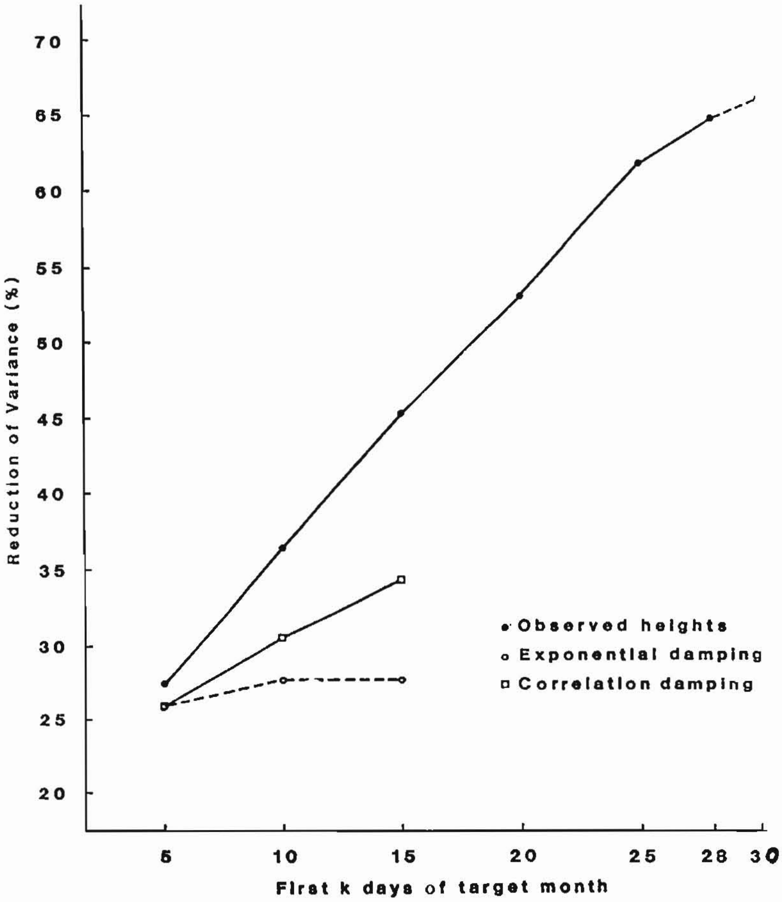

Fig. 16. Same as Fig. 6 (right) for various periods after target month begins for observed heights (long curve) and for prognostic heights simulated by exponential damping (dashed) or correlation weighting (solid).

All results given up to this point have assumed zero lead time; i.e., that the forecasts for SAT are made on the last day of the previous month. Table 5 summarizes results that would be obtained if the forecasts were made either 10 or 15 days earlier. In both cases antecedent predictors produce very poor results (RV's less than 3%). However, use of prognostic means simulated by the correlation method gives more respectable values, with an RV of almost 14% for a 10-day lead time.

Period of heights	RV(%) averaged for 50 stations	No. of significant stations
	(a) For data available 15 days before 1st day of target month	
Previous 15-day mean	0.5	2
Prognostic 15-day mean	7.5	22
	(b)*For data available 10 days before 1st day of target month*	
Previous 5-day mean	2.6	8
Previous 10-day mean	0.9	3
Prognostic 10-day mean	13.7	34

Table 5. Results of screening monthly mean temperature at 50 surface stations in the contiguous United States in Jan., Feb. and March, 1951-1980, as a function of 700 mb heights at 56 grid points over North America and adjacent oceans during various antecedent periods. All data are expressed as standardized anomalies and all significance is determined by a Monte Carlo technique. Prognostic means are obtained by weighting daily numerical forecasts by their correlations with observed heights from 1981 to 1983, assuming zero correlation at day 15.

6. Conclusion

This paper has described a series of statistical screening experiments designed to improve the accuracy of monthly mean surface air temperature (SAT) forecasts in the United States. Our principal results apply only to linear effects on 50 cities in the winter months of Jan., Feb. and Mar. during the 30 years: 1951-1980. They may be summarized as follows:

1. Only about 6% of the variance of SAT can be explained by the previous month's field of 700 mb height. This figure can be increased to almost 9% by adding the fields of last month's SAT and snow cover as potential predictors.

2. About 2/3 of the SAT variance can be specified from the field of concurrent mean 700 mb height. About 2% can be added by including the antecedent month's local temperature and another 2% by screening the entire field of previous SAT.

3. ENSO variables make no significant improvement to the results of 1. and 2. above. Thus, their effects on temperature are produced primarily through the circulation pattern.

4. The percent of SAT variance explained by heights for periods from 5 to 28 days *before* the beginning of the target month decreases with length of averaging period, indicating that recent data are more useful than older data.

5. The SAT variance explained by heights for 5-28 days *after* the target month begins increases with length of averaging period to a maximum at 30 days (specification). However, all shorter periods are better than the previous month's mean.

6. Previous heights in the form of antecedent or centered means make little contribution to SAT not already contained in future heights alone.

7. About 36% of the SAT variance can be explained by a perfect height forecast for the first ten days of the target month.

8. Simulating the skill of 1-10 day numerical predictions of 700 mb height indicates that they should explain at least 25% of the SAT variance, greater than any value for past heights.

9. If monthly forecasts have to be made about two weeks before the target month begins, about 10% of the SAT variance could be explained by use of daily numerical height prognoses.

Acknowledgement

This work was supported by the U. S. Department of Commerce, NOAA, National Weather Service, under Grant No. ATM-84AA-H00026, to the Cooperative Institute for Climate Studies, and by the U. S. Department of Energy under Contract No. DE-AS05-81EV10539. We thank the following individuals for data and assistance: Bob Livezey and Norm Maisel at CAC, Bob Kaylor, Güay-Hong Chen and Huug van den Dool at the University of Maryland, and John Walsh at the University of Illinois.

REFERENCES

Angell, J. K. and J. Korshover, 1983. Global temperature variations in the troposphere and stratosphere, 1958-1982. *Mon. Wea. Rev.,* **111**, 901-921.

Barnett, T. P. and R. Preisendorfer, 1987. Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.,* **115**, 1825-1850.

Donn, W. L., R. Goldberg and J. Adem, 1986. Experiments in monthly temperature forecasting. *Bull. Amer. Meteor. Soc.,* **67**, 165-169.

Epstein, E. S., 1987. How systematic are systematic errors? Preprints, *Eighth Conference on Numerical Weather Prediction,* Baltimore, MD, Amer. Meteor. Soc., pp. 460-465.

Gilman, D. L., 1983. Predicting the weather for the long term. *Weatherwise,* **36**, 290-297.

Harnack, R. T., C. Kluepfel and R. Livezey, 1986. Prediction of monthly 700 mb heights using simulated medium-range numerical forecasts. *Mon. Wea. Rev.,* **114**, 1466-1480.

Klein, W. H., 1983. Objective specification of monthly mean surface temperature from mean 700 mb heights in winter. *Mon. Wea. Rev.,* **111**, 674-691.

Klein, W. H., 1985a. Space and time variations in specifying monthly mean surface temperatures from the 700 mb height field. *Mon. Wea. Rev.,* **113**, 277-290.

Klein, W. H., 1985b. Specification of monthly mean anomalies of surface air temperature in Canada and Alaska. *Atmosphere-Ocean,* **23**, 155-176.

Klein, W. H., and R. Yang, 1986. Specification of monthly mean surface temperature anomalies in Europe and Asia from concurrent 700 mb monthly mean height anomalies over the northern hemisphere. *J. Climatology,* **6**, 463-484.

Namias, J., 1953. Thirty-day forecasting. A review of a ten-year experiment. *Meteor. Monogr.,* **6**, Amer. Meteor. Soc., 83 pp.

Rasmusson, E. M. and T. H. Carpenter, 1982. Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Wea. Rev.,* **110**, 354-384.

Roads, J. O. and T. P. Barnett, 1984. Forecasts of the 500 mb height using a dynamically oriented statistical model. *Mon. Wea. Rev.,* **112**, 1378-1388.

Ropelewski, C. F. and M. S. Halpert, 1986. North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Mon. Wea. Rev.,* **114**, 2352-2362.

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

Walsh, J. E., 1984. Forecast of monthly 700 mb height: Verification and specification experiments. *Mon. Wea. Rev.,* **112**, 2135-2147.

Walsh, J. E., D. R. Tucek and M. R. Peterson, 1982. Seasonal snow cover and short term climatic fluctuations over the United States. *Mon. Wea. Rev.,* **110**, 1474-1485.

Weare, B. C., 1986. An extension of the El Niño index. *Mon. Wea. Rev.,* **114**, 644-647.

Weare, B. C., 1987. Relationships between monthly precipitation and SST variations in the tropical Pacific region. *Mon. Wea. Rev.,* **115**, 2687-2698.