Objective Winter Temperature Forecasts: An Update and Extension to the Western United States[1,2]

Robert P. Harnack

Department of Meteorology and Physical Oceanography, Cook College—New Jersey Agricultural Experiment Station, Rutgers—The State University of New Jersey, New Brunswick 08903

(Manuscript received 13 July 1981, in final form 8 December 1981)

ABSTRACT

This study expands on previous studies (Harnack and Landsberg, 1978; Harnack, 1979) in that objective, statistical winter temperature forecast models are tested and verified for three additional winters (1979–81); models have been formulated and tested for the first time for the western one-third of the United States; new models have been formulated and tested using predictors defined for October; and Northern Hemisphere sea level pressure (SLP) has been tested as a predictor. An independent sample was used to test regression models.

The main results include: 1) Among the "November" type predictor models, the sea surface temperature (SST) only model continues to be superior to the others when tested on independent data, including those models using circulation predictors, and it performs significantly better than chance expectation. 2) The SST-only model performed much better in the eastern two-thirds of the United States than in the western third. 3) November SLP did not contribute to skillful winter temperature prediction as assessed by applying a model to 17 independent cases. 4) The October SST-only model showed slight skill relative to random chance but not compared to persistence. The other "October" type prediction models showed no skill. 5) The reliability of predictions increased considerably when only those predictions were verified in which the November SST-only model and persistence produced the same forecast. 62% of these forecasts have been correct (using three categories), which was superior to the performance of persistence alone (45% correct) or of the SST-only model alone (47% correct).

Geographical and yearly forecast performance differences are discussed as well as predictions made for the recent winter of 1980–81.

1. Introduction

Previous published papers (Harnack and Landsberg, 1978; Harnack, 1979) have reported on results of a winter temperature forecast experiment using objective methods applied to the eastern and central United States. The latter of these studies (hereafter referred to as H79) employed a variety of predictors (sea surface temperature, 700 mb height, and a Southern Oscillation Index), defined nine area-averaged winter temperatures for the region east of the Rocky Mountains as predictands, and used both empirical orthogonal function (EoF) analysis and screening multiple linear regression to formulate the forecast models. A variety of forecast models were formulated and tested, differing from one another in terms of predictor types available in the predictor pool and/or the length of winter period used to define the predictands, which ranged from one month (December only) to four months (December–March). Most predictors were defined as November means. Independent testing of the models was performed for all winters in the period 1973–78 (six winters). The principal conclusions of the H79 study are summarized as follows:

1) The forecasts improved as the length of winter period increased, when the circulation-plus-SST-component model was tested in this way. The predictions for December were correct 20% of the time, while the predictions for December–March were correct 56% of the time when forecasts were verified in a three-category format.

2) When the same winter length was used to define the predictands (i.e., December–February), but different predictor groupings were set aside for formulating the models, the SST-only model performed best (56% correct). It also performed significantly better than chance (33% at the 1% significance level using a binomial test and was the only one which performed better than a persistence control. Predictor groupings included circulation only, SST only, and combined circulation-plus-SST.

---

[1] Paper of the Journal Series, New Jersey Agricultural Experiment Station, Cook College, Rutgers—The State University of New Jersey.

[2] Portions of this paper were presented at the AMS sponsored First Conference on Climate Variations held 19–23 January 1981 in San Diego, California.

0027-0644/82/040287-09$06.25

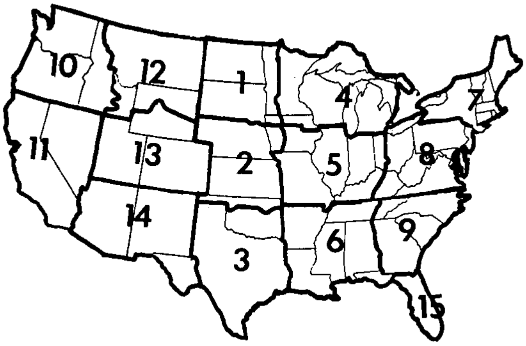

FIG. 1. Location of subareas with identifying numbers used in defining mean winter temperatures.

The current study expands on the previous studies in a number of ways. First, models now have been tested and verified for three additional winters (1979–81), bringing the independent sample size up to nine. Second, models have been formulated and tested for the first time for the western third of the United States. Third, models have been formulated and tested using predictors defined for October, instead of just November, so that the reliability of forecasts with an additional one month of lead time could be assessed. Finally, a much larger data sample was used to enlarge both the dependent and independent samples for formulating a new type of regression model by employing Northern Hemisphere gridded monthly-mean sea level pressure (SLP) available for the period 1899–78.

For the reader interested in a review of recent long range forecasting progress, the papers by Nicholls (1980) and Harnack (1981) should be consulted.

2. Data set description

Since the data sets employed and statistical methods used are described in detail in the H79 study, this discussion will be limited to summarizing the main aspects of the model formulation and testing, and to detailing only the extension of previous work.

The following data sets were used to define predictor variables for regression model types formulated for the H79 study and tested through the winter of 1980–81:

1) Mean November sea surface temperature (SST) for the stationary Atlantic Ocean Weather Ships (OWS) for the period 1949–80.

2) Mean November SST data for the eastern North Pacific Ocean for the period 1947–80. 5° by 5° latitude-longitude grid-point values for the area bounded by 25°N–55°N and 125°W–160°E were extracted from the Scripps Institution of Oceanography data set.

3) Mean November SST data (5° by 5° grid) for the eastern tropical Pacific Ocean (10°N–5°S, 85°W–165°W) for the period 1949–80. The reader should refer to H79 for further description.

4) Mean November Northern Hemisphere 700 mb heights on a 10° by 10° grid in the region bounded by 30°N–70°N and 0–180°W for the period 1947–80. This data set was obtained from the National Meteorological Center.

5) A Southern Oscillation (SO) index, as developed by Wright (1975), for the fall season (August–October) in the period 1949–80.

The foregoing data sets were used in various combinations, as in H79, to define the predictor pool available for regression analysis in order to formulate similar type regression equations for five western United States subareas and also for the Florida peninsula. These areas had not been included in the work reported earlier. Fig. 1 shows the locations and identifying numbers for all subareas used in this study. Subareas 1–9 were used in the H79 study, while subareas 10–15 were added in this one. Area-averaged temperatures for winter months (December–February) were determined for each of these subareas from historical, monthly-mean state climatic division temperature data obtained from the National Climatic Center. Area-averaged values, using all climatic divisions in the subarea, were computed for each winter month then values for the winter season were determined by simple arithmetic averaging of monthly values. Subarea boundaries were chosen arbitrarily. Temperature averages for these 15 subareas constituted the predictands.

In addition to expansion of the study to include the West and the Florida penninsula, new regression models were formulated using the same predictor types (except Atlantic OWS SST, which was becoming increasingly difficult to obtain), but defining the predictor variables for the month of October. Therefore, with the exception noted, October mean data of the same type, characteristics and source were used in this study as was used in the previous study when predictors were defined for November. In the "October" predictor models, the SO index was defined for the May–July period.

One additional type of data was used in a portion of this study: Northern Hemisphere sea level pressure data. November monthly-mean values for the same grid as that used for the 700 mb height data were extracted from a magnetic tape of SLP acquired from the National Center for Atmospheric Research. Corrections were made to this data as suggested by Trenberth and Paolino (1980), which were supplied by magnetic tape from the National Center for Atmospheric Research.

3. Procedure

For both the October type and November type predictor models, EoF analysis was performed separately on each of the data fields described in Section 2, as was done in the H79 study. This considerably reduced the number of predictors in the predictor pool available for screening by the regression analysis. The number of EoF's retained was arbitrarily set so that at least 80% of the variance for each field was explained. This criterion was a carryover from the H79 study, after which more sophisticated criteria have been employed for EoF analyses.

The Atlantic OWS SST variables were reduced from 7 variables to 2, the North Pacific SST from 50 variables to 7, the Eastern Tropical Pacific SST from 34 variables to 5, and the Northern Hemisphere 700 mb heights from 95 variables to 8. Further details of and rationale for the EoF analysis can be found in H79, especially Table 1. Important EoF spatial patterns are shown in that study as well.

In the current study, EoF analysis was additionally applied to the Northern Hemisphere SLP which also resulted in a reduction of predictor variables of this type from 95 to 8.

After the EoF's were extracted for each of the predictor fields, the resulting time series sets of amplitudes plus the SO index were used as predictors, together with the previously described set of predictands (mean winter temperatures for subareas) as input to multiple linear regression analysis. Regression models had been previously formulated for subareas 1–9 using November predictors in H79, but were formulated in this study for subareas 10–15 in the case of November models and for all subareas in the case of October models. As indicated in the H79 study (Tables 7 and 8) North Pacific SST EoF's tended to be selected more than either tropical Pacific or Atlantic SST EoF's.

TABLE 1. November-type winter temperature forecast models formulated and tested.

Model name	Predictor types	Number of predictors
Circulation and SST	Northern Hemisphere 700 mb heights (Nov)	8
	Pacific SST (Nov)	7
	Tropical Pacific SST (Nov)	5
	Atlantic OWS SST (Nov)	2
	Southern Oscillation Index (Aug–Oct)	1
SST only	Pacific SST (Nov)	7
	Tropical Pacific SST (Nov)	5
	Atlantic OWS SST (Nov)	2
Circulation only	Northern Hemisphere 700 mb heights (Nov)	8
	Southern Oscillation Index (Aug–Oct)	1
SLP	Northern Hemisphere sea level pressure (Nov)	8
	Southern Oscillation (Aug–Oct)	1

TABLE 2. October-type winter temperature forecast models formulated and tested.

Model name	Predictor types	Number of predictors
Circulation and SST	Northern Hemisphere 700 mb heights (Oct)	8
	Pacific SST (Oct)	7
	Tropical Pacific SST (Oct)	5
	Southern Oscillation Index (Jun–Jul)	1
SST only	Pacific SST (Oct)	7
	Tropical Pacific SST (Oct)	5
Circulation only	Northern Hemisphere 700 mb heights (Oct)	8
	Southern Oscillation Index (Jun–Jul)	1

The sample data used for the regression analysis (the dependent sample) consisted of the period 1950–72, or sample size of 23 years for most of the model types formulated. The exception was the model employing mainly November SLP EoF's as predictors, in which case the dependent sample numbered 63 years (1899–1961). For each model type formulated, a separate regression equation containing eight predictors was developed for each subarea. These predictors were screened from the pool of predictors shown in Table 1.

There were four November model types, varying only by the types of predictors in the predictor pool, which were formulated and then tested on the independent data sample. These model types are described in Table 1. The first three model types are the same as employed in the H79 study, except in the current study models were tested for an additional three winters for the eastern and central United States and for all nine independent sample winters for the western United States. These three model types represent an attempt to do comparative testing of models involving mean November information of just SST, just circulation, and a combination of both.

The fourth model type described in Table 1 was used in order to assess the effect of using a much larger data sample in model formulation and testing. Sea level pressure is the only plausibly useful variable, defined over a planetary scale domain and with a sufficient quantity of observations going into the means, that has a data sample exceeding 40 years.

There were three October model types formulated (Table 2). These follow along the lines of the first three November model types except all variables, except the SO index, are defined for October. The rationale for extending previous studies to a longer lead time in making winter forecasts was the encouraging results found for the November models and the obvious benefit that could accrue by finding skillful models with longer lead times.

TABLE 3. Verification scores (mean percent correct) using November-type winter temperature forecast models applied to an independent sample.

Predictor model	Number of predictors	Years	Mean percent correct using three categories	Mean percent correct using two categories
(November regression models)				
Circulation and SST	23	1973–81	34	51
Circulation only	9	1973–81	28	49
SST only	14	1973–81	47	62
SLP and SO	9	1962–78	31	41
November temperature (persistence)	1	1973–81	45	61

Each equation of each model type was applied to an independent sample consisting of the period 1973–81 (nine years) for the first three November model types, 1962–78 (17 years) for the SLP model, and 1973–79 (seven years) for the October model types. The varying periods used for the independent testing of the November and October model types were a result of data availability plus the fact that initial testing of the October models gave results that did not justify testing the models each year. On the other hand, the November models were tested in real-time, after 1976, since real-time winter temperature forecasts were desired. These were based on objective model forecasts made at the end of November.

4. Results

Table 3 gives the mean percent correct for all independent sample forecasts made for all 15 subareas by the November type models. The number of predictors in the pool is also shown for each model. The mean percent correct is given for forecasts made in terms of three categories (below, near, and above normal) *and* for forecasts made in terms of two categories (below and above normal). The category limits for three categories were determined by ranking observed winter temperatures from the period 1941–70 for each subarea and then dividing the data into terciles. For the two-category verification, limits were determined by finding the winter temperature mean for each subarea. Random chance dictates that about 33% of the three category type forecasts are expected to be correct while about 50% of the two-category type forecasts are similarly expected to be correct by chance alone. Only the SST model rises much above the expected level, in line with previous results. For the three-category verification, the 47% correct figure is down from the 56% reported in H79. However, not only does the updated figure include three additional recent winters, but also includes verifications using the Western subarea equations for the first time. More will be said on the effect of including the West later on. The verification for a persistence control is included in Table 3 as well. Again, only the SST model is competitive, showing only a slight, insignificant advantage over persistence overall. Later, it will be shown that there are important east–west differences in the relative scores.

A *t*-test was performed to determine if the mean percent correct for each model including persistence was significantly better than chance expectation. The *t*-statistic was computed from $t = (\bar{P}_0 - \bar{P}_e)/(S_p^2N^{-1})^{1/2}$, where $\bar{P}_0$ is the observed mean percent, $\bar{P}_e$ the expected mean percent correct, N the number of cases, and S_p^2 the standard deviation estimated from

$$S_p^2 = \sum_{i=1}^{N} (P_i - \overline{P_0})^2/(N-1), \tag{1}$$

where P_i is the percent correct in the ith year.

The expected mean percent correct ($\bar{P}_e$) was set at 33% for the three-category verification. The computed *t*-statistic was then compared to the *t*-value taken from the *t*-distribution with $N - 1$ degrees of freedom in order to determine significance.

The *t*-test was applied in this study with the assumption that the winters were independent cases. This is not an unreasonable assumption considering the work of Madden (1977), who found that the spectra of winter season temperatures for most locations in North America resembled a "white noise" spectra. It should be noted that this significance test does not make any assumption about independence of winter temperatures between geographical areas, only independence between winters. Only the November SST predictor model performed significantly better than chance at the 5% level, including persistence.

The new type November model developed in the current study, which employs Northern Hemisphere SLP EoF's plus the summer SO index, did not perform better than chance for the 17 independent sample winters used. This result plus the lack of skill relative to chance for the circulation only model, which employed 700 mb height EoF's plus the SO index, suggests that antecedent circulation predictors are not useful for directly predicting winter temperature anomalies. SST predictors, however, continue to be useful. More on this point in later sections.

In order to see if there are important regional differences in the verification scores, Table 4 is presented. This stratifies the mean percent correct values

TABLE 4. West/East stratified verification scores (mean percent correct) using November-type winter temperature forecast models applied to an independent sample.

Predictor model	Number of predictors	Years	Mean percent correct using three categories (west/east)	Mean percent correct using two categories (west/east)
(November regression models)				
Circulation and SST	23	1973-81	31/36	47/53
Circulation only	9	1973-81	29/27	44/51
SST only	14	1973-81	36/51	58/63
SLP and SO	9	1962-78	32/31	38/44
November temperature (persistence)	1	1973-81	49/43	69/57

by West (subareas 10–14) versus East (subareas 1–9 and 15). Important differences show up for the SST-only model. The East score for the three-category verification is only slightly down (from 56 to 51%) when the three most recent winters are used in addition to the earlier six winters previously used as the independent sample. The regional comparison with persistence indicates that the SST-only model is superior in the East (51% vs 43%), while persistence is superior in the West (49% vs 36%). It should be noted that over the long term, fall to winter temperature persistence is significant for much of the West but much less so for the East (Namias, 1978a). A discussion of specific winter forecasts is reserved for later in this paper.

Table 5 gives the verification scores for the October model types. Surprisingly, the SST-only model still shows slight skill relative to random chance (42% vs 33%), although not relative to persistence (42% vs 47%). Given the relatively small sample sizes, these differences were not shown to be statistically significant. The other models show no skill when applied to this independent sample. The East/West stratification is shown in Table 6. Again, the SST-only model performs much better in the East, and about equal to persistence there.

The regional differences noted here between verification scores for the western third and eastern two-thirds of the United States, are contradictory to those found by Namias (personal communication). His semi-objective winter forecasts made for the United States have done much better for the far West than for the East. The primary predictor in his several step methodology is the antecedent and projected North Pacific SST field (Namias, 1976). In the current study, the SST-only models employed used North Pacific SST predictors extensively, so the difference between the Namias' results and this experiment is at first glance puzzling. Several possible explanations come to mind:

1) Namias uses stepwise multiple regression (screening) rather than eigenvector methods to derive his circulation and SST equations.

2) The subjective element in the Namias' forecasts not only allows for personal experience to benefit the final forecast product but also allows more flexibility and physical insight in applying large-scale air-sea interaction concepts to specific cases.

3) The forecasts made in this study were verified for sizeable area averages only, while Namias uses a 99-point grid for verification and makes the forecast map with sufficient detail to account for known terrain effects. Another possible contributing difference is the regional emphasis by Namias on the West in making the final forecast.

Fig. 2 gives the number of zero-, one- and two-category errors for each subarea when the November SST-only model was tested on the nine recent winters (1973–81). This again illustrates the superiority of eastern forecasting by this model, but obviously gives more detail. The smallness of the sample size prohibits any firm conclusions to be based on this anal-

TABLE 5. Verification scores (mean percent correct) using October-type winter temperature forecast models applied to an independent sample.

Predictor model	Number of predictors	Years	Mean percent correct using three categories	Mean percent correct using two categories
(October regression models)				
Circulation and SST	21	1973-79	27	50
Circulation only	9	1973-79	30	50
SST only	12	1973-79	42	60
October temperature (persistence)	1	1973-79	47	55

TABLE 6. West/East stratified verification scores (mean percent correct) using October-type winter temperature forecast models applied to an independent sample.

Predictor model	Number of predictors	Years	Mean percent correct using three categories (west/east)	Mean percent correct using two categories (west/east)
(October regression models)				
Circulation and SST	21	1973–79	37/21	46/51
Circulation only	9	1973–79	23/33	31/59
SST only	12	1973–79	29/49	54/63
October temperature (persistence)	1	1973–79	34/51	34/66

ysis but the overall best forecasts have been produced thus far for the southeast subarea (subarea 9) and the mid-Atlantic subarea (subarea 8), when the distribution of errors is considered. The Great Lakes region (subarea 4) is next best. The worst overall forecasts by this model have been produced for the Northern Rockies region (subarea 12). It is interesting to compare these results, for winter season forecasts, to those of Madden and Shea (1978) who estimated the potential predictability of *monthly* temperature by station (see their Fig. 3). Their results for January indicate the best *potential* predictability for an area stretching from the Gulf Coast to the mid-Atlantic states, where the SST-only model was generally performing best; and also in portions of the West (especially Northwest), where the SST-only model was mediocre. Perhaps estimates of *seasonal* temperature predictability would bring the two results even more into line, but only a much larger sample size for this study would be expected to result in *very* similar results.

In order to see the year-to-year performance characteristics of some of the models, Fig. 3 is presented. The number of correct forecasts by two of the November type models, using three temperature categories, are plotted on the graph for each winter tested. The SST-only model has been superior to the circulation-only model in seven of the nine years. Also the SST-only model has performed at least nominally better than chance expectation (i.e., five correct) in five of the nine years, with only one year below the expected value. For the circulation-only model, only two winters have been forecast better than expectation, but five have been forecast worse. It should be mentioned that while the sample size is rather small, the period considered contained a wide variety of temperature patterns, and several were quite anomalous over large areas (see Harnack, 1979, 1980).

The forecast made by the November SST-only model for the most recent winter of 1980–81 was moderately successful. The objective forecast map is shown in Fig. 4. The observed map, given in terms of subarea values, is shown in Fig. 5. The 'warm' West, 'cold' East aspect of the winter temperature pattern was handled well by this model, but the rel-

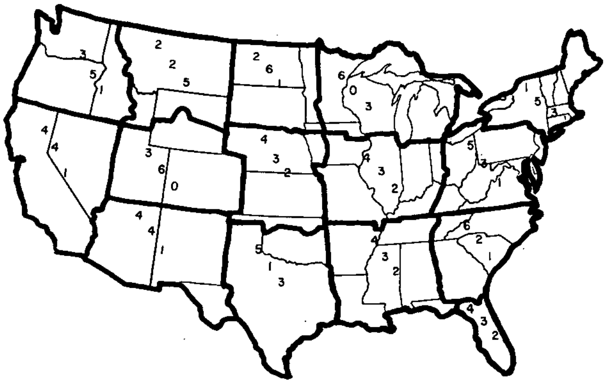

FIG. 2. Number of zero-, one- and two-category errors obtained by subarea when the November SST-only winter temperature prediction model was applied to the independent sample (1973–81).

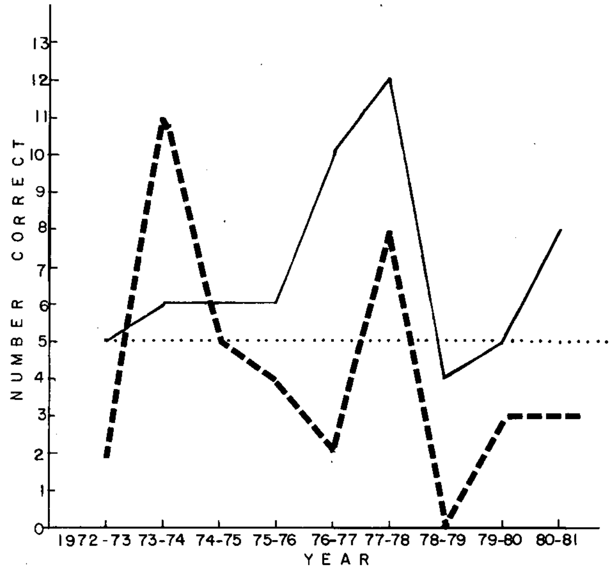

FIG. 3. Number of correct category forecasts obtained by year when the November SST-only model (solid) and circulation-only model (dashed) was applied to the independent sample. The total number possible correct is 15 and the expected number correct by chance is 5, which is denoted by a dotted line.

ative sizes of these temperature categories were reversed from forecast to observed, resulting in five subareas having two category errors. A subjective, experimental forecast made in real-time from the use of the objective forecasts, persistence and climatology is shown in Fig. 6. The corresponding observed map is shown in Fig. 7. This forecast fared better than the objective SST-only model forecast, since the two-category error area is greatly diminished. While subjective modification of objective guidance would not always result in improving the forecast, experience with the models tested so far suggests that this is so in general.

One mainstay of the subjective method used to finalize the forecast has been to rely *more* on those subarea forecasts made with the November SST-only model when the forecast category is the *same* as the temperature category forecast by the persistence control. For the nine test winters, this condition has been met 37% of the time. Of these instances, 62% have been correct (for three-category forecasts), which is better than either persistence (45%) or the model (47%) has done separately when evaluated over all forecast areas. By starting the forecast map with the forecasts for subareas having this type of agreement, then using just persistence in those areas where it tends to do well or just the SST-only model forecast where it has been doing well, the final forecast product can be mostly completed. In addition, experience with the coherence and teleconnection patterns of United States surface temperature is used to decide on the forecast for areas where there are contradictory indications. A study by Rasmusson *et al.* (1981) describes an objective analysis of United States surface temperature which clarifies these patterns. Antecedent circulation is not used since the circulation-only model has not shown any overall skill.

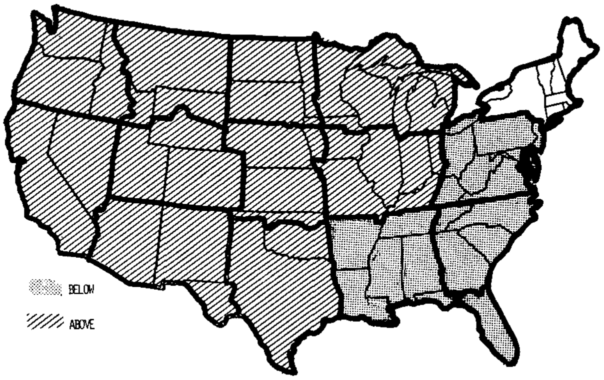

FIG. 5. Observed winter (December–February) temperature categories for 1980–81 by subarea.

A word or two is appropriate here regarding the nature of North Pacific large-scale air-sea interactions that may have taken place during the winter of 1980–81. During November 1980 a large "cold pool" (i.e., below normal SST anomalies) stretched from west of the dateline to about 145°W, while warmer water (i.e., above normal SST anomalies) were in place to the east. This resulted in a rather large east-west SST anomaly gradient from 140 to 150°W over the 30 to 50°N latitude belt. From all indications, by examining monthly SST anomaly maps contained in the NOAA publication *Oceanographic Monthly Summary*, this situation persisted into the winter of 1980–81. The overlying atmospheric circulation over the North Pacific during much of this winter, especially in January, was similar to that observed during the winters of 1976–77 and 1977–78 which are discussed elsewhere (Harnack, 1980). The anomalous component of flow at 700 mb during the late fall and early to mid-winter had a pronounced southerly component to it in the 130–150°W sector. This association between anomalous east-west SST anomaly gradients and anomalous meridional flow at 700 mb was found to be important in winter in an earlier study (Harnack and Broccoli, 1979), and was evident during the late 1980, early 1981 period. Numerous case studies (e.g.,

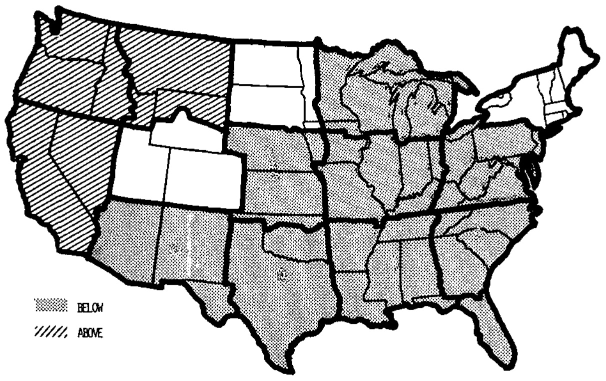

FIG. 4. Forecast winter (December–February) temperature categories for 1980–81 using the November SST-only prediction model.

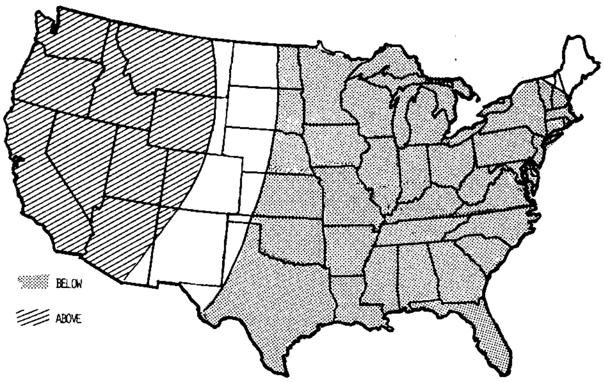

FIG. 6. Subjective Harnack winter (December–February) temperature category forecast for 1980–81.

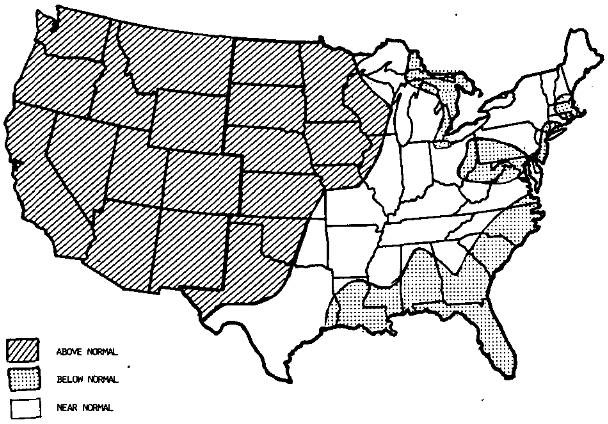

FIG. 7. Observed winter (December–February) temperature category map for 1980–81.

Namias, 1978b; Harnack, 1980) and the earlier study (Harnack, 1979), which the current study extends, explained how anomalous southerly flow in the eastern North Pacific (i.e., 130–150°W) results in an amplified downstream ridge in the West, trough in the East situation over North America. This in turn usually gives the West a rather mild winter and the East a rather cold one. It is apparent that the described type of large-scale air-sea interaction over the North Pacific and circulation teleconnection from the eastern North Pacific to North America was operating during the winters of 1976–77, 1977–78 *and* 1980–81. Of course, the mean circulation during these winters had some dissimilarities as well, and have been largely unexplained. Nevertheless, the SST-only model has done rather well in predicting the generally warm West, cold East temperature pattern in each of these winters (see Fig. 3). As discussed in the H79 study (see Table 8), two of the most important predictors in the regression equations for the SST-only model are the second (FP2) and third (FP3) EoF's of the November North Pacific SST field. The FP2 amplitude for November 1980 was large negative, since the SST anomaly field was similar to that presented in Fig. 14 of the H79 study (i.e., cold water in the central North Pacific, strong east–west SST anomaly gradient in the eastern North Pacific, and warm water just off the West Coast). A negative FP2 amplitude *contributes* generally to a prediction of a cold winter in the East and a warm winter in the West when the SST-only model is employed. The case studies, including physical reasoning, indicate the reasonableness of this objectively derived relationship.

The FP3 spatial coefficient pattern combined with the correlation pattern of FP3 with winter temperatures in the eastern United States (see Figs. 18 and 19 of H79) present too complex a situation for a simple physical explanation to be useful in describing events of the winter of 1980–81. All that can be said is that cold water over a large region of the North Pacific contributes in balance to a *negative* FP3 amplitude, which in turn contributes *statistically* to a winter prediction of below normal temperatures in the eastern third of the United States.

In summary, the November 1980 North Pacific SST anomaly field had much in common with that of November 1976 and November 1977 which resulted in the objective predictions by the SST-only model of winter temperatures for the United States for each of the following winters to be similar. These forecasts were rather successful since (i) the major characteristics of the SST anomaly field persisted from November into winter in each of these cases, (ii) the overlying atmospheric circulation for winter was highly meridional over the eastern North Pacific in a sense expected from earlier studies, (iii) the downstream circulation over North America was highly amplified with a pronounced West Coast ridge and deep East Coast trough, and (iv) the associated surface temperature pattern, as expected, showed a characteristic warm West, cold East pattern. Obviously, the situation described has occurred enough in the dependent sample cases (1950–72) that application of the SST-only model to these independent cases has resulted in useful predictions of winter temperatures without explicitly incorporating all of the physical linkages that may have been operating.

One last point to make about model performance on the independent sample is that for operational usage, statistical models like those developed here could be updated each year, so that sample regression coefficients approach their population values with time. The models should be better "tuned" by this procedure and perhaps perform better than indicated here.

5. Conclusions

Previous work on objective winter temperature forecast modelling was extended geographically to include the western third of the United States; models were tested on the three most recent winters (1979–81); new, but statistically similar models were formulated and tested using mainly October data; and the long-term Northern Hemisphere sea level pressure data set was used for the first time in objective winter season forecast models.

The main results of this study may be summarized as follows:

1) The November type predictor model which uses only sea surface temperature predictors continues to be superior to the others when tested on independent data, including those models using circulation predictors, and it performs significantly better than chance expectation (at 5% level using binomial test as in H79 study). The SST-only model is also competitive with a persistence control overall

(47% vs 45% correct using three temperature categories). It should be noted that no other model tested was deemed to be significantly better than chance at the 5% level including persistence.

2) The SST-only model performs much better in the eastern two-thirds of the United States than in the western third (51% vs 36%). It performs better than persistence in the East, but worse in the West.

3) The model employing November sea level pressure as the predictors performed poorly (31%) relative to random chance (33%), when applied to independent data, thereby reinforcing the opinion that the sole use of primarily circulation type predictors for objective predictions of winter temperature is not warranted.

4) Among the October type predictor models formulated and tested, only the SST-only model showed slight skill (42%) relative to random chance overall, although not relative to persistence (47%).

5) The reliability of predictions increased considerably when only those predictions were verified in which the November SST-only model and persistence produced the same forecast. 62% of these forecasts have been correct, with the correct condition being met 37% of the time.

6) The geographical distribution of category errors by the November SST-only model had some similarity to that expected from predictability studies such as that of Madden and Shea (1978).

7) The November SST-only model performed moderately well for the winter of 1980–81, which it had for the somewhat similar winters of 1976–77 and 1977–78. It was reasoned that similar large-scale air-sea interactions took place over the eastern North Pacific Ocean during each of these winters.

The current study reinforces the idea that North Pacific sea surface temperature should be used for making winter temperature forecasts over the United States, at least for forecasts produced at the end of November, while use of antecedent circulation should be deemphasized. Of course, given the relatively small size of the independent sample (nine winters) additional testing is necessary in order that complete confidence may be placed in the results obtained. A more systematic use of Atlantic SST might further improve this type of forecast.

Acknowledgments. This study was made possible by support provided by the Climate Dynamics Program, Division of Atmospheric Sciences, National Science Foundation under Grant ATM-78-15434 and as a part of NJAES Project No. 13405, supported by the New Jersey Agricultural Experiment Station. The author wishes to thank Ms. Cathy Biddulph for general assistance, Ms. Jeremi Harnack for drafting, Mr. John Lanzante for computer programming, and Ms. Ruth Smith for typing the manuscript.

REFERENCES

Harnack, R. P., 1979: A further assessment of winter temperature predictions using objective methods. *Mon. Wea. Rev.*, **107,** 250–267.

——, 1980: An appraisal of the circulation and temperature pattern for winter 1978–79 and a comparison with the previous two winters. *Mon. Wea. Rev.*, **108,** 37–55.

——, 1981: Principles and methods of extended period forecasting in the United States. National Weather Association Monograph, 1-81. 34 pp. [Available from National Weather Association, 4400 Stamp Rd, Marlow Heights, MD 20031].

——, and A. J. Broccoli, 1979: Associations between sea surface temperature gradient and overlying mid-tropospheric circulation in the North Pacific region. *J. Phys. Oceanogr.*, **9,** 1232–1242.

——, and H. E. Landsberg, 1978: Winter season temperature outlooks by objective methods. *J. Geophys. Res.*, **83,** 3601–3616.

Madden, R. A., 1977: Estimates of autocorrelations and spectra of seasonal mean temperatures over North America. *Mon. Wea. Rev.*, **105,** 9–18.

——, and D. J. Shea, 1978: Estimates of natural variability of time-averaged temperatures over the U.S. *Mon. Wea. Rev.*, **106,** 1695–1703.

Namias, J., 1976: Seasonal forecasting experiments using North Pacific air/sea interactions. *Preprints 6th Conf. Weather Forecasting and Analysis*, Albany, Amer. Meteor. Soc., 13–16.

——, 1978a: Persistence of seasonal temperatures up to one year. *Mon. Wea. Rev.*, **106,** 1557–1567.

——, 1978b: Multiple causes of the North American abnormal winter 1976–77. *Mon. Wea. Rev.*, **106,** 279–295.

Nicholls, N., 1980: Long-range weather forecasting: value, status and prospects. *Rev. Geophys. Space Phys.*, **18,** 771–788.

Rasmusson, E. M., P. Arkin, W. Y. Chen and J. B. Jalickee, 1981: Biennial variations in surface temperature over the United States as revealed by singular decomposition. *Mon. Wea. Rev.*, **109,** 587–598.

Trenberth, K. E., and D. A. Paolino, 1980: The Northern Hemisphere sea level pressure data set: trends, errors and discontinuities. *Mon. Wea. Rev.*, **108,** 855–872.

Wright, P. B., 1975: An index of the Southern Oscillation. Climatic Research Unit Report 4, University of East Anglia, Norwich, England, 22 pp.