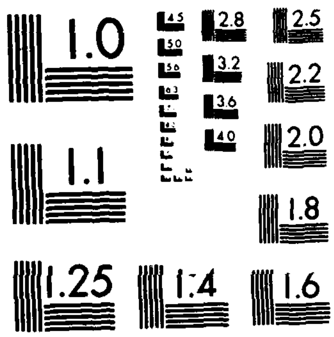
AD-A143 576 CLIMATE PREDICTION PART 1 CYCLONE FREQUENCY(U) VIRGINIA UNIV CHARLOTTESVILLE DEPT OF ENVIRONMENTAL SCIENNES B P HAYDEN JUL 84 TR-30 N00014-81-K-0033 1/2

UNCLASSIFIED F/G 4/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

CLASSIFICATION OF COASTAL ENVIRONMENTS

TECHNICAL REPORT 30

CLIMATE PREDICTION
PART I: CYCLONE FREQUENCY

BRUCE P. HAYDEN

DEPARTMENT OF ENVIRONMENTAL SCIENCES
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA

JULY 1984

OFFICE OF NAVAL RESEARCH
COASTAL SCIENCES PROGRAM

CONTRACT NO. N00014-81-K-0033
TASK NO. 389-170

CLASSIFICATION OF COASTAL ENVIRONMENTS

TECHNICAL REPORT 30

CLIMATE PREDICTION

PART I: CYCLONE FREQUENCY

BRUCE P. HAYDEN

DEPARTMENT OF ENVIRONMENTAL SCIENCES
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA

JULY 1984

OFFICE OF NAVAL RESEARCH
COASTAL SCIENCES PROGRAM

CONTRACT NO. N00014-81-K-0033
TASK NO. 389-170

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to make acknowledgements in four areas: data collection, model construction and development, forecasting and report generation. The hours spent by Phyllis Ross, Stephanie Lent and Ronald Broome are greatly appreciated. William Smith is largely responsible for the development of the computer programs needed to do the analyses and make the forecasts. Forecasts for the period of hindcast trials and operational forecasting were done by Ron Broome, Bob Wayland, Richard Wayland and Todd Hawes. Graphics were prepared by William Smith and Richard Wayland. Finally special thanks are in order for Wilma LeVan who, as always, makes the seemingly endless task of report generation bearable.

ABSTRACT

Climate prediction models based on multivariate analyses of cyclone frequencies are constructed from historical data (1885-1960) and evaluated for forecast skill on independent data (1960-1983). Cyclone frequencies are predicted for six-month duration seasons at 87 locations over eastern North America and the western North Atlantic from 27.5° to 55°. Three types of principal components models are constructed and tested. Model I uses unrotated principal component axes; Model II uses rigid rotation of the component axes; and, Model III uses oblique rotations of the component axes.

Forecast skill averages 75% correct for 2 category measure of forecasts. Skill based on a chance model would yield only a 50% score. Magnitude forecast skill is also demonstrated. No seasonal "cycle" in forecast skill is noted, i.e., all seasons are predicted with about the same level of skill. Forecast skills are highest off the east coast of the U.S., southern Canada, the northern plains of the U.S. and over the southwestern part of the U.S. east of 100° W. No trends in skill scores are found over the 1960-1983 period of forecast trials.

The various rotation schemes have little effect on overall model or global skill but there are some local skill differences, i.e., there are some areas that are better forecast with one rotation version than with the others.

Cyclone frequency climates are shown to be predictable in the time scale of six-month duration seasons. Forecast skills exceed those reported for any other climate variable.

INTRODUCTION

The University of Virginia Climate Forecast Model (Hayden and Smith, 1982) is based on multivariate analyses of cyclone frequencies. Spatial fields of cyclone frequencies are predicted for six-month duration seasons. The model covers eastern North America and the western North Atlantic. Predictability is due to season-to-season persistences in the spatial patterns of the frequencies of cyclones. Strong persistences in storm frequency and track are found from one six-month period to the next. Hayden and Smith (1982) showed that the model out-performed chance, simple persistence, damped persistence, and climatology as forecasts. Evaluation of a battery of forecast skill scores indicated there was predictability of both the sign and the magnitude of the anomalies. Hayden (1981a) used a jackknife procedure to generate forecasts for the 95-year period of record by withholding different periods for independent data forecast trials. No secular trends in model skill were found. It was assumed that the model was stable and not just a quirk of the particular dependent data period.

In the forecast model about 50% of the variance is involved in cyclone frequency patterns that persist, in some measure, from one season to the next. Based on a 2-by-2 test to evaluate forecast skill, i.e., tests of forecasts of above or below the long-term mean, a 75% skill score is expected and achieved. A 50% skill could be achieved by coin flipping. Several versions of the original cyclone frequency forecast model (Hayden and Smith, 1982) have since been constructed and tested on the same independent data period and on four years of operational forecasting. This report details the new model versions and the extensive forecast trials on independent data (hindcast and operational forecast modes). Details of the model are found in Hayden and Smith (1982). The present report supplements that earlier work.

BACKGROUND

Lorenz (1973) stated, "Regardless of what might be indicated by theory, a conclusive proof that partial predictability exists at a given range would be afforded by any demonstration that at least one forecasting procedure exhibits skill at that range." This rather pragmatic view of the prediction problem is especially appropriate where climate prediction is concerned. To date, theoretical studies indicate a limit of about two weeks for the time domain of weather forecasting. Where longer range predictions of the "state" of the atmosphere are concerned, specificity in the time domain must be relinquished. The prediction objective then becomes the specification of the average state of the system for some suitable time interval (month, season, year, etc.). With this modified prediction objective in mind, the required techniques become more stochastic and less deterministic. This necessity is augmented by the fact that suitable theories permitting deterministic forecast models for months, seasons and years are not available at present.

In the absence of a deterministic basis for climate forecasting, one is left with the need to identify some mode of persistence in the atmospheric system such that knowledge about the current and recent states of the atmosphere permits estimation of future conditions. Most efforts to identify such persistences in temperature and precipitation data-time series have failed or the magnitude of the resulting forecast skill is so small, and the number of forecast trials so few, that it is impossible to distinguish the forecast model from a model based on chance. The Climate Analysis Center's monthly and seasonal forecasts are based on persistences in the thickness fields. The perception is that the general circulation may exhibit persistences that are not apparent in station temperature and rainfall. The research group at the Scripps Institution under Jerome Namias' direction base their predictions on the persistences in sea-surface temperature fields which, in turn, serve as a "memory" for the atmosphere through thermodynamic couplings. Our work at the University of Virginia is based on identified persistences in the fields of occurrences of cyclones over eastern North America (Hayden and Smith, 1982). It is clear that occurrences of cyclones are not independent of structure or thickness fields so our work is in some sense like that of the Climate Analysis Center but the forecasts do not always agree so real differences exist.

Over the last several years we have completed an extensive forecasting and verification effort. This report summarizes the results of this effort. We are convinced that sufficient success has been demonstrated that Lorenz's (1973) criterion of "conclusive proof" has been fully met and we can advance the theory that *climate is at least partially predictable*. Equally important, however, is the need to study the causes of the persistences and the nature of failures in persistence. This awaits further work.

The approach taken in our work is not new. The concept of analyses of the general circulation via study of "centers of action" had its champion in T. Bergeron. He referred to such study as dynamic climatology (Bergeron, 1930).

> . . . a dynamic climatology should describe the frequencies and intensities of well-defined systems

that are more or less closed in a thermodynamic sense.

Bergeron's concept of dynamic climatology differed from that of Hesselberg whose concept is close to the definition now generally accepted.

> Dynamic climatology must be concerned with the quantitative application of the laws of hydrodynamics and thermodynamics . . . to investigate the general circulation and state of the atmosphere, as well as the average state and motion for shorter time intervals

The outcome of the Hesselberg approach is best observed in the computer general circulation models (GCMs). Although GCMs look promising in identifying probable future states of the atmosphere associated with altered boundary conditions, they seem less likely to provide useful prediction capabilities for the monthly, seasonal, and year-to-year levels of the forecast problem. With the aid of modern computers and statistical techniques, the systematic spatial and temporal variations in the centers of action of the general circulation can be identified. The present work is offered as evidence of the value of this approach. Given Bergeron's concept of dynamic climatology and C. S. Durst's definition that *climate is the synthesis of the weather,* we conclude that the fundamental elements of climate are the various extant features of the general circulation rather than the more commonly assumed fundamental elements of weather (temperature, pressure, humidity, etc.). The task of climate prediction is then to specify future states of the general circulation and its centers of action in a stochastic sense. Given useful prediction, statements about associated fields of the fundamental elements of the weather may be possible on climatological time scales. Forecast trials employing this concept have proved successful and will be discussed in subsequent reports.

MODEL DEVELOPMENT

Three versions of the UVa Climate Forecast Model have been constructed. The original model (Hayden and Smith, 1982) used principal components analysis (PCA) to decompose the records of seasonal patterns of cyclone frequencies into orthogonal representations of the original data. The temporal persistences of these orthogonal representations (principal components) are used in making the forecasts. In the two later versions of the model, the constraint of orthogonality was 1) eased and 2) removed. In the former case (the second version of the model) the property of orthogonality was retained but the axes (principal components) were rigidly rotated with the constraint that variance explained by each of the selected lower order components be maximized. This is known as the VARIMAX rigid rotation. In the third version of the model the constraint of orthogonality is removed from lower order principal components and each axis is rotated such that each explains the greatest portion of residual variance unexplained by the sum of all of the lower order rotated components. This variation is called the PROMAX oblique rotation. In this report the unrotated principal components version is referred to as MODEL I; the VARIMAX rigid version is referred to as MODEL II; and the PROMAX oblique version is referred to as MODEL III.

For details on the properties and relative merits of various types of rotations of principal components the reader is directed to Richman (1983a, 1983b). Richman (1981) has also shown that rotated principal components give more faithful representations of meteorological data fields. Our studies show modest but consistent 2-by-2 forecast skill improvements with relaxation of the orthogonality constraint and the capacity to forecast some geographic locations with Model II and Model III that were not possible with Model I.

MODEL DATA

Monthly cyclone frequencies for the years 1885-1984 were tabulated from monthly charts of the "Tracks of the Centers of Cyclones at Sea Level" published by *Monthly Weather Review* and in recent years by *The Mariners Weather Log*. Multiple entries of a given storm in a grid cell were ignored. Grid cells south of 27.5°N were not included in this study because early forecast trials showed no forecast skill in this region. The 87 grid cells forecasted are indicated by the black dots in Fig. 1. Data spatial inhomogeneities due to the variable density observation network used to make the original storm track charts were ruled out as a problem in earlier work (Hayden, 1981b). Frequencies were not adjusted for latitude variations in grid-cell area because of distortions involved in such adjustments (Hayden, 1981c). For the purpose of constructing and testing the prediction model, the data matrix was divided into a dependent (1885-1960) part from which the principal components were calculated and the forecast models constructed, and an independent (1960-1980) part which was reserved and used to evaluate forecast skill. The post-1980 years were forecast in real time. Real time forecasts were generally completed two to three weeks following the close of a month. This time was needed to acquire the charts of cyclone tracks from NOAA, extraction of data from the charts, and running of the models. Alternative lead time could be planned and evaluated for changes in forecast skill. Lag correlation studies indicate that sufficient variance is explained out to a lag of one year and that useful forecasts with longer lead times merit study. Tests of shorter lags, i.e., one month lag indicate little or no forecast skill at that time scale.

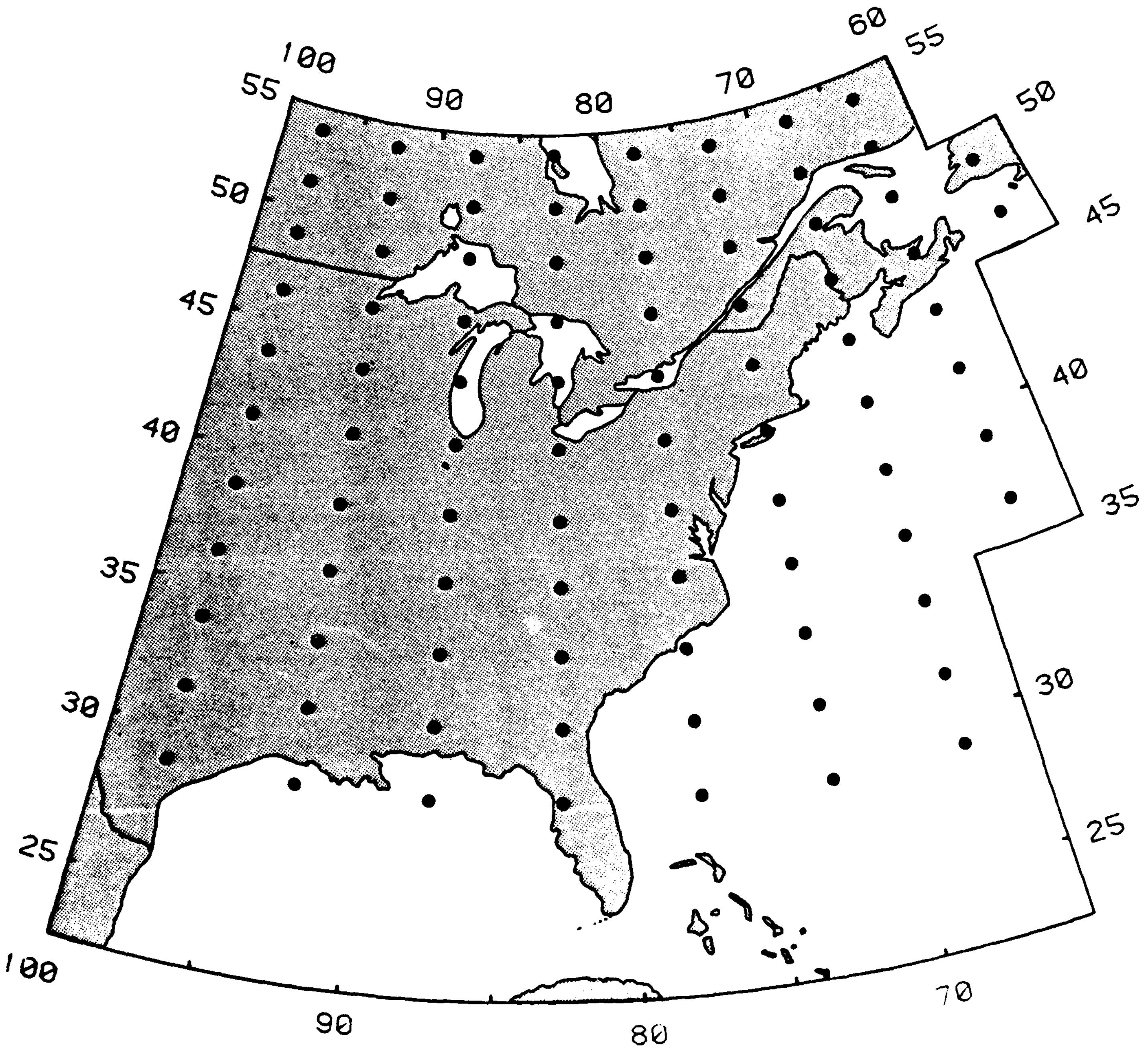


Fig. 1. Chart of the study area. 2.5° latitude by 5.0° longitude grid cell centers are indicated. There are 101 rectangular grid cells in tbe study area. Only those grid cells north of 27.5 N are used in this study.

MODEL CONSTRUCTION

Figure 2 shows, schematically, model construction. The first stage in the construction of the models was data preparation. The archives of cyclone frequencies were first divided into two parts. All the data from 1885-1959 were reserved for model construction (the dependent data). The data for the years 1960-1980 were reserved for forecast trials (independent data -- hindcasts). Data for the post-1980 period were used in real time to make forecasts (independent data -- operational forecasts).

Monthly cyclone frequency data are composited into six-month seasons. Twelve six-month seasons are defined. The principal components of cyclone frequencies for each of the 12 seasons are then calculated. The first five of these components for each season are then subjected to VARIMAX and PROMAX rotations. The case weightings for each vector for each season for each year of the dependent data record are calculated and reserved.

The vector case weightings are used to derive the one-season lag regression equations. These regression equations are used to estimate the case weightings for one season from the known case weighting for the previous season. The regression equations in Model I differ from those of Model II and Model III. In Model I,

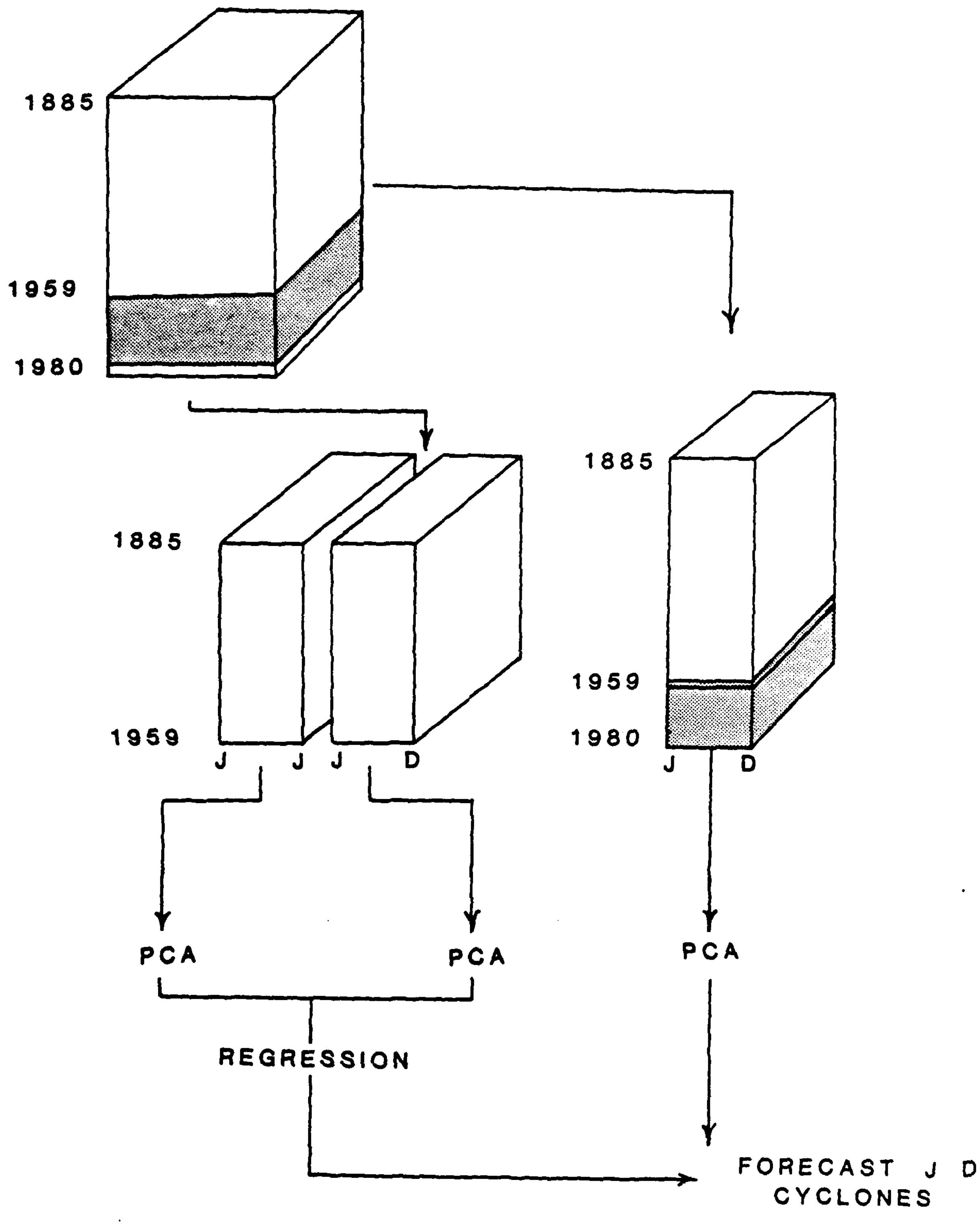


Fig. 2. Schematic of assembly of the cyclone frequency prediction models. Clear portions of the cubes represent dependent data (shaded = independent data). PCA refers to principal components analysis. J-J = January-June; J-D = July-December.

the case weightings for the two seasons are regressed for each component but no cross component regressions are used because the orthogonality of the components and their season-to-season similarity always resulted in near zero correlations between seasons. In Models II and III within- and cross-correlations are examined and the regression with the highest correlation is selected for use. In all cases (Models I, II and III) if there are correlations below 0.35 the term is not used in the equation. Previous trials showed that rarely was there a model forecast skill when the correlations were below 0.35. This constituted a pre-screening and thus a reduction in the number of models that required development and testing.

Using the regression equations, the case weightings for each vector for each model version (Models I, II and III) are estimated and used in the forecast equations. The general form of these equations is given in Hayden and Smith (1982) as

$$C_s = X + a_1 O E_1 + a_2 O E_2 + \ldots + a_5 O E_5 \quad \{1\}$$

where C_s is the matrix of predicted cyclone frequencies for each grid cell for the season to be forecasted; X is the matrix of long-term (1885-1959) mean cyclone frequencies at each grid cell for the season to be forecasted; O is the matrix of standard deviations of X at each grid cell for the season to be forecasted; E_i are the principal components for the season to be forecasted (non-rotated or rotated depending on model version being constructed); and a_i is the forecasted case weighting calculated from the one-season lagged regression equations.

Each term in the equation may be considered an individual model. As five components are used in construction of these models each term may be evaluated for forecast skill. The additive combinations of terms can also be evaluated. A large number of possible model configurations is thus possible. Only the models with all terms included are reported on here. Model I has four terms and Models II and III have five terms.

THE MODELS

Earlier we (Hayden and Smith, 1982) published the details of the models to predict cyclone frequencies for the October-March and April-September six-month seasons. The component parts of each of the 12 six-month season models constructed for all three versions of the model (I, II and III) are on file at the University of Virginia. Each model consists of the data matrixes listed in Table I.

TABLE I

Summary of the Forecast Model Matrixes*

MATRIX	N	DEFINITION OF THE MATRIX
X	87	Cyclone frequency means
O	87	Standard deviations of X
E_j	87	Predictor eigenvector variable loadings
E_{j+1}	87	Predictand eigenvector variable loadings
F_j	75	Predictor case weightings
F_{j+1}	75	Predictand case weightings
R	5	F_j vs F_{j+1} regression coefficients

* = 12 six-month season models for versions I, II and III
N = number of elements in the matrix

Examples of the matrixes for the October-March season in Table I follow. Figure 3 shows matrix X for 1885-1960 long-term means. Figure 4 shows the matrix of the standard deviations (O) of X. Figure 5 shows the matrix E and Figure 6 shows the column matrix (F) by year. The archives of the forecast models and forecast products are voluminous and do not lend themselves to reproduction in technical reports. They are available for inspection and study at the University of Virginia.

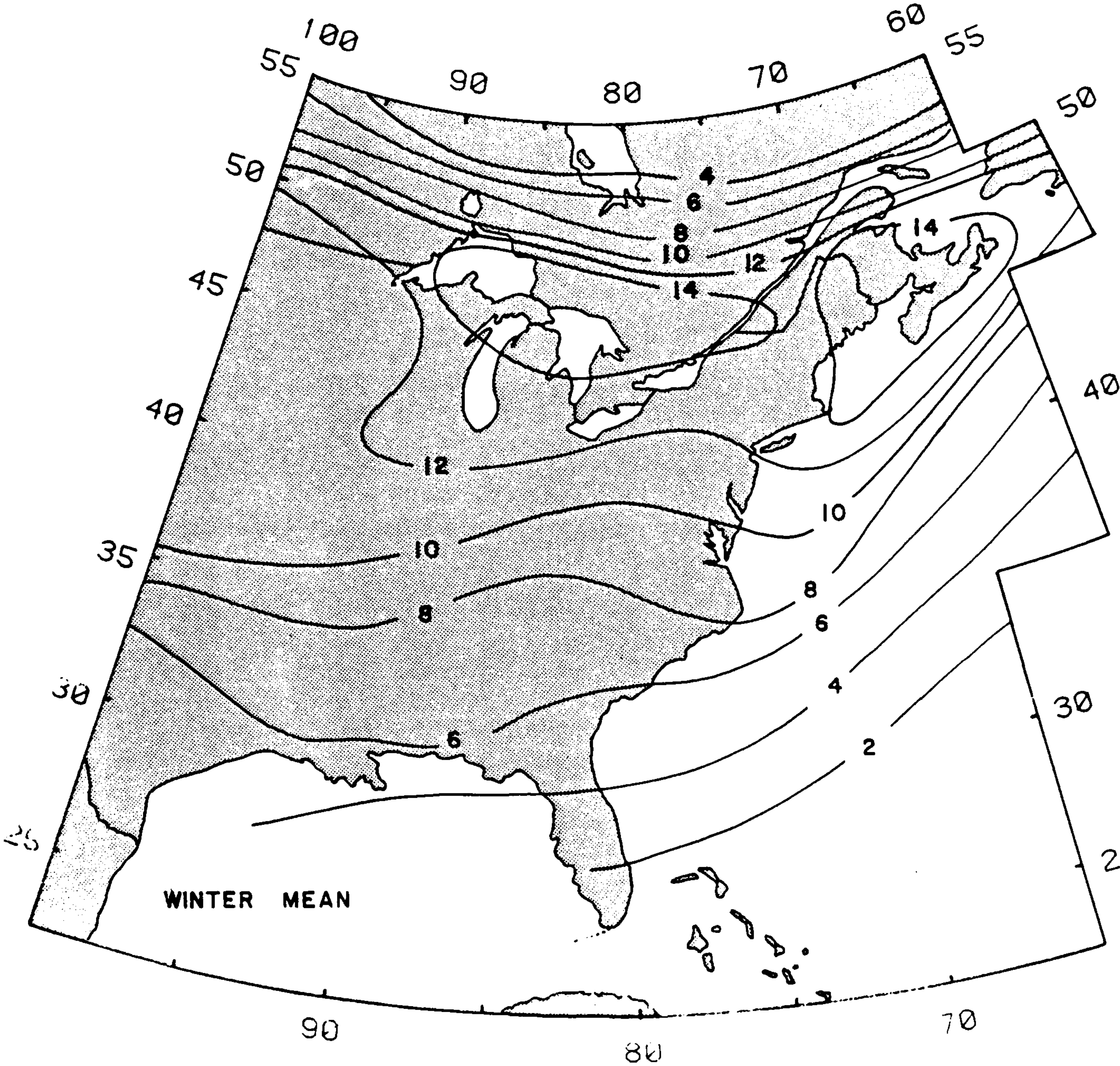


Fig. 3. The matrix X of October-March long-term mean cyclone frequencies displayed in map form. The units are cyclones per grid cell.

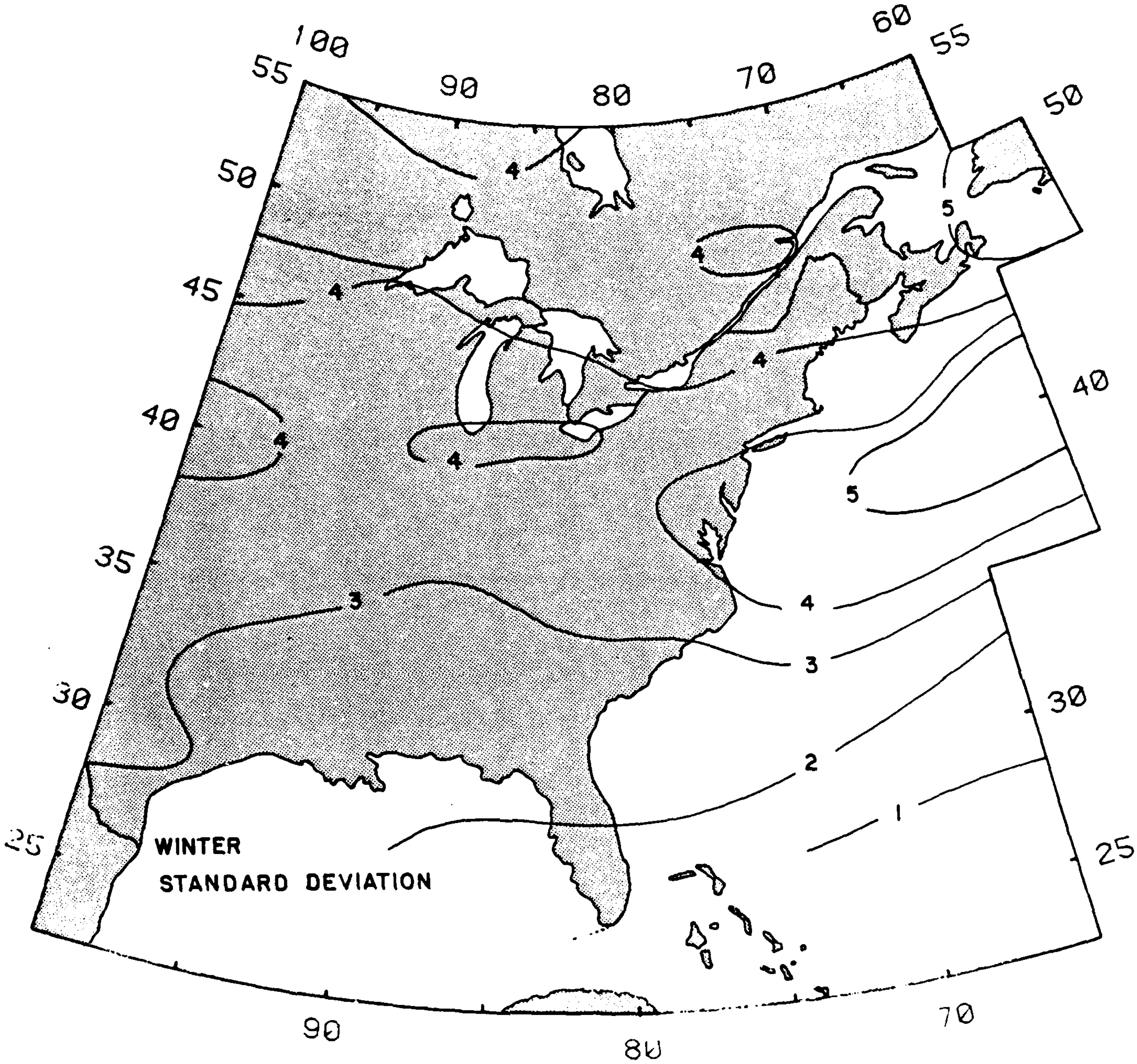


Fig. 4. The matrix (O) of the standard deviations of cyclone frequencies for the October-March season. The units are cyclones per grid cell.

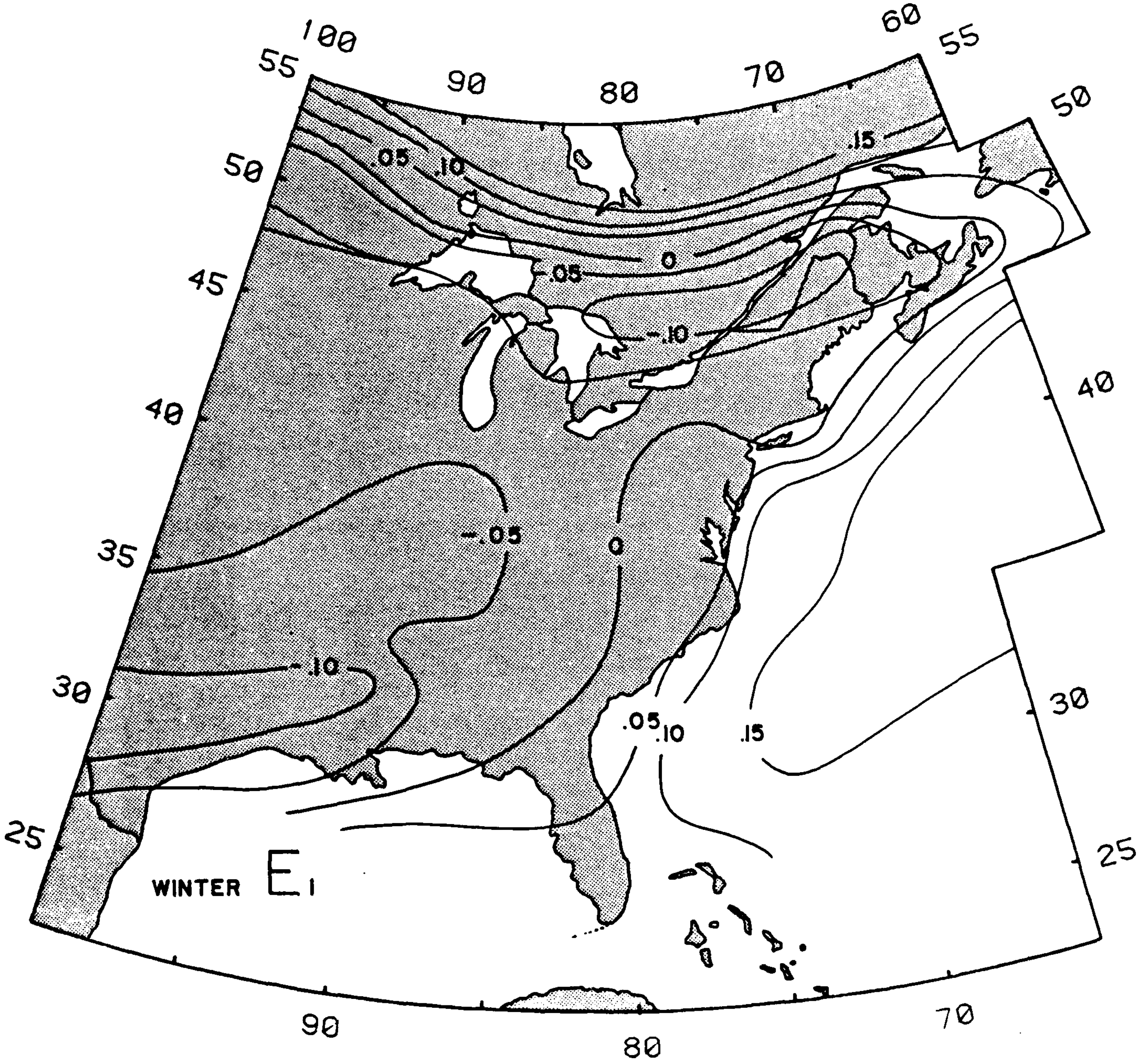


Fig. 5. The matrix E for the first principal component of the winter (October-March) season. The values plotted are dimensionless.

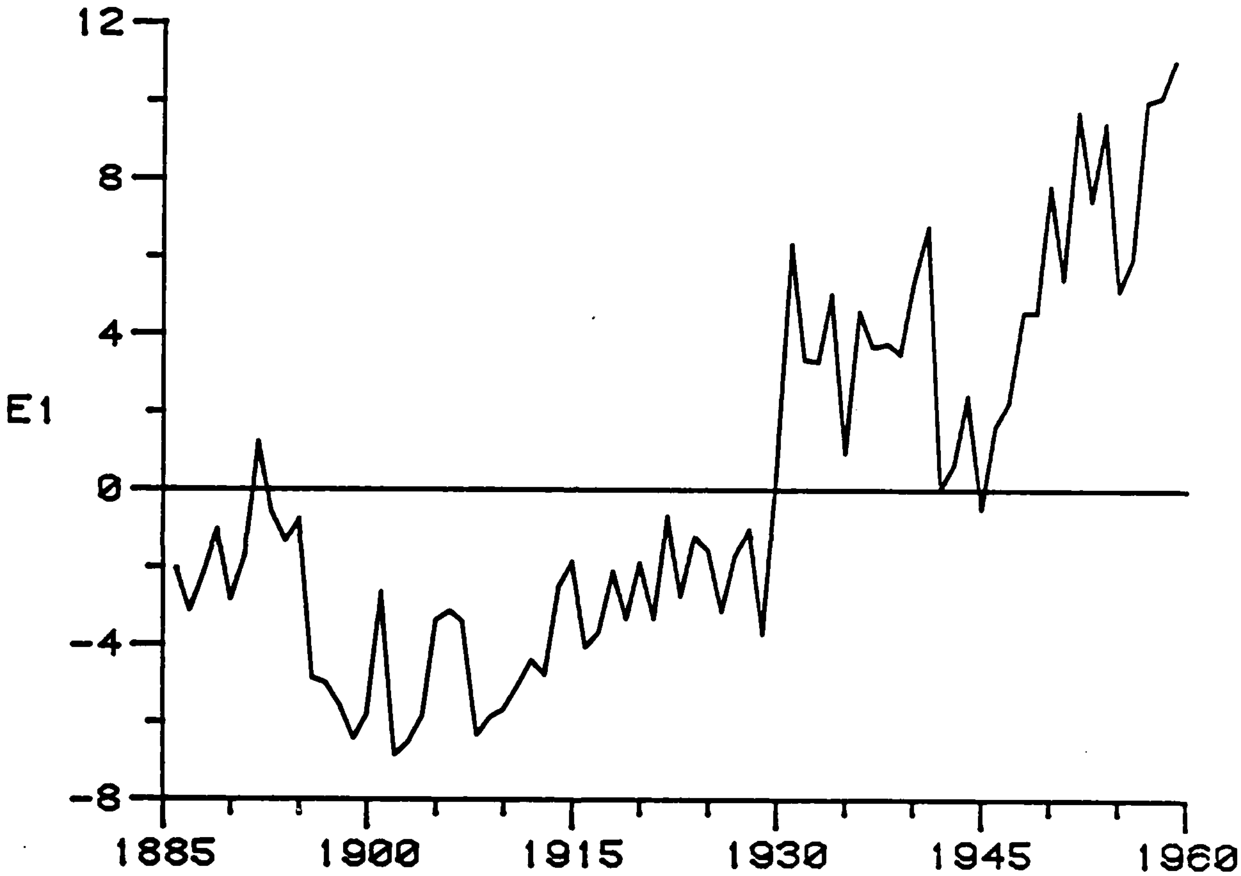


Fig. 6. The column matrix F of case weightings of the first principal component of cyclone frequencies for the October-March season. The values plotted in the time series are dimensionless.

THE FORECASTS

The Hindcast Period: 1960-1980

In order to generate a forecast for a six-month interval, the case weightings on the principal components of the previous six-month period must be calculated. This requires that principal components used in the forecast irclude data for the previous six-month period. In the case of the first forecast of the independent data period (January through June 1960), the calculated principal component case weightings for the July through December 1959 period were entered into the dependent data period regression equations, and the predicted case weightings for the January to June (1960) period were derived. For this first forecast the dependent data period contained all the months needed to predict the first six-month season of the independent data period. In subsequent forecasts new principal components analyses had to be run to generate the case weightings needed as input into the regression equations. At no time were data for the independent data period included in the regression equation development. All forecasts were made for time beyond that used to build the models.

The Operational Period: 1980-1983

Charts of the tracks of the centers of cyclones for each month are prepared by NOAA at the end of each month. They are released and are publicly available about 15 days after the close of the month. On receipt of the charts, frequencies per grid cell are counted and entered into the data base. Principal components are then found for the six months just concluded and case weightings for each component calculated. The regression equations derived for the dependent data period (1885-1960) are used to estimate case weightings for the upcoming six-month season. The forecasted weightings are then used in Equation 1 to estimate cyclone frequencies in coming seasons. Operational forecasts were begun in 1980.

Forecast Products

Two forecasts are presented here to illustrate the nature of the forecast products generated. Both were made on an operational basis. The forecast for October to March 1980-1981 was selected because it was extreme in the sense of having largely negative departures from the mean forecast almost everywhere and the magnitude of the negative anomaly forecasted was large. The second forecast selected for illustration was for the September to February 1981-1982 period. This forecast contains both large positive and large negative anomalies from the mean. Three products are returned from the forecast. First, the long-term mean cyclone frequencies are presented in map form. Second, the predicted anomalies in cyclone frequencies for each grid cell are displayed in map form (Fig. 7 and 8). The third product is a map of the predicted anomalies added to the means (Fig. 9 and 10).

The range of forecasted anomalies generally averages from six to ten cyclones per grid cell. As typical maximum values of the means for a six-month season are on the order of 12 cyclones per grid cell, the forecasted anomalies are large in relative magnitude. The contoured anomaly fields (Fig. 7 and 8) are interesting in that one type of axis of maximum values and two types of axes of minimum values are evident. The axis of maximum values along the east coast of the U.S. (Fig. 8) can be directly interpreted as an axis along which more than the normal number of cyclones is likely to be observed if the forecast is correct. The axis of absolute minimum values "negative storm track" e.g., as in the track extending eastward from Colorado (Fig. 7 and 8), is interpreted as an axis along which fewer than the normal number of storms are expected. Finally, within an area of forecasted negative anomaly, there may be axes of local "maxima" or small negative values, e.g., the trace of small negative values across the Great Lakes in Figure 7. Thus while storms might be less frequent than normal, those that do occur would tend to move along this track. The three different types of tracks are illustrated with different symbols in the illustrations.

Clearly the charts of forecasted anomalies do not provide all the information that is needed to interpret the forecast, so we added the forecasted anomaly to the long-term mean (Fig. 9 and 10). The resulting chart has positive values everywhere and so the interpretation difficulties of "negative tracks" are no longer present. The resulting axes of maximum values can directly be interpreted as the forecasted preferential location of the storm tracks for the forecasted season. While forecast skill will be discussed in a subsequent section it should be noted that both these forecasts were successful. The sign of the anomaly was forecasted correctly in 74.2% of the 87 grid cells in the October to March 1980-1981 forecast and 89.7% of the 87 grid cells were correctly forecast in the September to February 1981-1982 forecast.

To show each of these products for each model version and for each forecast made would require the display of thousands of maps. This is beyond the scope of this technical report. All of the maps are on file at the University of Virginia in the author's archives. The subsequent observations and verifications of each forecast are also saved for study.

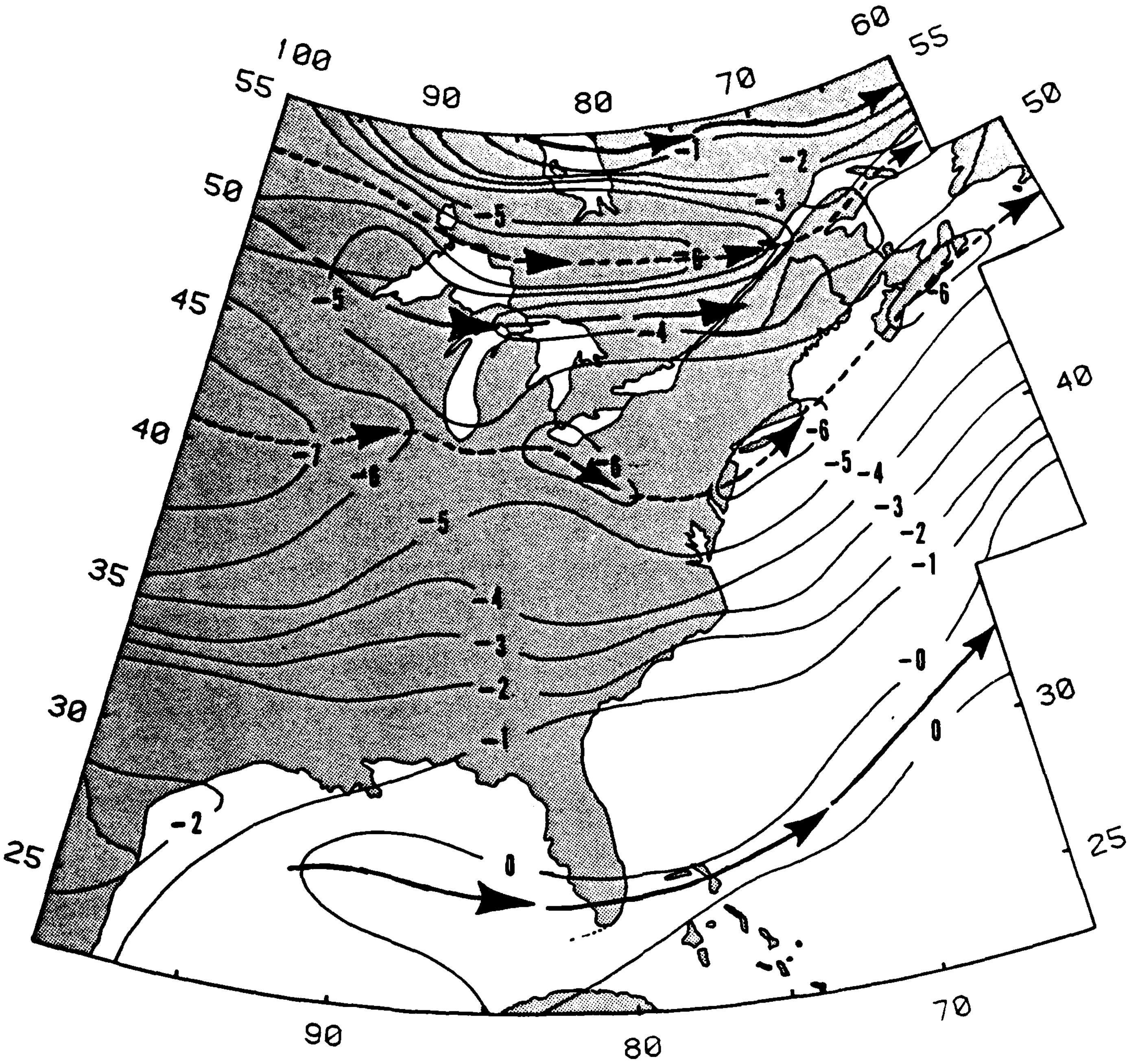


Fig. 7. Operationally forecasted cyclone frequency anomalies for October-March 1980-1981 (Model I). Forecast was issued on 24 October 1980. The units are cyclones per grid cell. Solid arrows indicate axes of maximum positive anomaly; short dashed arrows indicate axes of maximum negative anomaly; long dashed arrows indicate minimum negative anomaly axes.

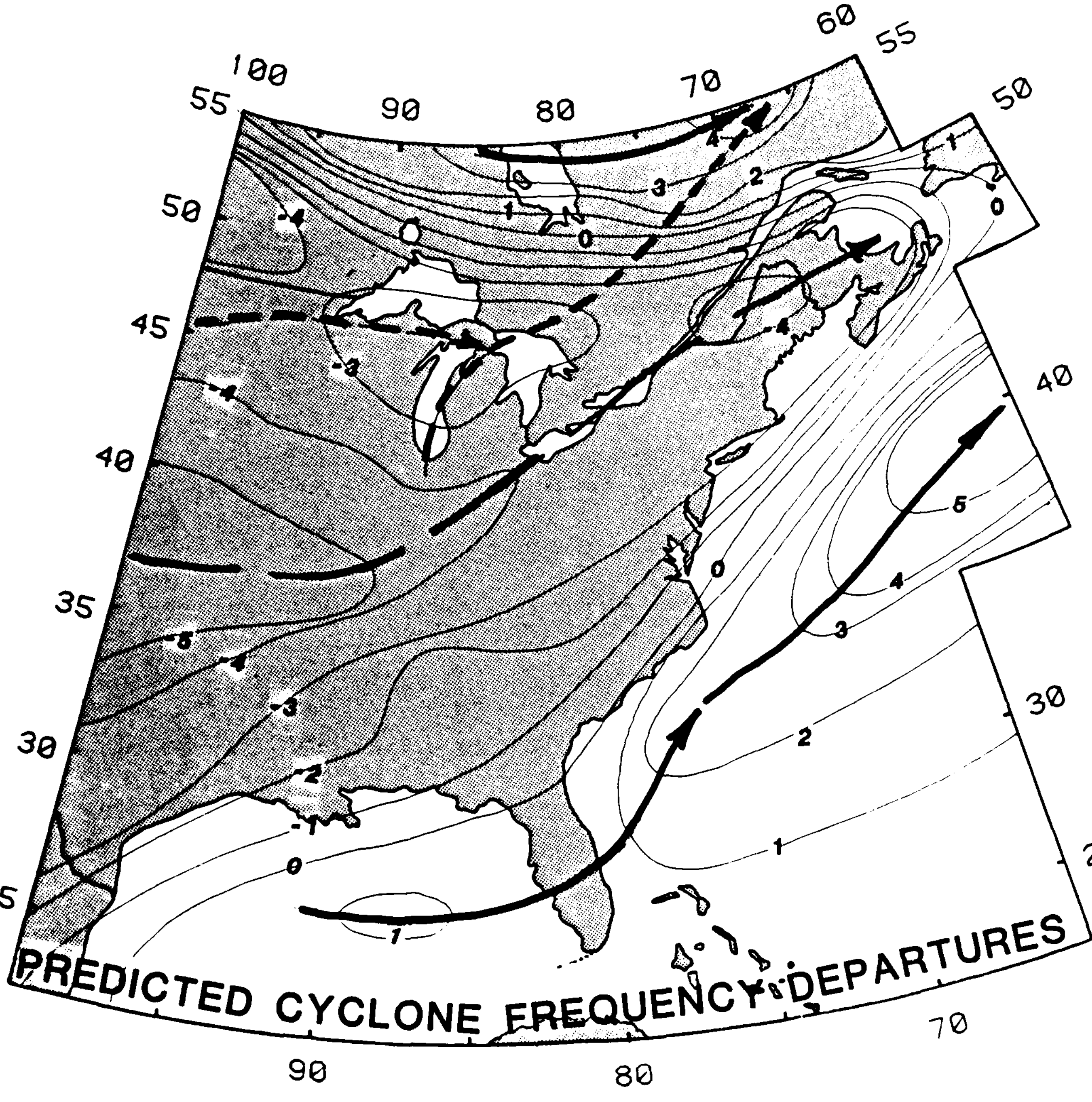


Fig. 8. Operationally forecasted cyclone frequency departures from the long-term mean for the September-February 1981-1982 season (Model I). Solid arrows indicate axes of maximum positive anomaly; long dashed arrows indicate axes of maximum negative anomaly; short dashed arrows indicate axes of minimum negative anomaly.

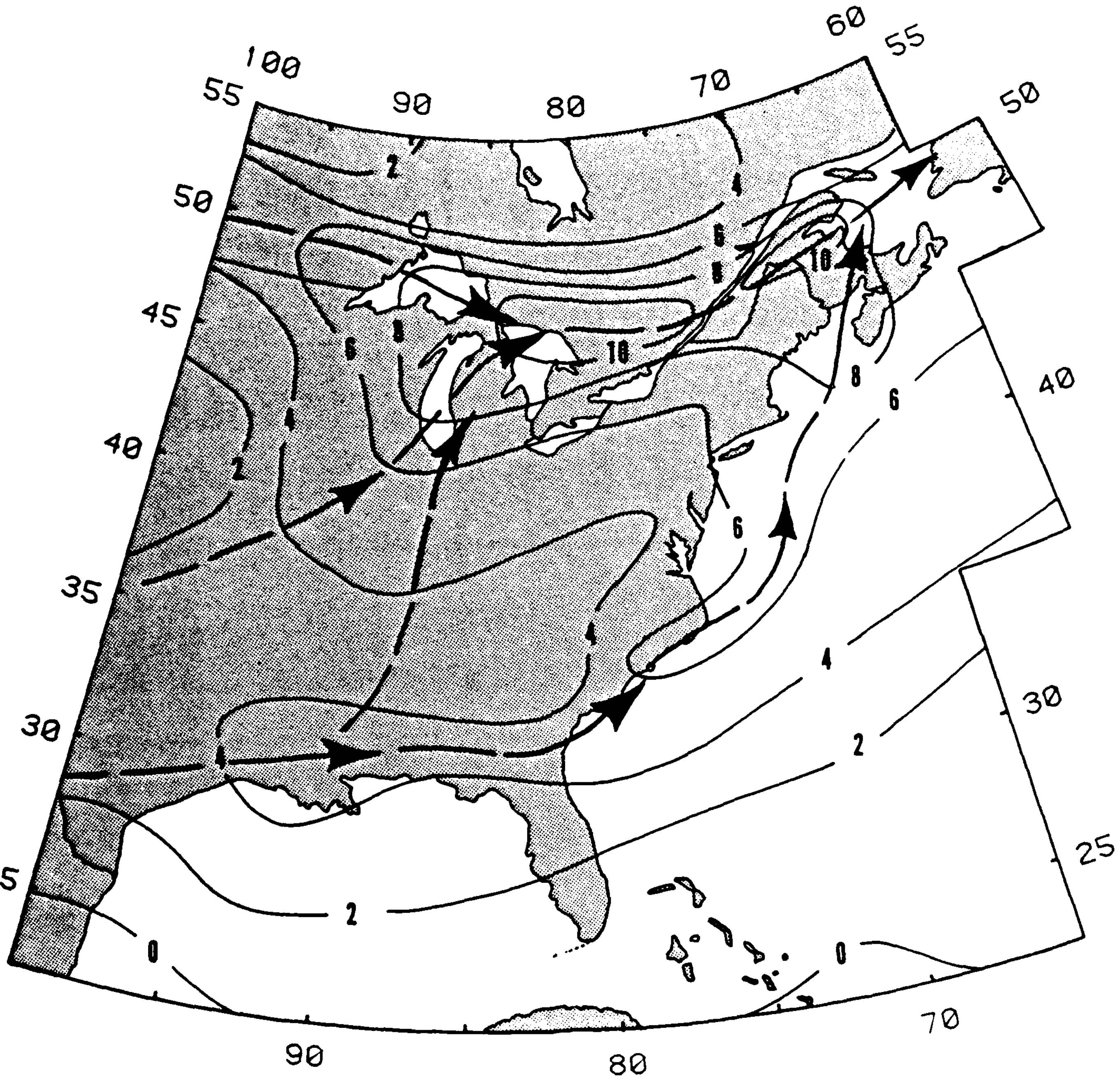


Fig. 9. Operationally forecasted cyclone numbers for the October-March 1980-1981 season (Model I). The units are cyclones per grid cell. Arrows indicate axes of local maximum frequencies.

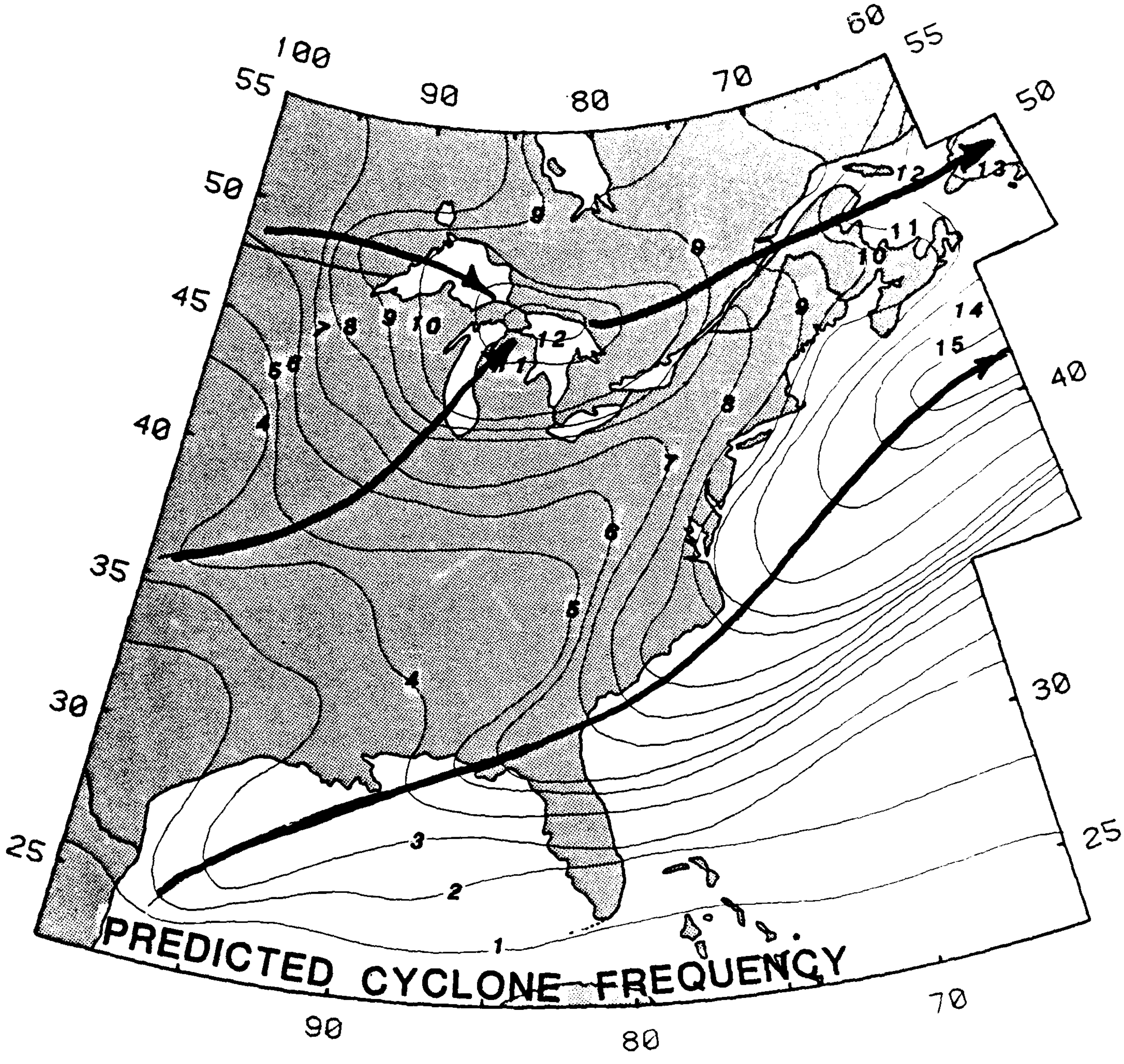


Fig. 10. Operationally forecasted cyclone numbers for the September-February 1981-1982 season (Model I). The units are cyclones per grid cell. Arrows indicate the axes of local maximum frequencies.

Inventory of Forecasts

Table II lists the number of forecasts made for the independent data period and the operational period for MODELS I, II, and III. We made 21,924 forecasts for each model version for the independent data period, and 2,958 forecasts were made using each model version during the period of operational forecasting. Comparisons of these forecasts with observations form the basis for assessing the forecast skill of the models constructed.

TABLE II

Inventory of Forecasts

	Forecast Period					
	(1960-1980)			(1980-1982)		
MODEL	I	II	III	I	II	III
Grid cells (A)	87	87	87	87	87	87
Seasons (B)	12	12	12	12	12	12
No of years (C)	21	21	21	3*	3*	3*
AxBxC (total forecasts) for Models I, II and III	(21,924)			(2,958)		

* 1983 June-Nov and July-Dec forecasts were not verified in time for this report.

MEASURES OF FORECAST SKILL

Numerous methods have been advanced to quantify estimates of forecast skill (Brier and Allen, 1951; Vernon, 1953), and as noted by Brier and Allen the method selected depends on the purpose of verification. The purpose here is to establish the level of reliability of the forecast scheme relative to the climatological means as forecasts. A battery of tests of forecast skill is reported here. Two types of forecasts are made and evaluated: category and magnitude forecasts. In most trials on climate forecasts magnitude forecast skills are not reported. Rather, various categorical measures are reported (e.g., 2, 3, and 4 category tests). Magnitude measure obviates the need for complex categorical measures.

Percent Correct Score

The percent correct score is the simplest measure of forecast skill. This measure is used to assess the skill of forecasts where only two types of forecast are used, i.e., above or below the mean. This is sometimes referred to as the 2-by-2 or sign test. Chance alone would dictate a percent score of 50%. In the present study 21 years are forecast in the independent data forecast trials (1960-1980). As these forecasts were made after 1980 the term hindcasts is applied. Table III gives the probabilities that various 2-by-2 percent correct scores could occur by chance alone.

TABLE III

CHANCE PROBABILITIES IN 2-BY-2 TRIALS

NO. CORRECT FORECASTS IN 21 TRIALS (%)	PROBABILITY OF EXCEEDING BY CHANCE ALONE
21 (100%)	.0000004
20 (95%)	.000011
19 (90%)	.00012
18 (86%)	.00075
17 (81%)	.0036
16 (76%)	.014
15 (71%)	.040
14 (67%)	.095
13 (62%)	.20
12 (57%)	.34

For each forecast period cyclone frequencies are estimated for 87 locations (grid cells). These 87 cannot be considered mutually independent trials of the model. The most conservative standard for acceptance or rejection of a trial is the .05 probability level at a given location or grid cell. This test would be considered "over conservative" by Livezey and Chen (1983). Earlier tests of the model (Hayden and Smith, 1982) indicated that magnitude forecast skill was present in a model if the

2-by-2 percent correct score based on 21 trials equaled or exceeded 67%. The reader should view subsequent statements on model skill in light of these standards. A 71% skill score standard at each grid cell (a local skill score) is very conservative. A 71% average skill score for the entire 87 grid cell field (a global skill score) is even more conservative. Nonetheless these standards are exceeded by the present model.

Heidke Skill Score [H]

The Heidke skill score is also a measure of skill in a 2-by-2 or sign test. The Heidke skill score is calculated as follows:

$$H = (R-E)/(T-E) \quad \{2\}$$

where R is the number of correct forecasts, T the total number of forecasts, and E is the expected number correct by some standard such as chance. The Heidke skill score resembles the percent correct score but is scaled over a range of 0 (no skill) to 1.0 (perfect skill). Many investigators prefer the Heidke skill score over the percent correct score, but the percent score is more widely understood. Arithmetic interconversion between the two measures is $H=(\%-50)\times 2$ where % is the percent correct skill score. Both skill scores are reported here to facilitate model evaluation.

Deviation Skill Score [D]

The deviation skill score is calculated as follows:

$$D = (d_e - d_f)/(d_e \quad (3)$$

where df is sum of the deviations between forecasted and observed values and de is sum of the deviations expected by the mean as the forecast. The deviation skill score (Vernon, 1953) is used in non-category forecasts where the magnitude of the anomaly is forecasted. In the deviation skill score the deviations of the forecast from the observed occurrences are weighted linearly. The larger the error the larger the penalty. Small forecast errors are rewarded over larger ones.

Quadratic Skill Score [Q]

The quadratic skill score is calculated as follows:

$$Q = (d_e - d_f)/ d_e \quad (4)$$

where the terms are as described above. In the quadratic skill score (Vernon, 1953) the penalty to the forecaster varies with the square of deviation of the forecast from the observations. Here the penalty for large errors is severe. Ideally one would like a high percent correct skill score and a high quadratic score.

AAE and RMSE

The average absolute error (AAE) is the average error irrespective of the sign of the forecasted anomaly relative to the mean. This value is compared to the average absolute error of the mean as a forecast. A direct error reduction relative to the mean as a forecast expressed as a percentage can then be calculated. In the case of the root mean square error (RMSE) the deviations of the forecasts from the observations are squared, summed, and divided by the number of forecasts; then the square root is taken. A reduction of the root mean square error of the mean as a forecast is desired for the model forecast. If the sign of the forecast is correctly made all the time then the minimum root mean square error can be insured with a forecast of the historical average absolute error of the mean as a forecast. The average absolute error of the forecast, if forecasts are normally distributed, can be used to divide the distribution into quarterlies for 4-by-4 skill tests.

Local Skill

The term local skill is reserved for geographic or point skill. It is the average skill at a point over time. In the present study, forecasts were made for 87 grid cells (Fig. 1). Local skills are reported for each grid cell. Under ideal circumstances local skill should pass a 0.05 test of statistical significance (Table II). The % of correct forecasts needed to pass the 0.05 level at an individual grid cell is dependent on the number of forecast trials. Twenty-one trials is the standard used in Table II.

Global Skill

When local skills are aggregated or spatially averaged, a single skill score "representing" all localities is reported. This score is referred to as a global skill score. Two types of global skill scores are defined here. As the forecast models are constructed for six-month duration seasons and 12 such seasons are defined, we then have within-model global scores and between-model global scores. Thus we have a global skill score for the six-month season beginning in April and ending in September and also a global skill score which averages all possible six-month season models.

Global skill scores are convenient in that a single number can be forwarded as a most general measure of model reliability. However, it should be remembered that forecast skill varies from season to season and from place to place. These variations must be understood if the models are to be properly evaluated and, more importantly, used. Because skill at one site may not be independent of skill at adjacent locations, great care must be exercised in specifying statistical significance for global measures of skill. Global skills reported in the absence of reported local scores may be misleading. A very conservative standard and one recommended here is that the average global % skill score is as large as required to pass a local test of skill (see Table II).

ASSESSMENT OF FORECASTS

The Mean as the Forecast

Forecasts are usually expressed relative to the mean as the alternate and simplest forecast. Where the distribution is normal the mean tends to be the most frequent occurrence. While mean might well be a prudent and conservative forecast, the mean is not always a good forecast. To examine the mean as a forecast we used the 1885-1960 cyclone frequency means for the various six-month seasons as forecasts for the six-month seasons between 1960 and 1982. Figures 11 and 12 illustrate the average absolute and root mean square errors of the means as forecasts. It is clear from both measures that the mean as a forecast varies with season and that there is a secular trend toward the mean as a progressively better forecast. Between 1960 and 1982 the root mean square errors have fallen from about 5 cyclones per grid cell to about 2.5 cyclones per grid cell.

The reasons for the decline in average absolute and root mean square errors of the means as forecasts are unclear. We conclude that variability has declined because the departures from the mean have fallen. Whittaker and Horn (1981) tabulated cyclogenesis over North America and found a general decline in cyclogenesis. The overlap between their data and ours is plotted in Fig. 12. Apparently the decline in cyclone frequency variability is associated with fewer cyclones developing and perhaps the "clipping" of extreme occurrences. Whittaker and Horn suggest that the decline over North America is compensated for elsewhere in the Northern Hemisphere but they are not able to detail the compensation. If the downward trend is real, then it would follow that the mean has become a more difficult standard to better. As will be seen in later sections, model

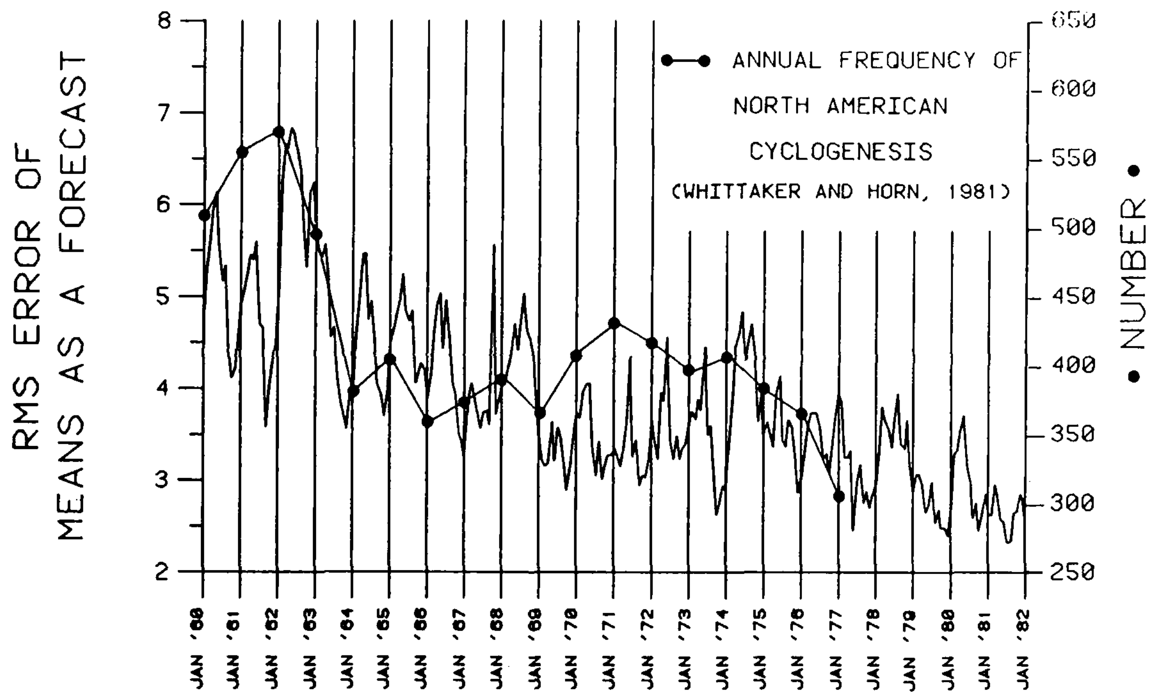


Fig. 11. RMSE for the (1885-1960) means as a forecast by season and year (1960-1982). The line with circles is the trend in the annual frequency of North American cyclogenesis (after Whittaker and Horn, 1981).

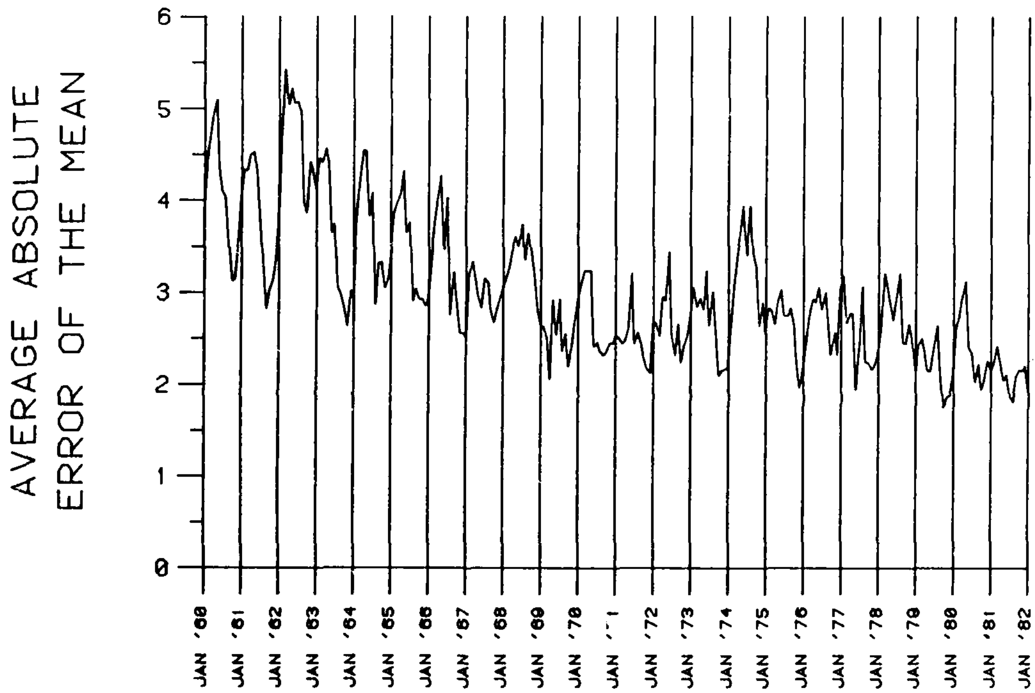


Fig. 12. AAE of the mean as a forecast by season and year (1960-1982).

forecast skill does not show a secular decline. Forecast skill of the models being tested remained high during the period of improvement of the mean as a forecast. We interpret this to indicate that model forecast skill is not sensitive to magnitude of the departure from the mean represented by the observed conditions.

Magnitude Versus the Sign of Forecasted Anomalies

The quadratic skill score measures how well the forecast model predicts the size of the departure from the observed conditions with penalty proportional to the square of the departure from the mean. The percent skill score measures how well the forecast model predicts whether the departure will be + (above the mean) or - (below the mean). Clearly, a model that does a good job of predicting the magnitude of the anomaly should also do a good job of predicting the sign of the anomaly. The reverse is not necessarily true. Accordingly, we have plotted the quadratic skill scores of Model I for all 12 six-month season forecasts for the period 1960-1983 against the percent correct skill scores for the same period (Fig. 13). When percent correct forecast skill falls below 60%, quadratic skill is negative. The relationship is strongly linear; however, care should be exercised when percent correct skill falls below 60% because skill in forecasting the magnitude of the anomaly cannot be demonstrated.

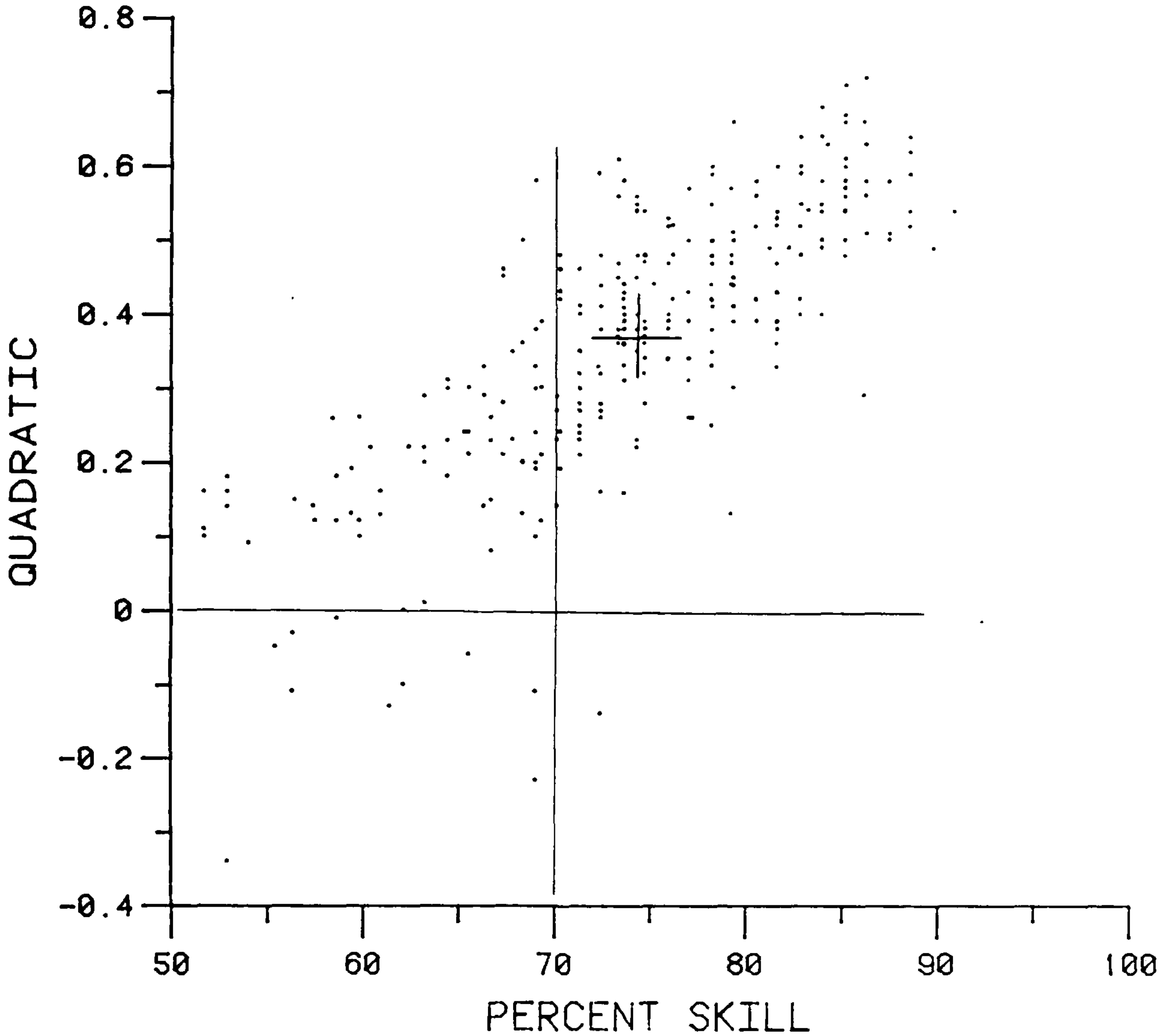


Fig. 13. The relationship between percent correct and quadratic skill scores (1960-1983) for Model I across all 12 six-month season forecasts. Cross indicates the means for the two measures of skill; horizontal line is the zero quadratic skill level; vertical line is the 0.05 significance level for a local test of skill.

limit of percent correct skill that is associated with quadratic magnitude skill (see Fig. 13). Areas with skill less than 60% are not contoured. The grid cells indicated with a black circle are those grid cells where 21 correct forecasts were made in trials. This 100% correct score occurs in regions of generally high forecast skill and they are not outliers due to chance.

Four areas of excellent skill in all seasons are found: 1) off the east coast of the U.S.; 2) in areas north of 50^{o} N latitude; 3) across the northern plains; and 4) an area extending northeastward from the southern plains. These four areas represent four important storm tracks that are not evident in the charts of the means of cyclone frequencies (Hayden, 1981a and Hayden and Smith, 1982). The central region of the eastern U.S. is generally forecast with a skill of at least 70% but small regions of lower skill occur in some seasons.

If we use actual local skill scores as a proxy for the attribute of predictability (see Madden and Shea, 1978) then the geography of skill presented here is at odds with that reported by others. Madden finds that predictability is highest in coastal areas and declines toward the interior of the country. This is not the case for cyclone frequency prediction. Predictability does not decrease toward the interior of the continent or in the offshore direction and skill along the coast is generally lower than in adjacent areas.

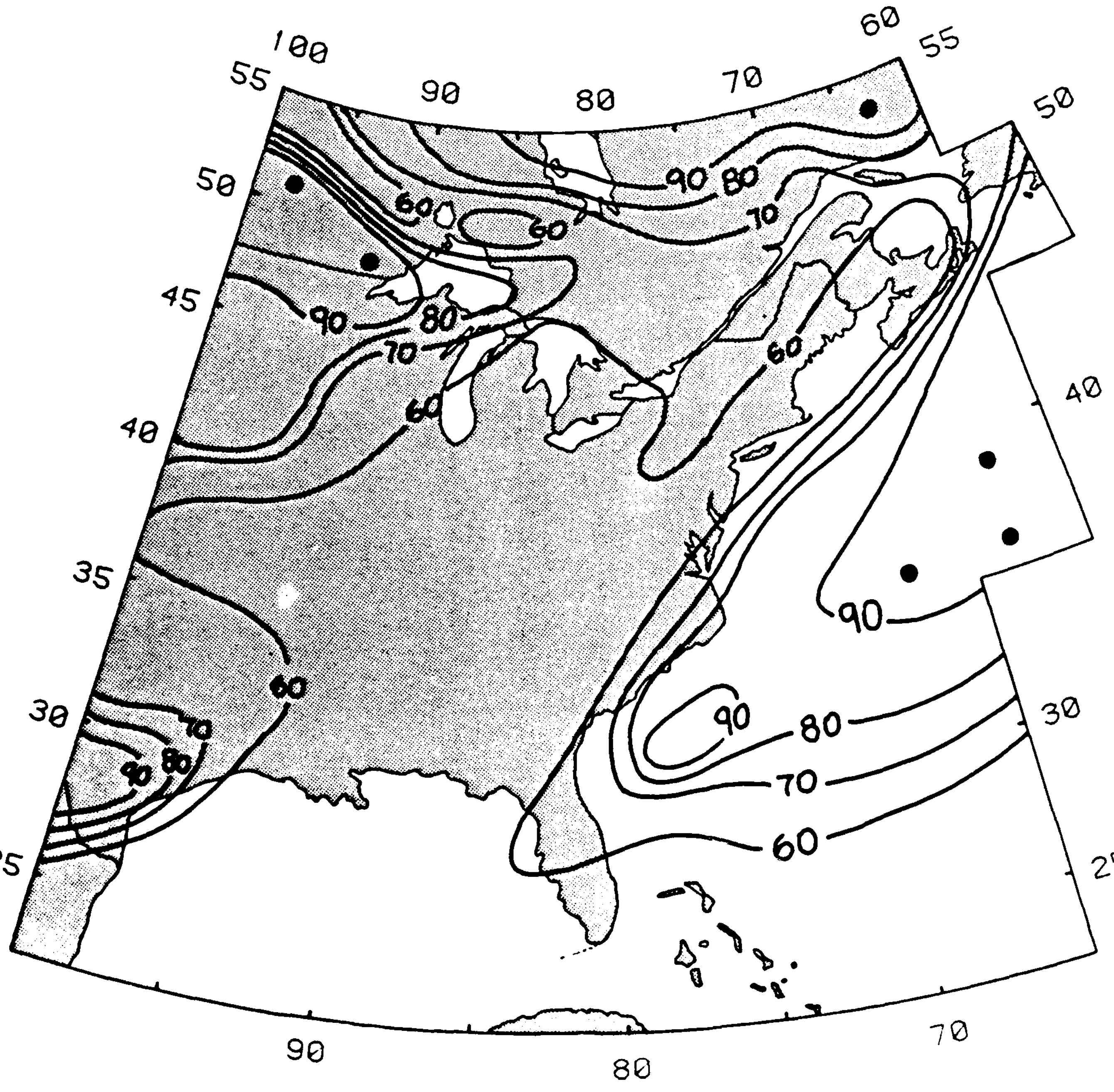


Fig. 14. January-June local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

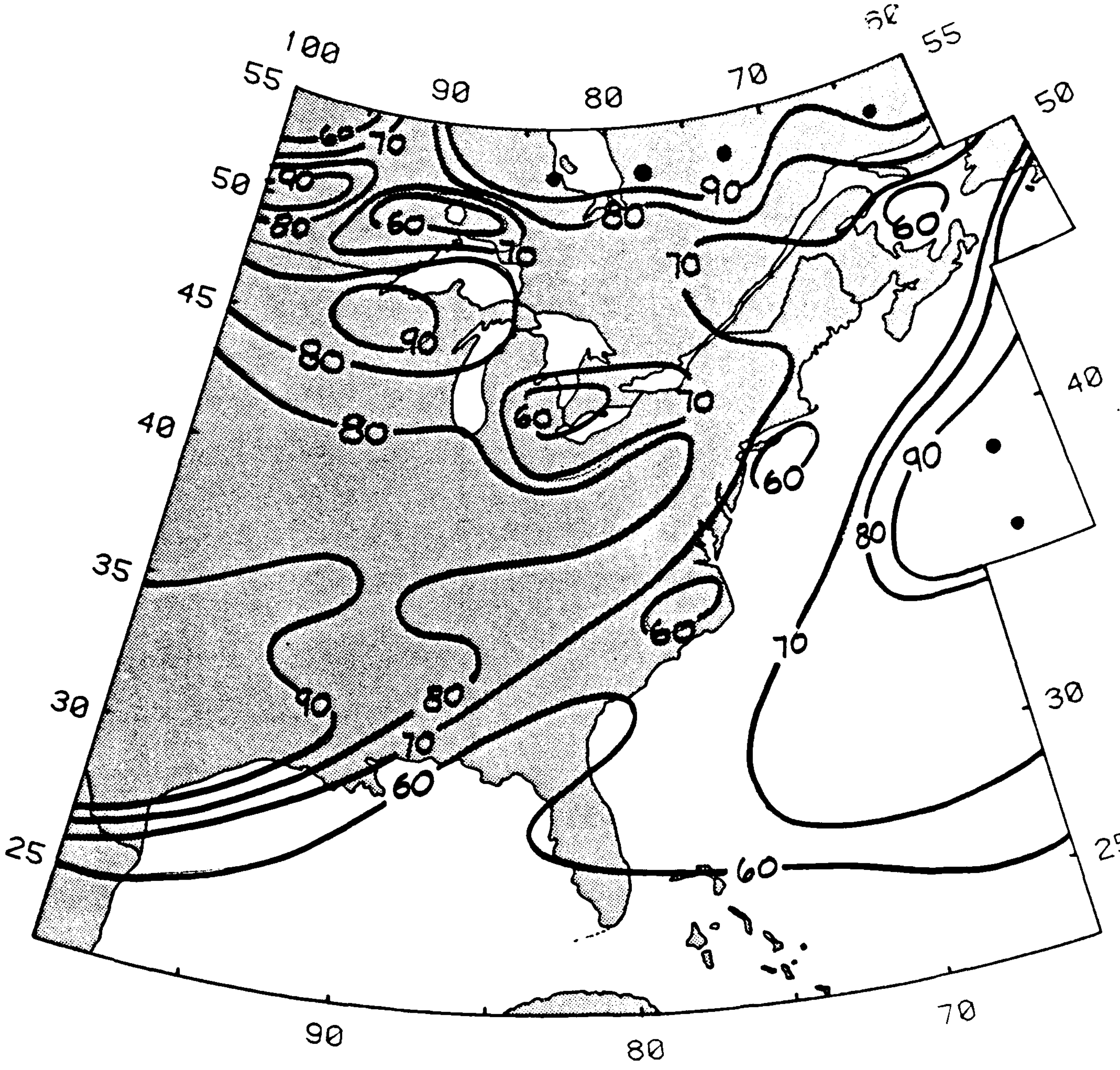


Fig. 15. February-July local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

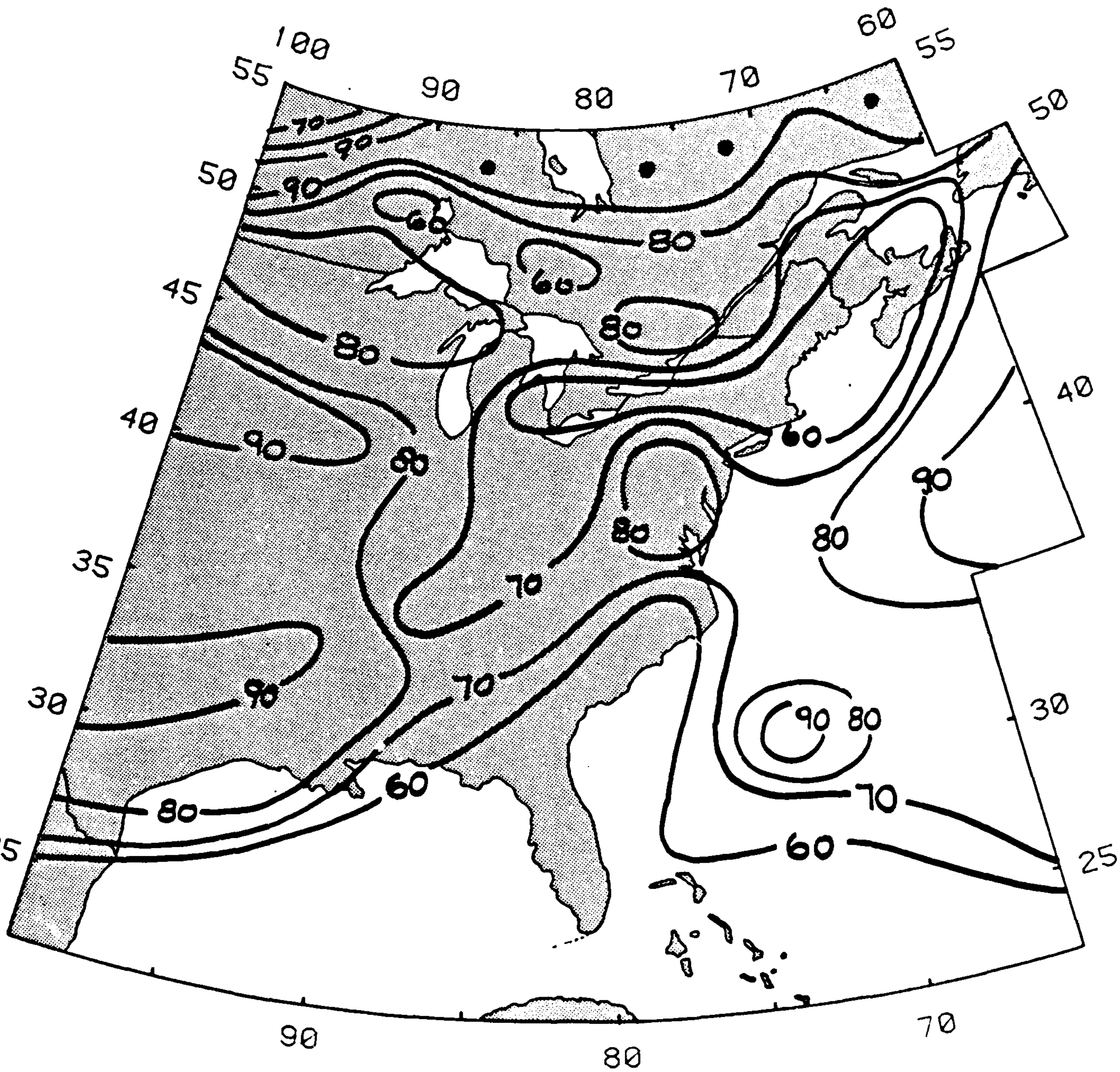


Fig. 16. March-August local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

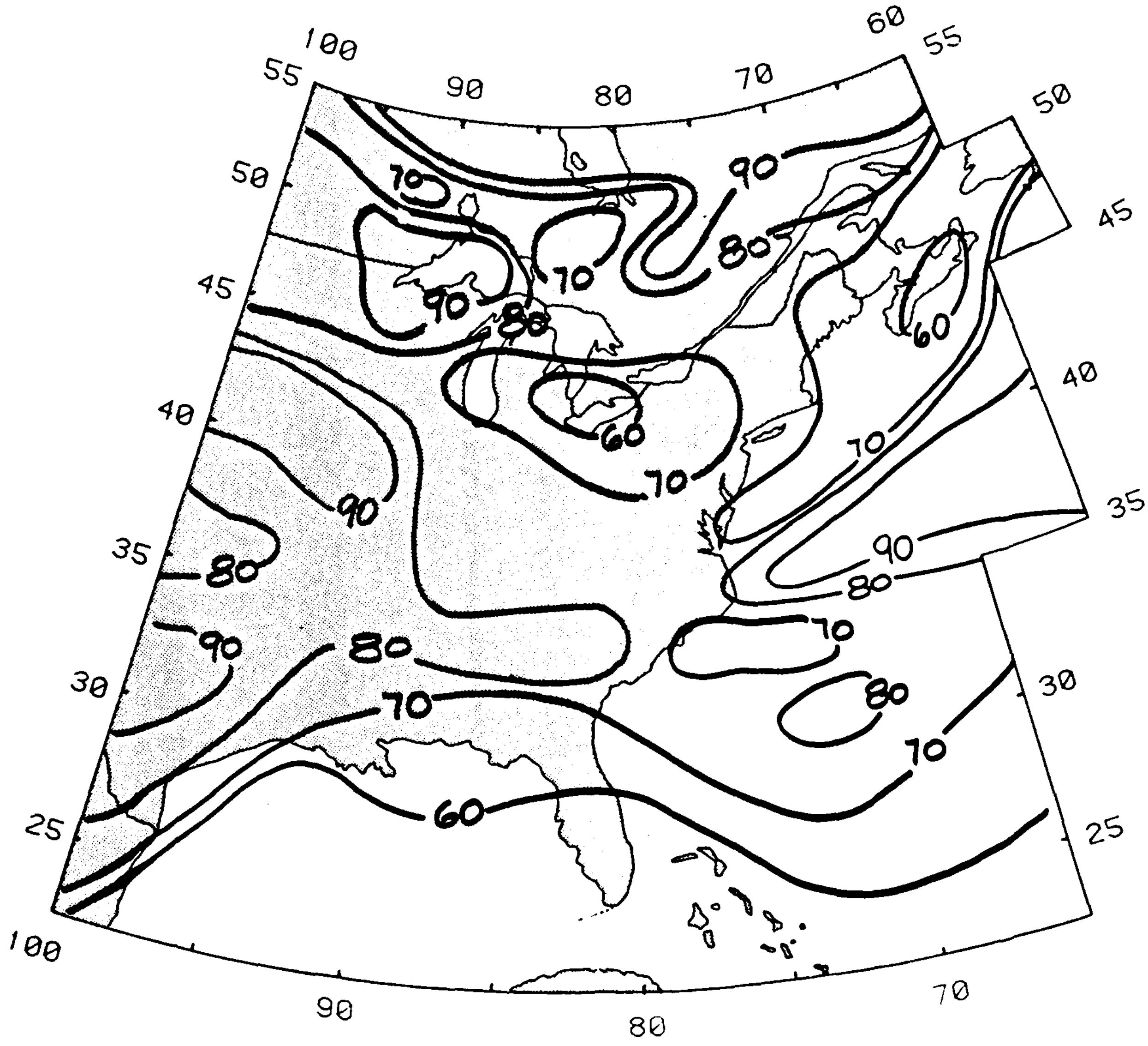


Fig. 17. April-September local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

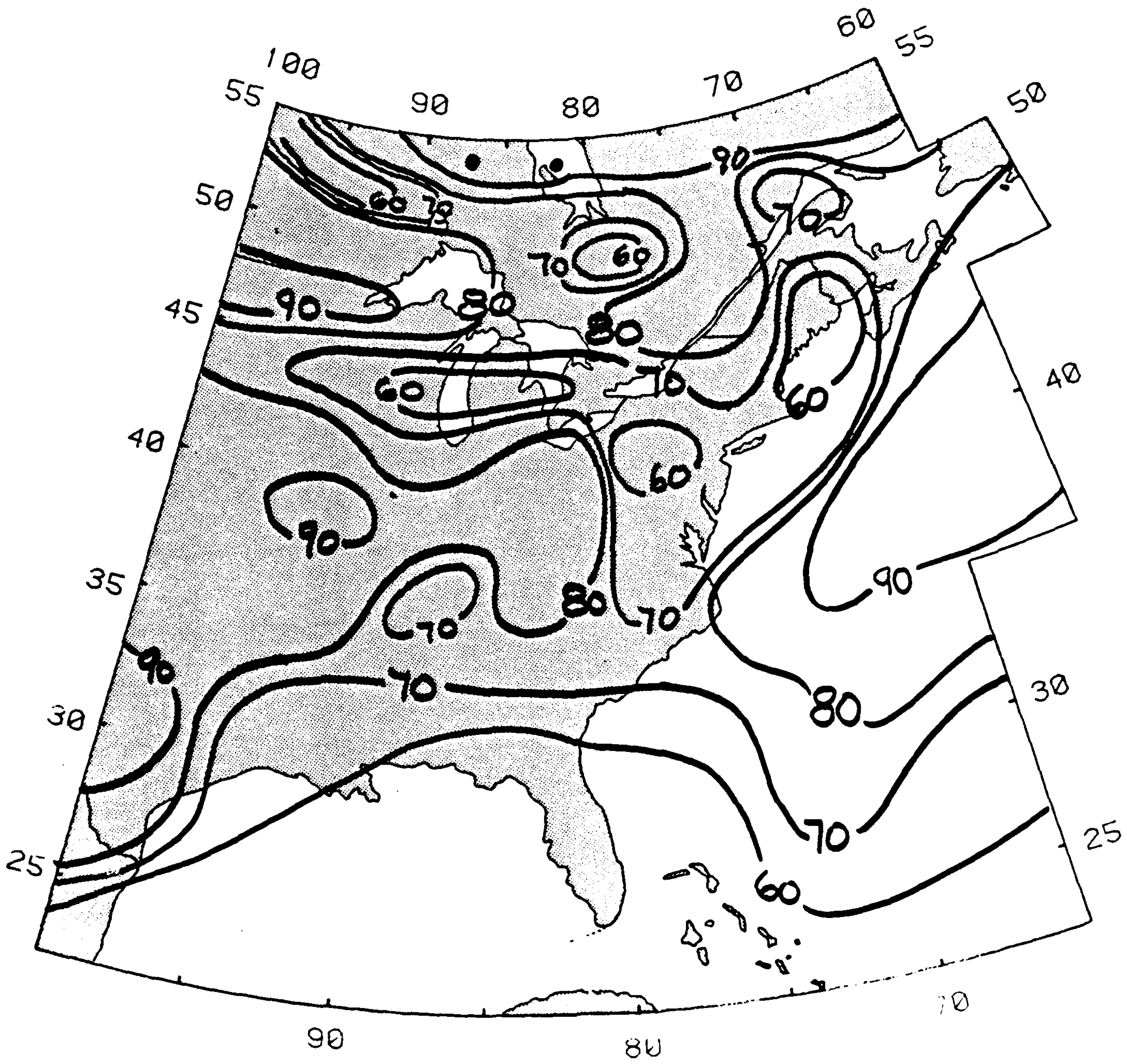


Fig. 18. May-October local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

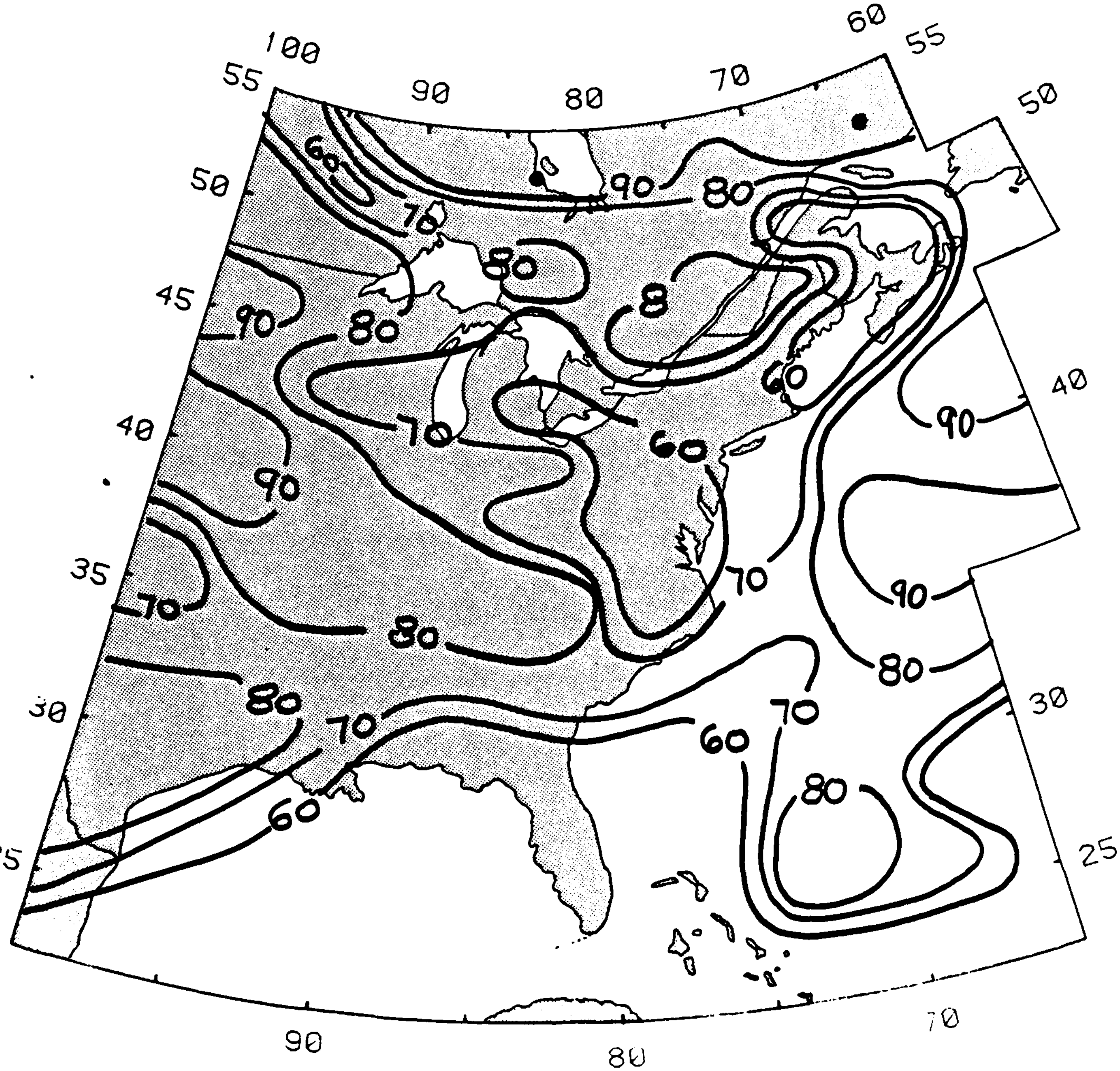


Fig. 19. June-November local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

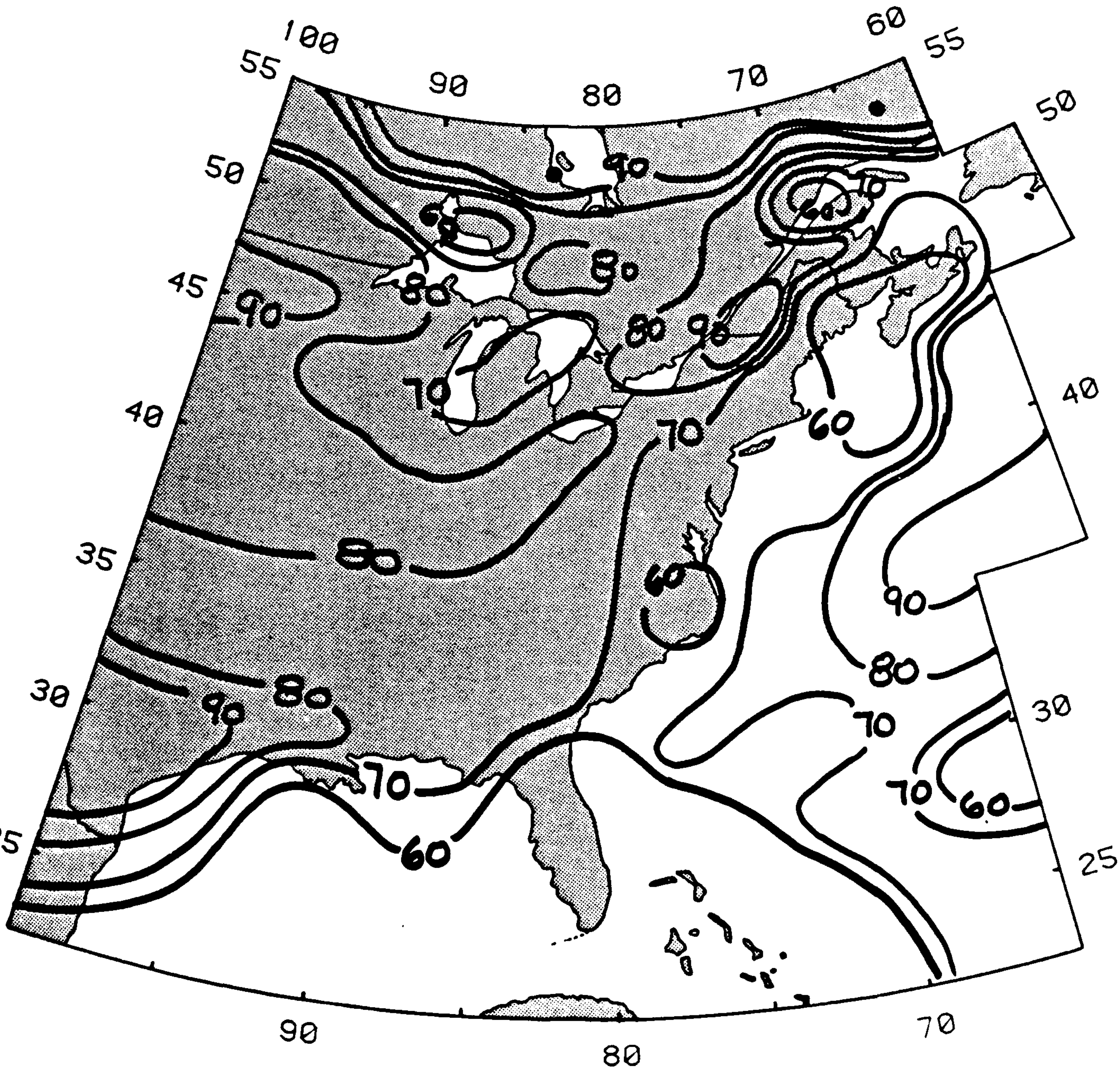


Fig. 20. July-December local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

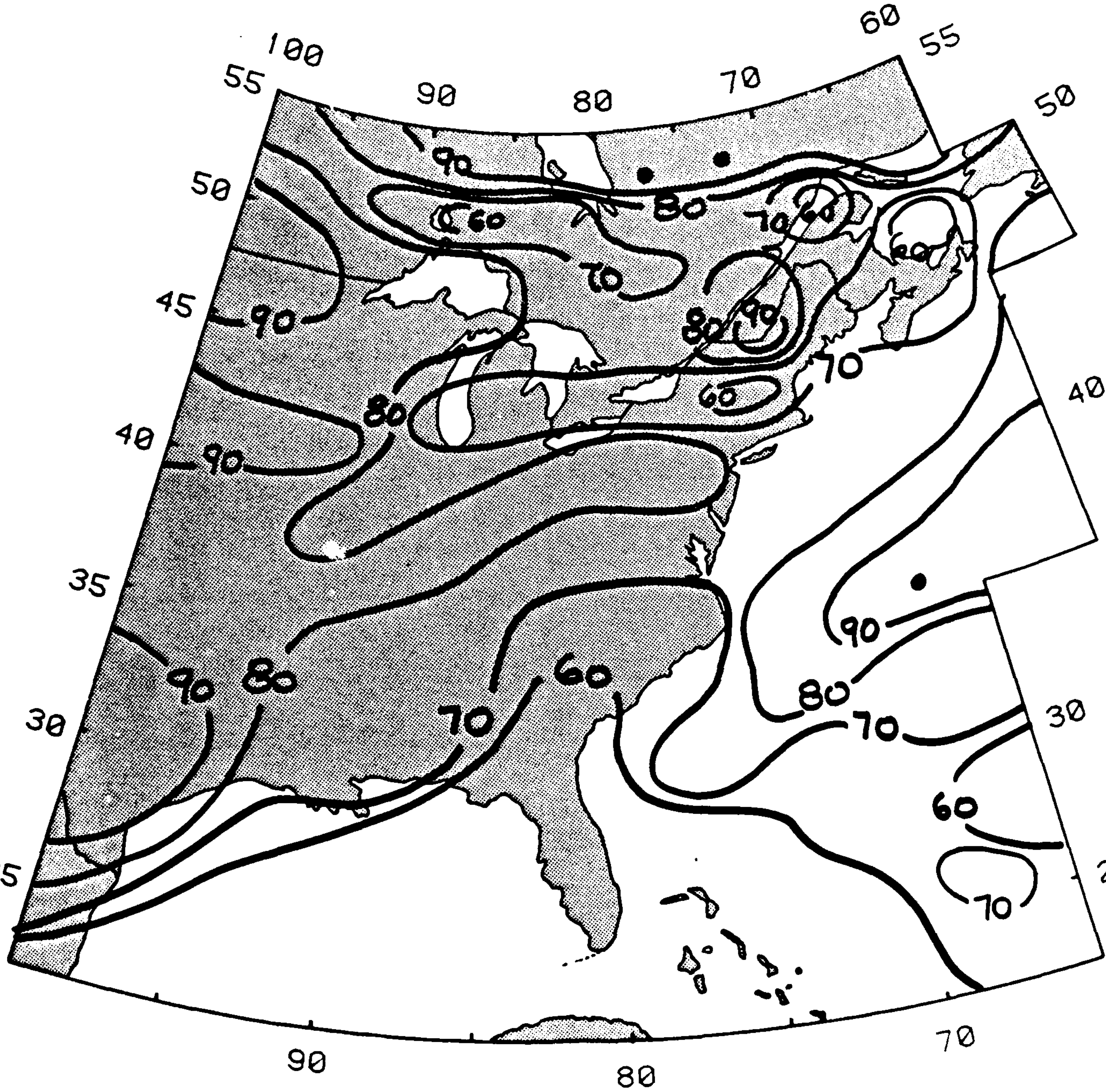


Fig. 21. August-January local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

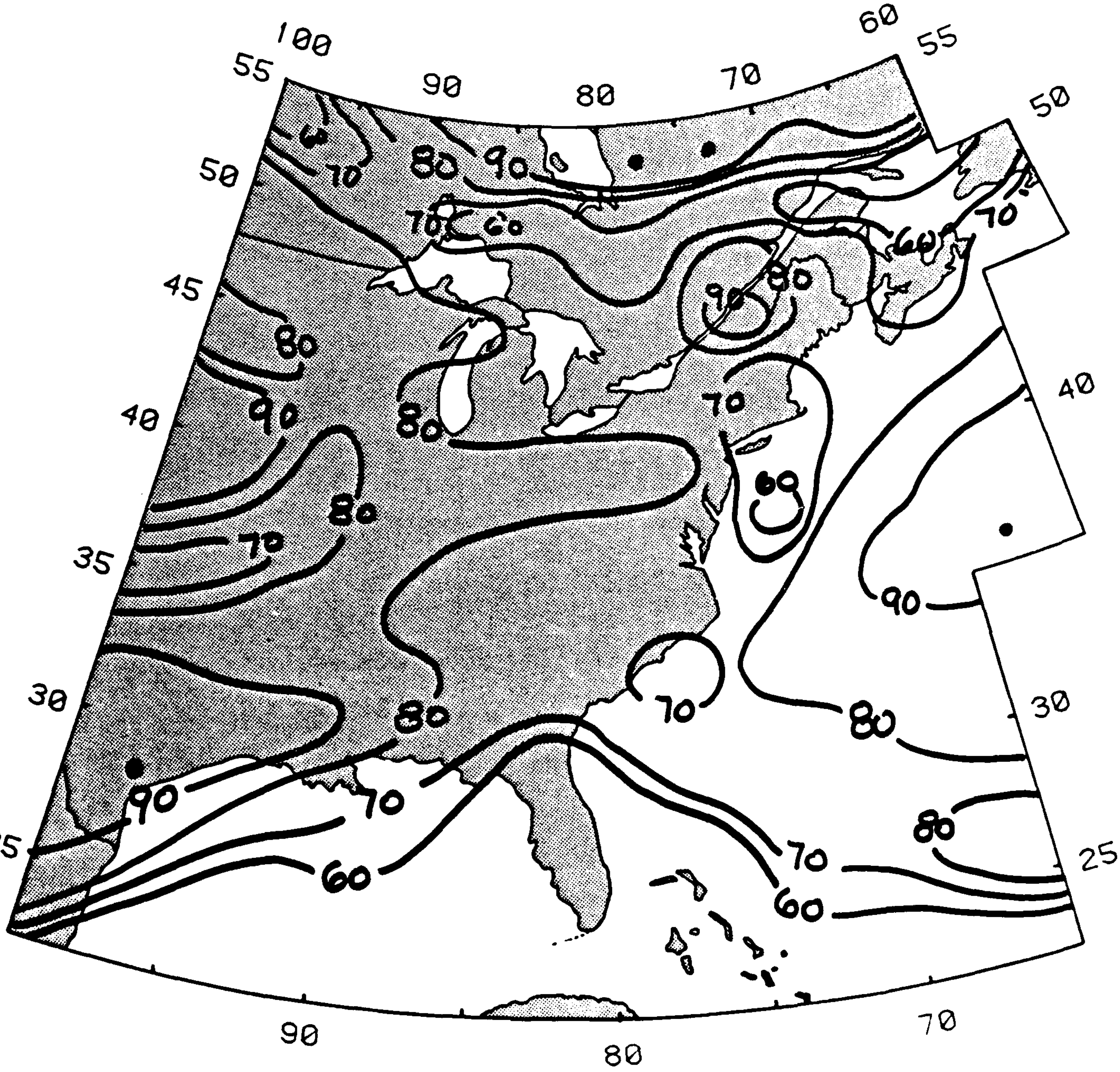


Fig. 22. September-February local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

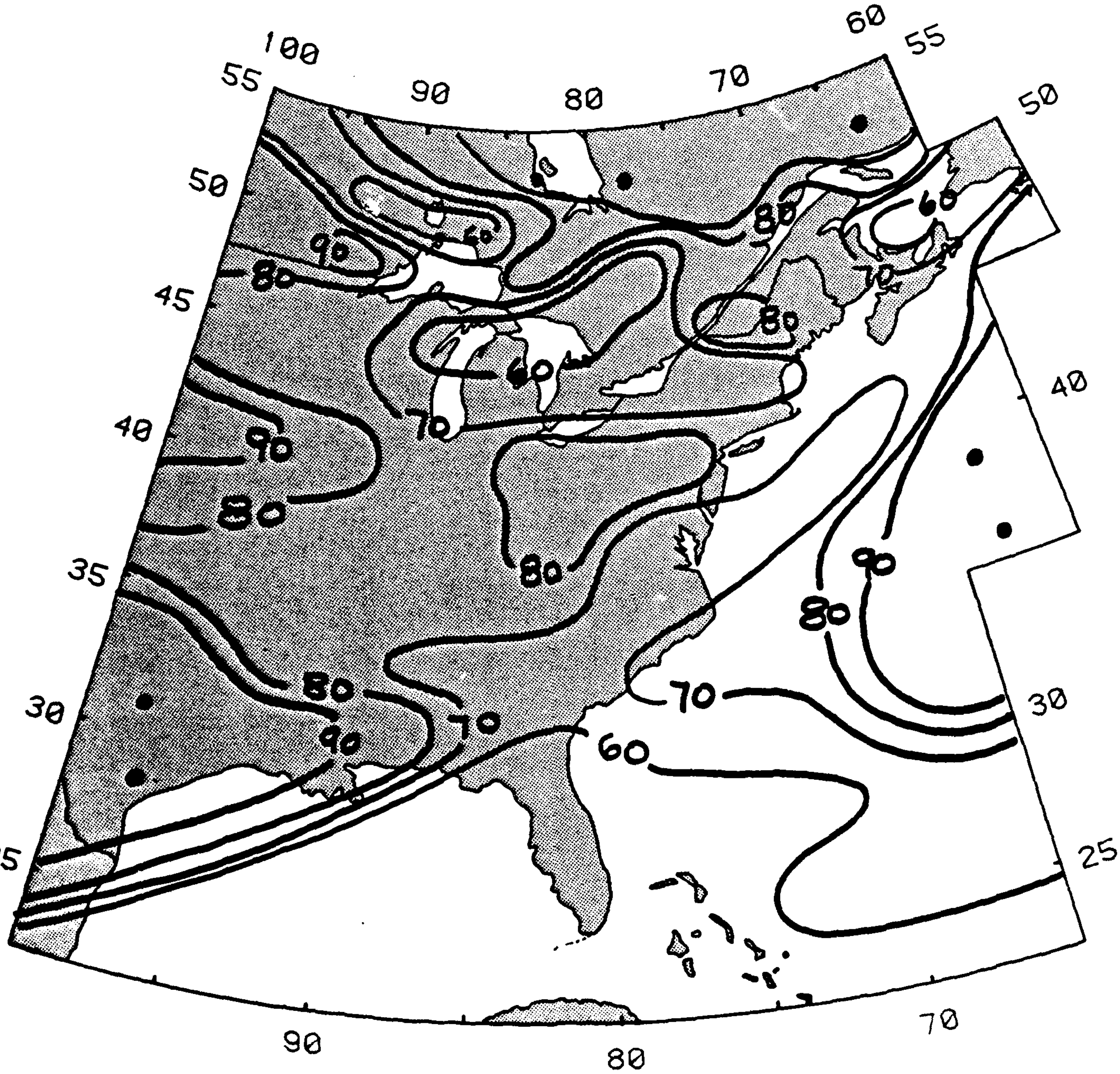


Fig. 23. October-March local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

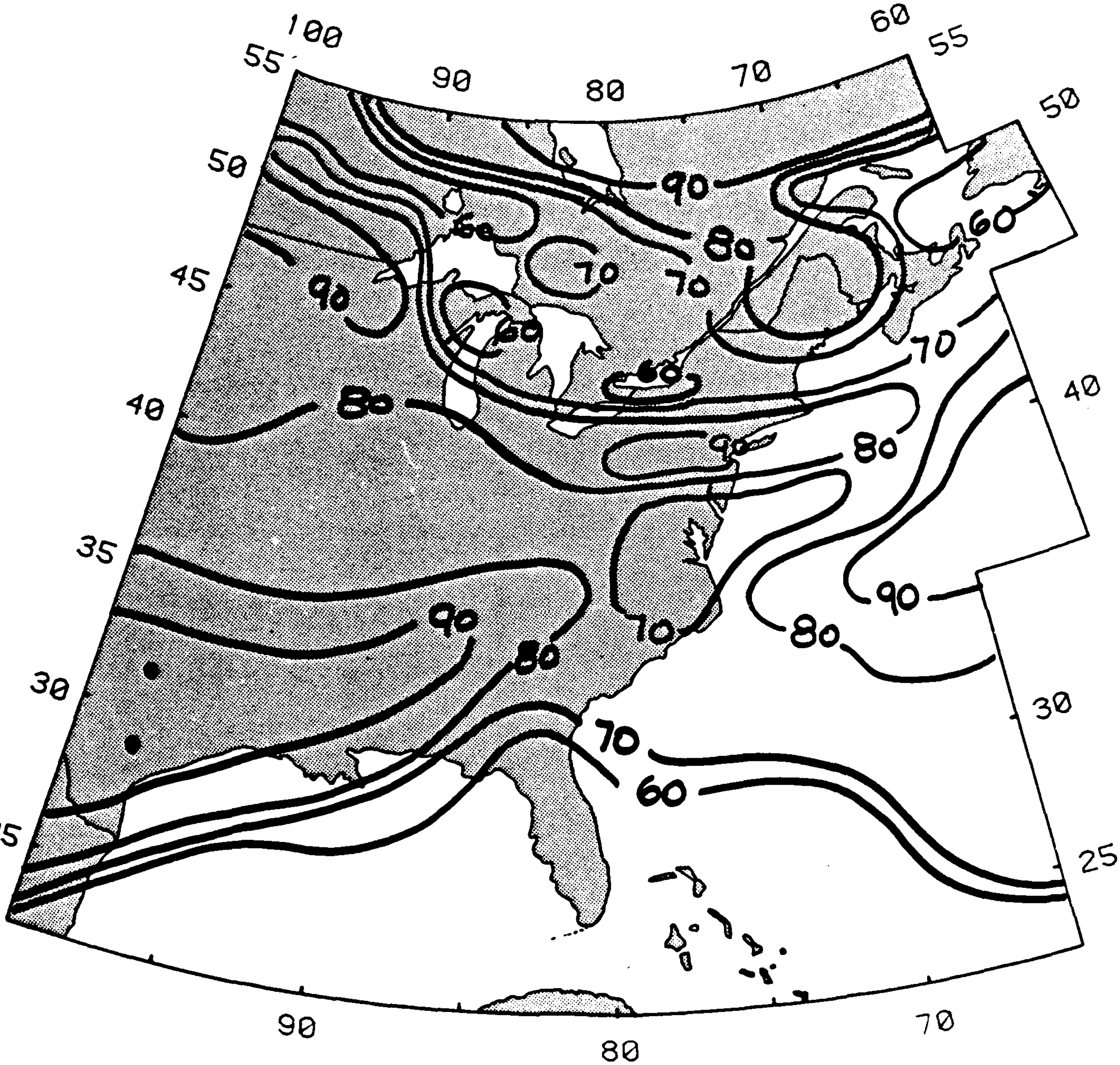


Fig. 24. November-April local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

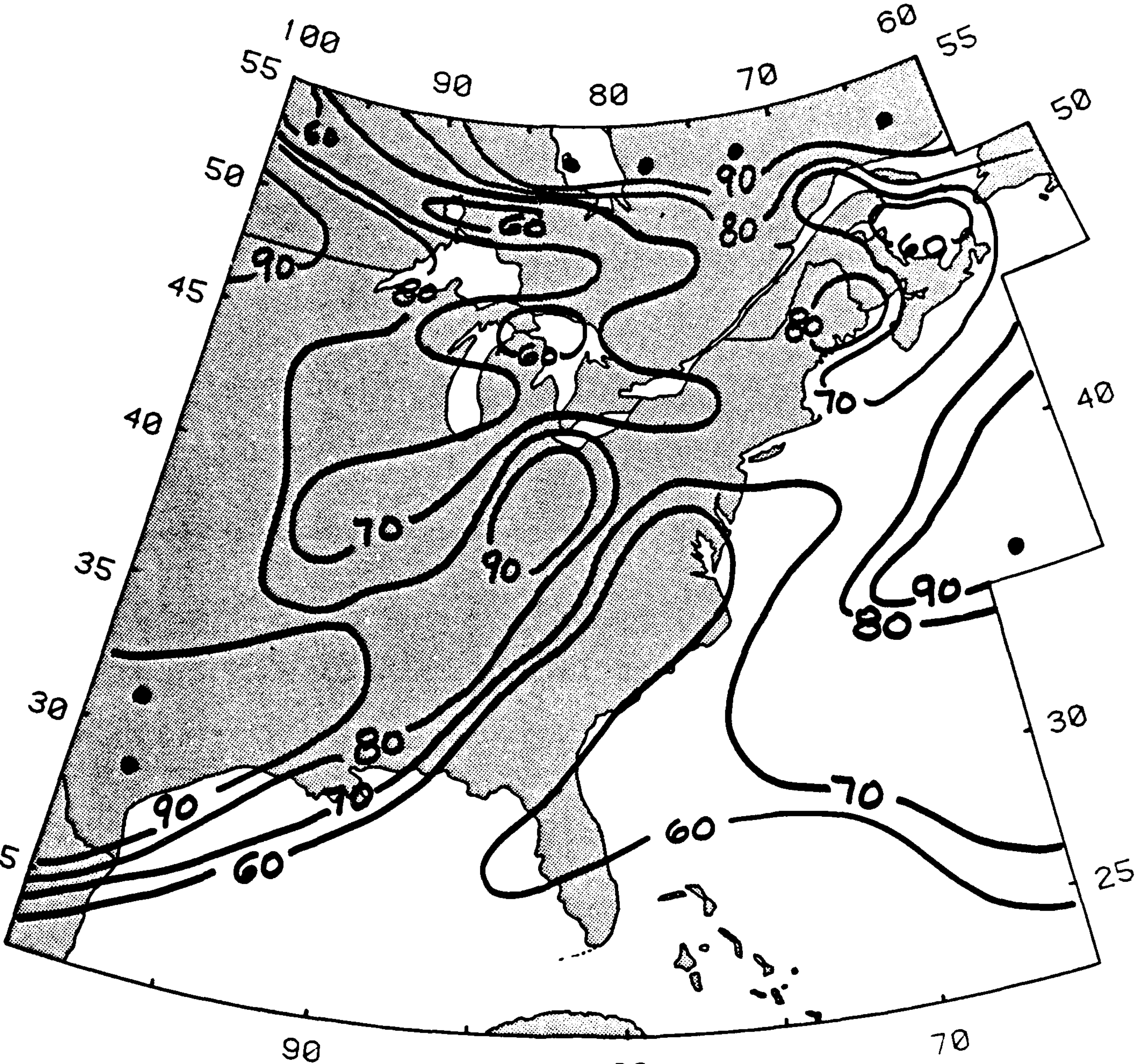


Fig. 25. November-April local percent correct skill scores (1960-1980) for Model I. Skills less than 60% are not contoured. 100% correct scores are indicated by black circles.

Global Skill

Figures 26, 27 and 28 show the global percent correct skill score by season and year for Models I, II and III. The three time series of forecast skill are similar in gross form as well as in most of the details. Some important differences are evident. Model II (Fig. 27) had a failure in the mid-1960s that was not present in Model I or Model III (Figs. 26 and 28). Model III had a failure in 1978 that was not evident in either Models I or II. The failure in mid-1975 is present in all three models but Model III was clearly the best forecast of the three that season. In contrast, peaks in the three curves are congruent. These differences are important in that by running all three models for each forecast differences will be revealed and possible forecast failure may be forewarned.

The most serious kind of forecast failure is the general decline in forecast skill. Such a depression of skill occurred in the mid-1970s and lasted about three years. During this three-year period the numbers of cyclones increased and the variability in cyclone numbers also increased. Apparently a mode of variation occurred that the models were not able to predict. In earlier studies (Hayden, 1981a) we used a jackknife procedure

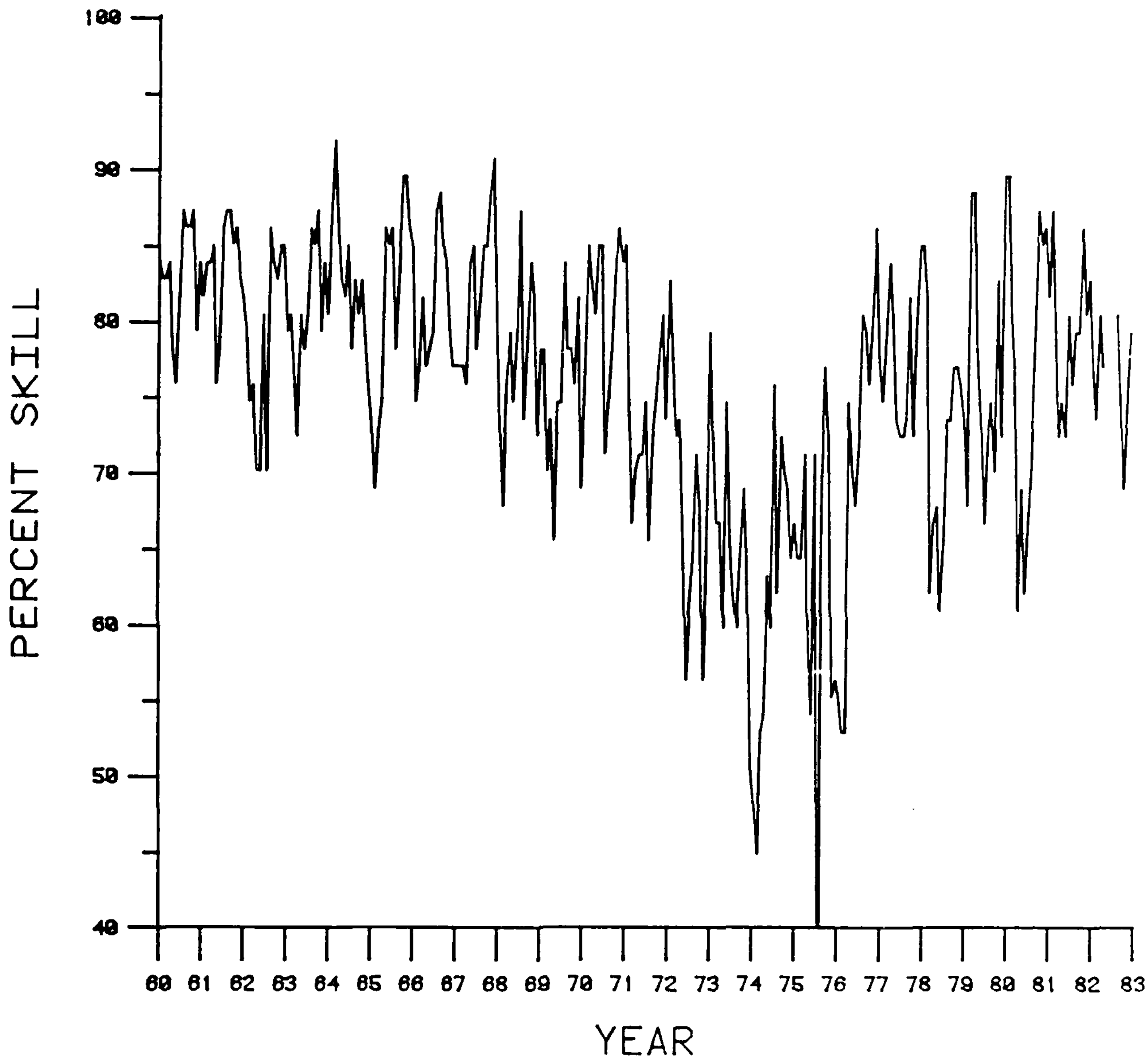


Fig. 26. Model I global percent correct skill scores (1960-1983) by season and year.

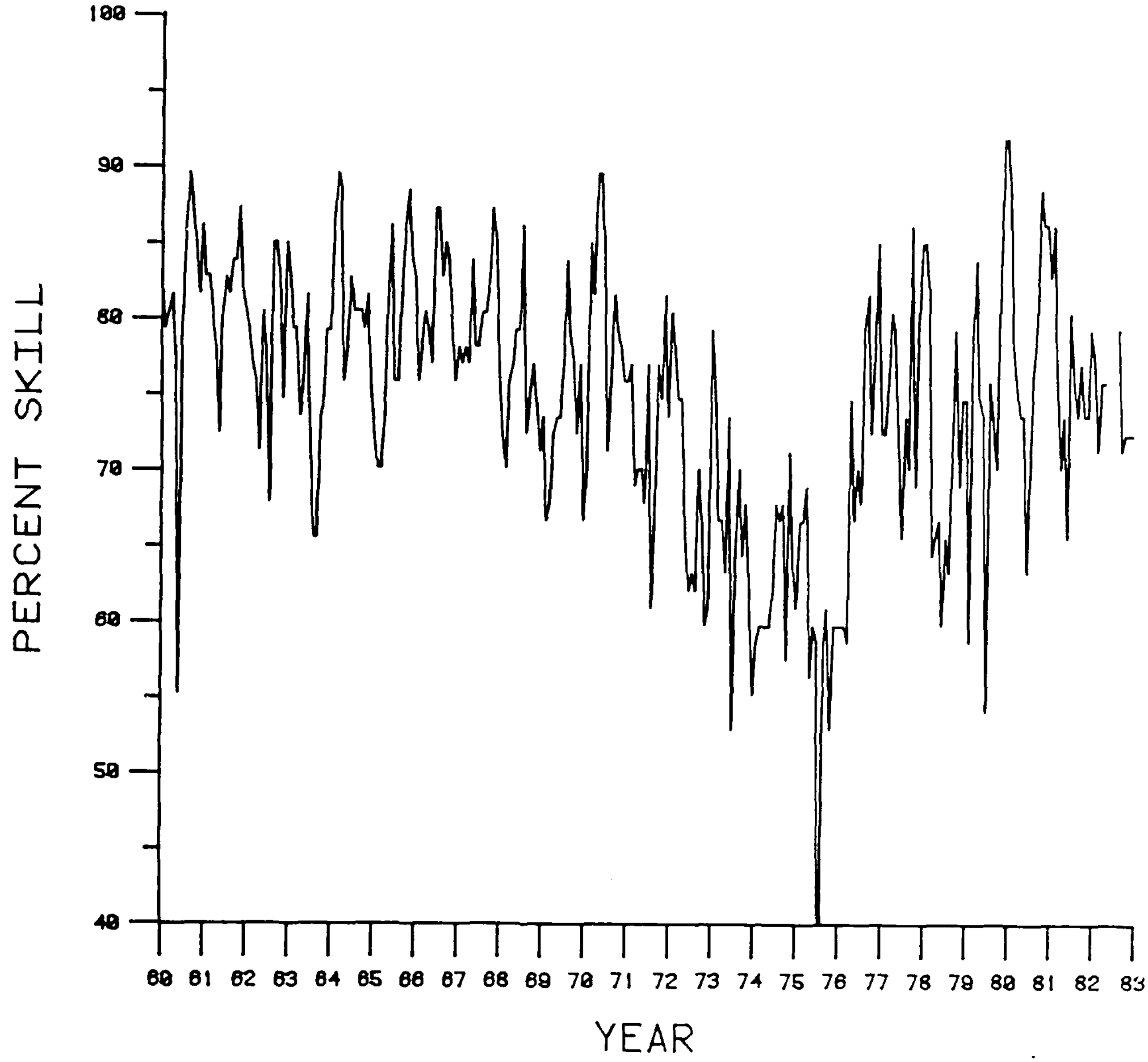


Fig. 27. Model II global percent correct skill scores (1960-1983) by season and year.

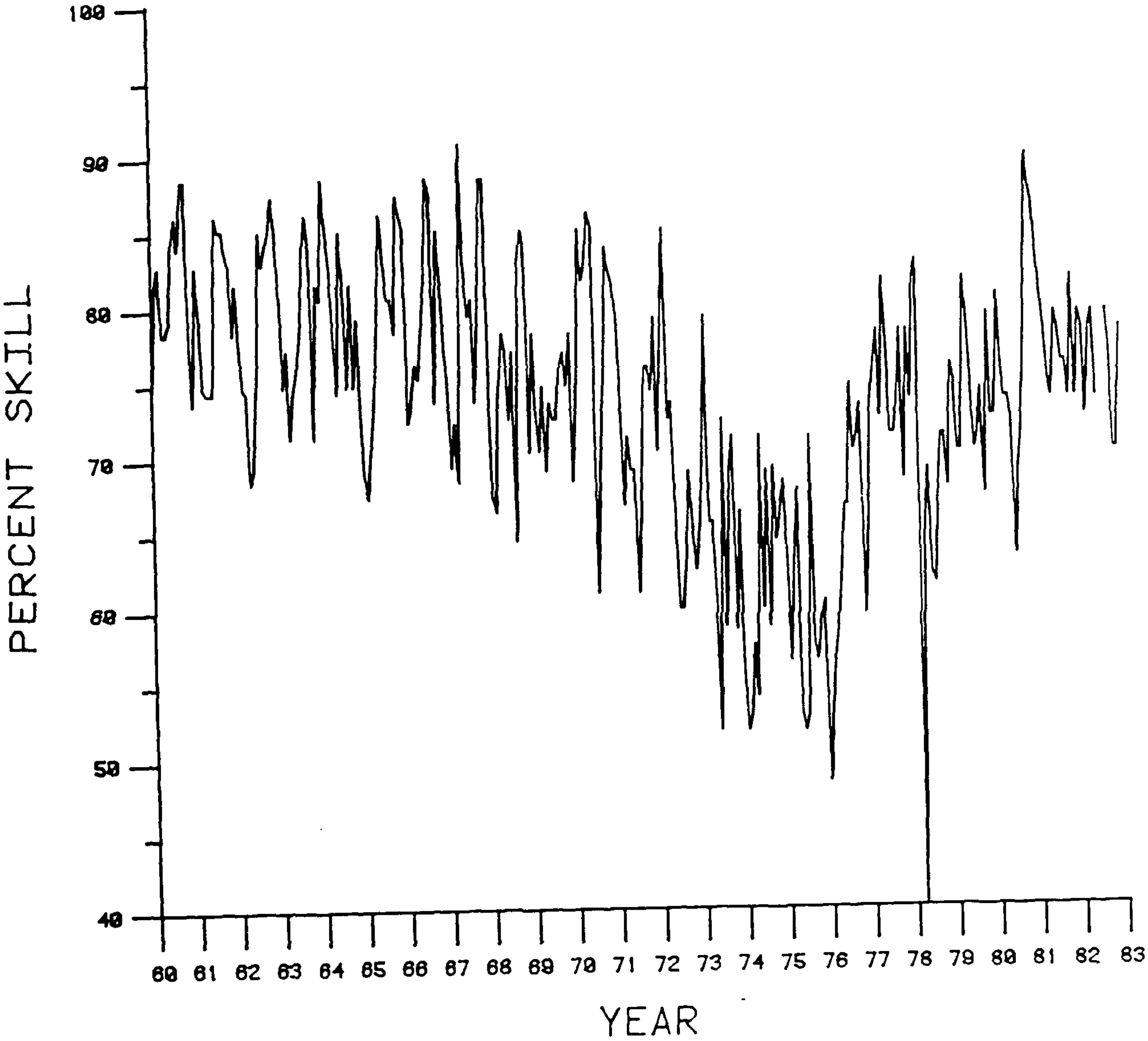


Fig. 28. Model III global percent correct skill scores (1960-1983) by season and year.

to predict October-March seasons from 1885 to 1979. No comparable period of poor forecast skill was found. It is probable that the period 1973-1976 was anomalous relative to the 1885-1960 period. In effect the mid-1970s anomaly has no counterpart in the training data. Following the decline, forecast skills returned to the high levels of the earlier part of the record.

Global percent correct skill scores for Model versions I, II and III for each of the 12 seasons forecasted are presented in Table IV. Scores for the independent data period (1960-1980) are given. Values in the parentheses in Table IV and in subsequent tables are the percent skill scores for the hindcast and forecast periods taken jointly. A strong seasonal trend in forecast skill is not present. Forecasts which include the three summer months tend to have a slightly lower score than those that include no summer months. It is not clear why the differences are significant. On an average basis a global skill score of about 75% is indicated. Model III out-performs Models I and II by several percentage points. The highest score earned (77.3%) and the highest low skill score earned (77.8%) are found for Model III. Scores that include the forecasts from the operational period are higher than those for the hindcast period alone.

The greatest discrepancy between the three models is found for the January to June forecast season. Skill in Model I was 67.8% while Model III had a score of 77.3%. No other case of such an extreme difference is found. A range of 2 to 3% is common. Model stability is indicated across seasons, from model to model and from hindcast to operational forecast periods.

TABLE IV

Global Percent Correct Skill Scores 1960-1980

SEASON	MODEL I	MODEL II	MODEL III
JANUARY-JUNE	67.8(69.1)	75.4(76.1)	77.3(77.8)
FEBRUARY-JULY	75.2(75.1)	74.7(74.7)	76.0(76.1)
MARCH-AUGUST	75.2(75.2)	74.8(74.7)	76.6(76.0)
APRIL-SEPTEMBER	75.2(75.2)	72.7(72.9)	73.8(73.8)
MAY-OCTOBER	72.4(72.4)	73.0(72.0)	75.5(74.7)
JUNE-NOVEMBER#	70.9(71.1)	71.1(7.15)	74.5(74.4)
JULY-DECEMBER#	72.4(72.4)	73.0(73.2)	75.2(75.0)
AUGUST-JANUARY	72.5(73.1)	75.5(75.7)	75.8(76.2)
SEPTEMBER-FEBRUARY	75.9(76.3)	75.6(76.1)	76.5(77.0)
OCTOBER-MARCH	74.5(75.2)	74.9(75.4)	75.2(75.8)
NOVEMBER-APRIL	75.6(75.7)	76.5(76.6)	76.7(77.2)
DECEMBER-MAY	75.2(75.8)	75.5(75.9)	76.6(77.2)
AVERAGE	73.6(73.9)	74.6(75.8)	75.9(76.0)

( ) = 1960-1983; # 1983 omitted

Heidke Skill Score

Global Heidke Skill Scores are a simple linear transform of the percent skill scores. Figures 29, 30 and 31 show the global Heidke skill scores by season and year for Models I, II and III. These time series are, in all respects excepting scale, identical

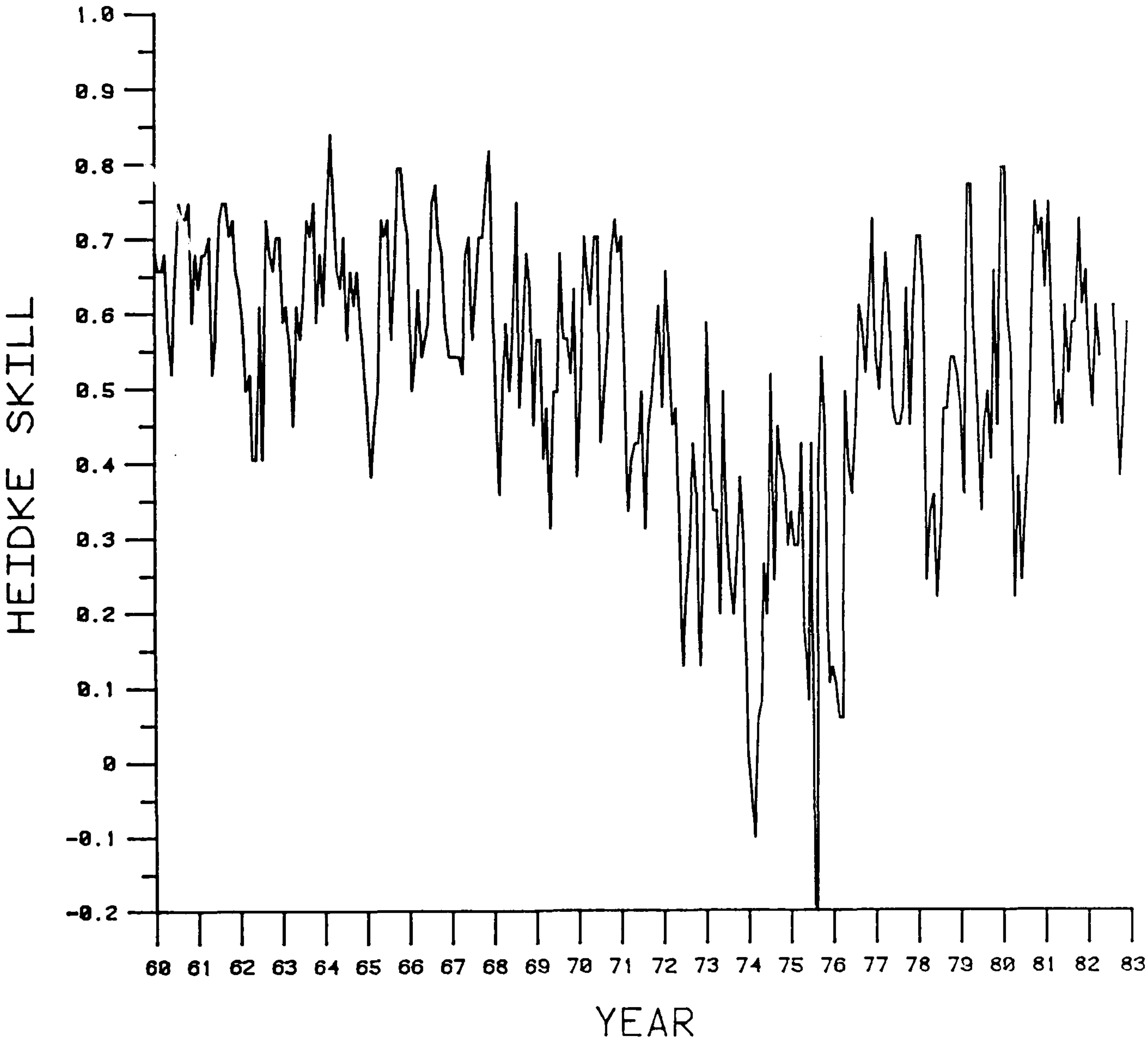


Fig. 29. Model I global Heidke skill scores (1960-1983) by season and year.

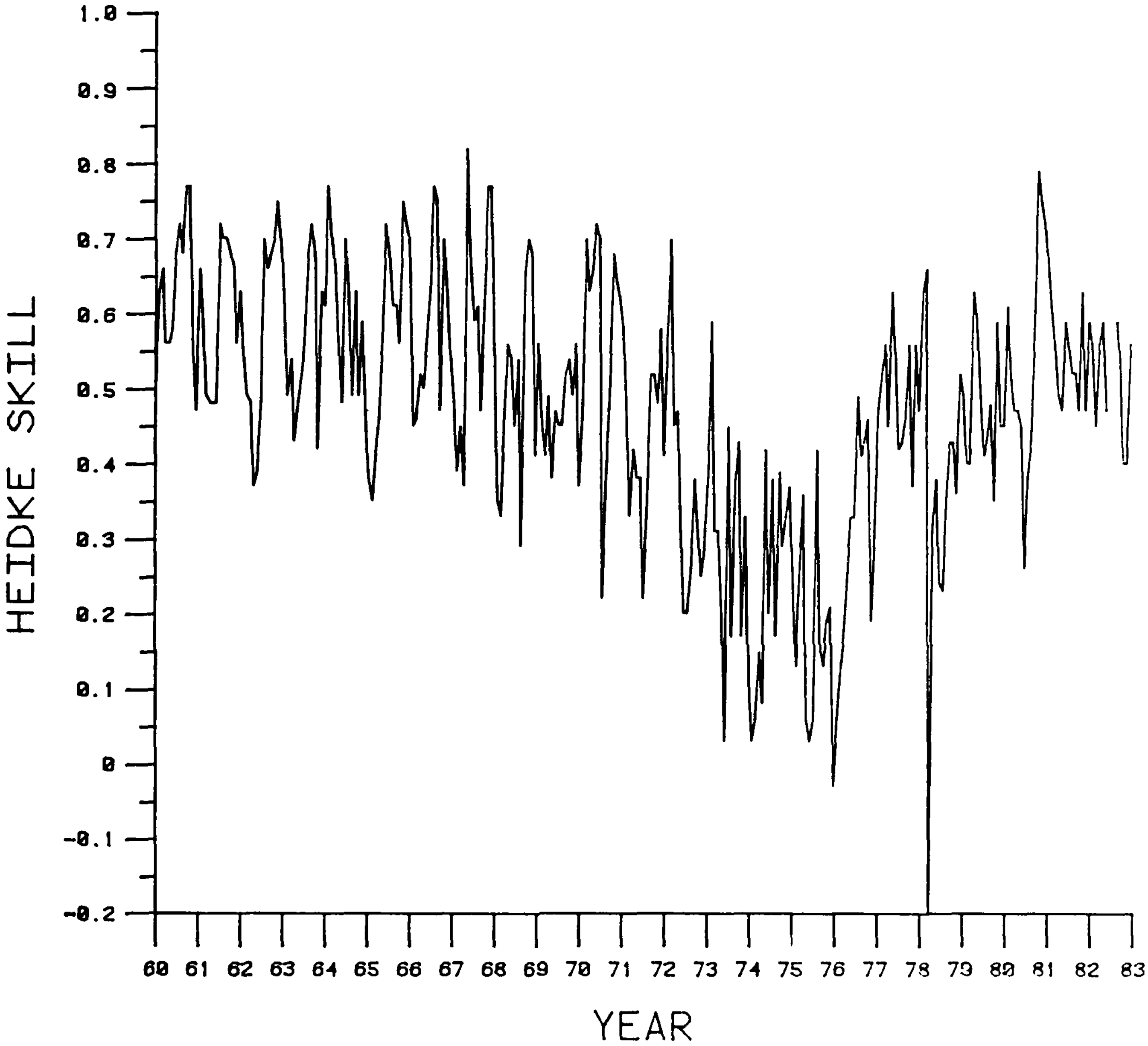


Fig. 30. Model II global Heidke skill scores (1960-1983) by season and year.

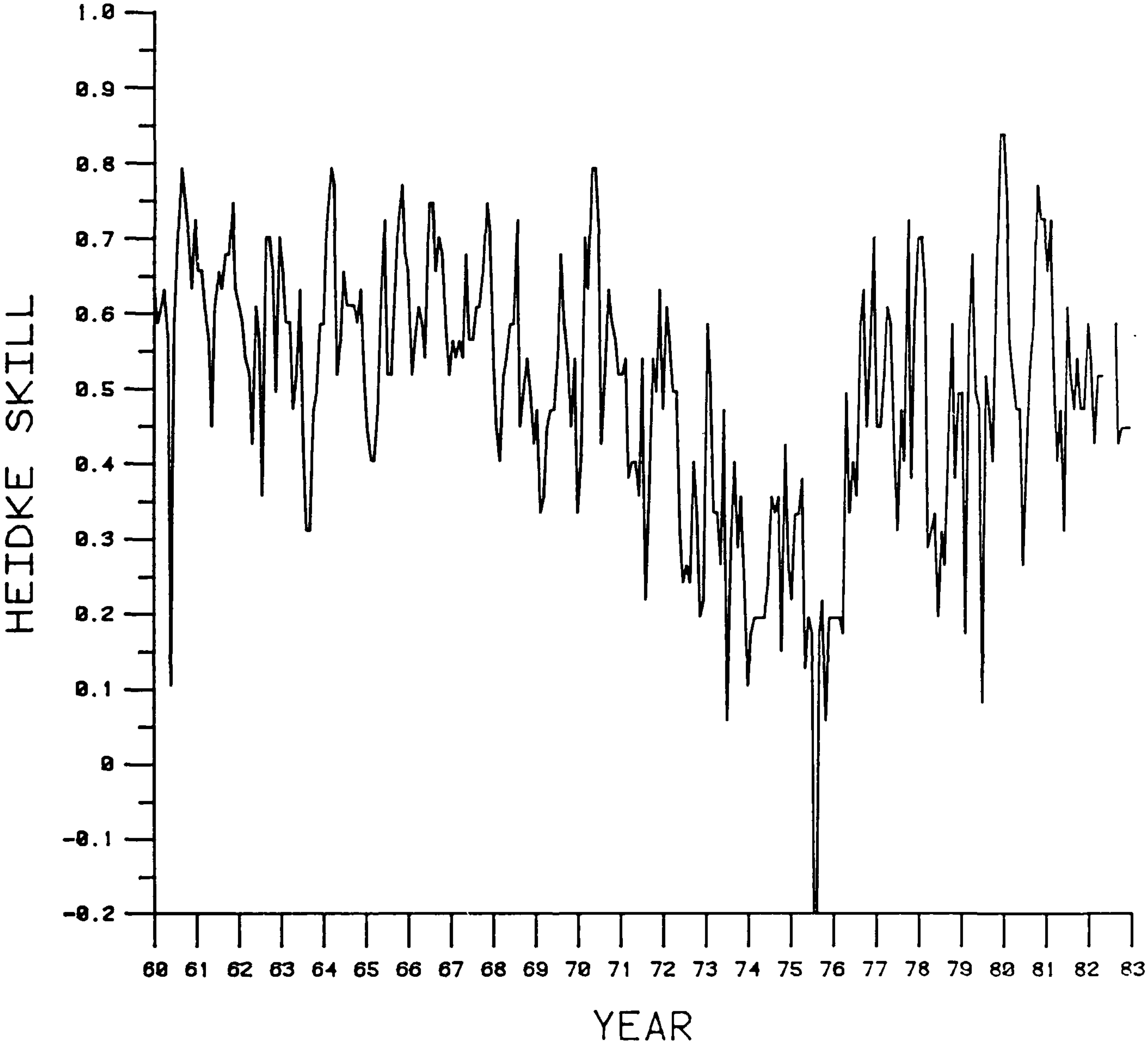


Fig. 31. Model III global Heidke skill scores (1960-1983) by season and year.

to Figures 26, 27 and 28. The comments on these earlier figures apply here as well. Table V gives the global Heidke skill scores for Models I, II and III for each of the 12 seasons and the hindcast and hindcast plus operational periods. The conclusions drawn from Table IV apply also to Table V.

TABLE V

Global Heidke Skill Scores 1960-1980

SEASON	MODEL I	MODEL II	MODEL III
JANUARY-JUNE	.355(.381)	.508(.521)	.545(.556)
FEBRUARY-JULY	.503(.502)	.493(.493)	.519(.522)
MARCH-AUGUST	.503(.503)	.495(.491)	.532(.519)
APRIL-SEPTEMBER	.502(.502)	.454(.458)	.475(.475)
MAY-OCTOBER	.448(.447)	.460(.464)	.510(.494)
JUNE-NOVEMBER#	.418(.422)	.422(.429)	.490(.488)
JULY-DECEMBER#	.444(.447)	.459(.464)	.504(.500)
AUGUST-JANUARY	.449(.463)	.509(.514)	.515(.525)
SEPTEMBER-FEBRUARY	.514(.525)	.513(.521)	.529(.540)
OCTOBER-MARCH	.492(.505)	.497(.507)	.504(.516)
NOVEMBER-APRIL	.511(.514)	.529(.531)	.535(.543)
DECEMBER-MAY	.505(.517)	.510(.517)	.532(.544)
AVERAGE	.472(.478)	.492(.516)	.518(.520)

( ) = 1960-1983; # 1983 omitted

Deviation Skill Score

Global deviation skill scores by season and year are given in Figures 32, 33 and 34. In general, scores are high and are always

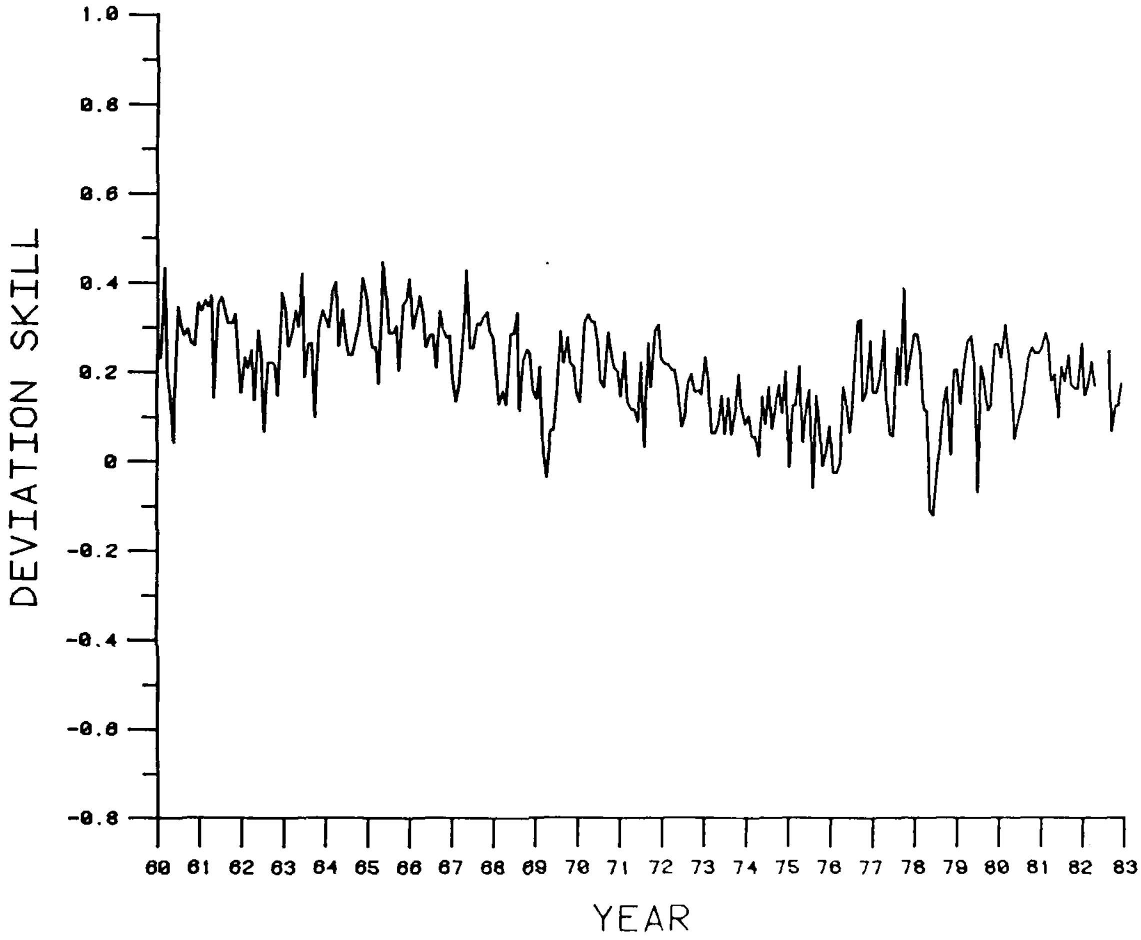


Fig. 32. Model I global deviation skill scores (1960-1983) by season and year.

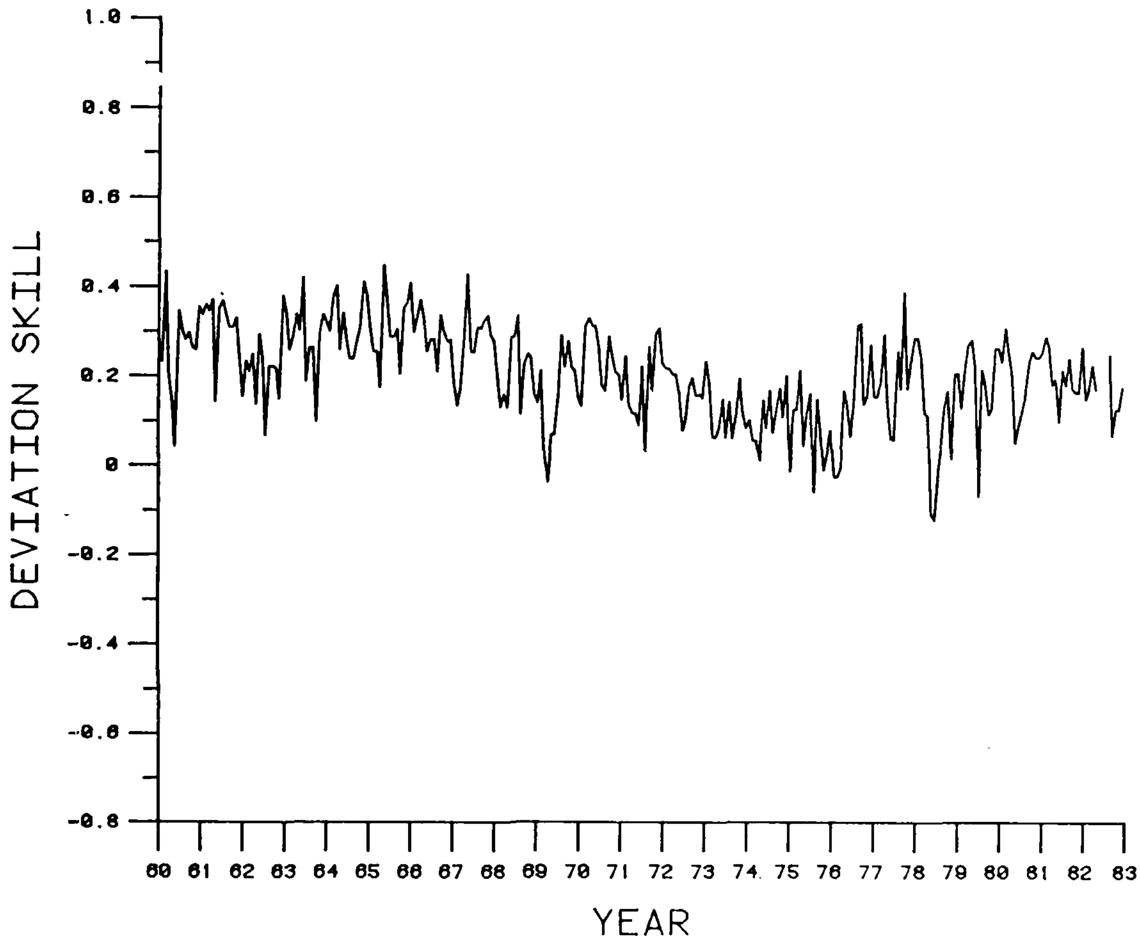


Fig. 33. Model II global deviation skill scores (1960-1983) by season and year.

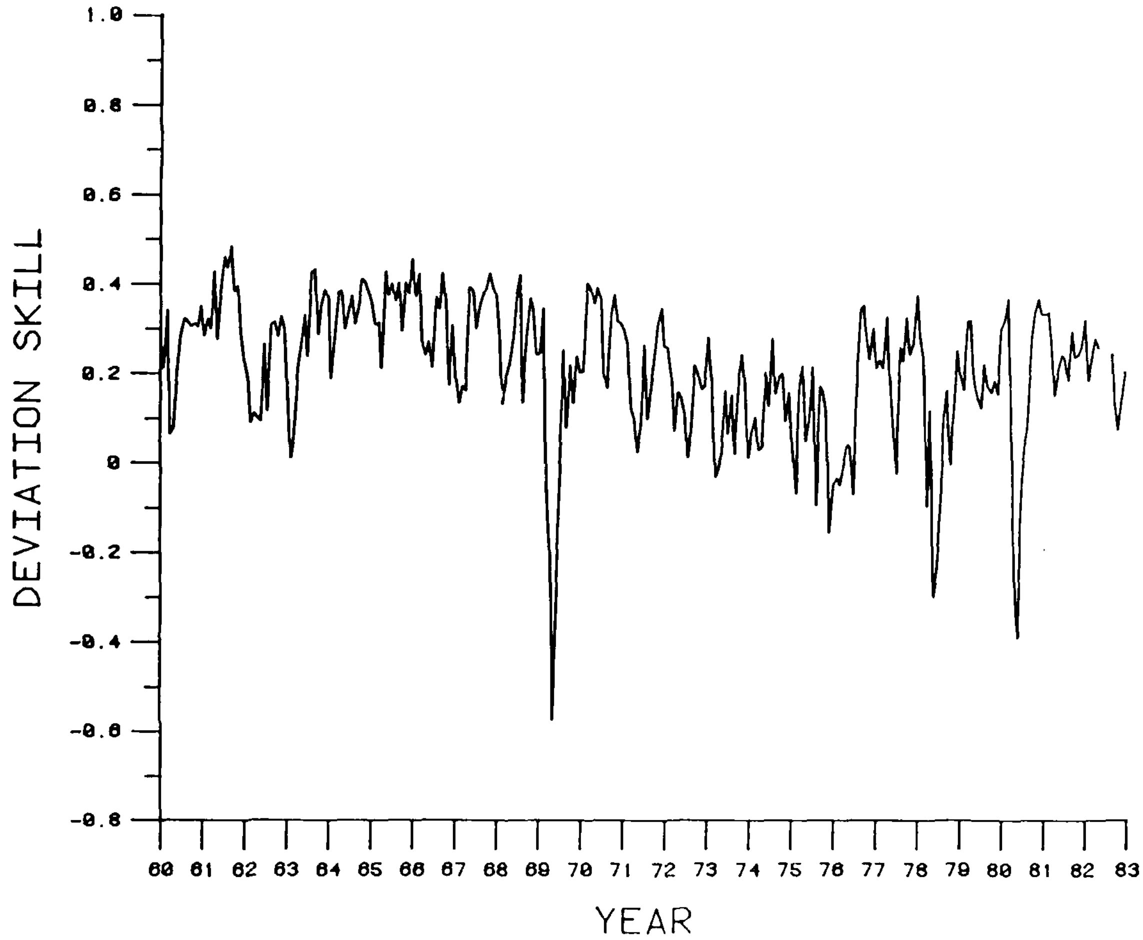


Fig. 34. Model III global deviation skill scores (1960-1983) by season and year.

positive. Negative scores indicating forecast failures occurred in only 5% of the 286 forecasts (Fig. 32). There is then almost no possibility that this outcome could have occurred by chance alone. The time histories of the deviation skill scores for the three models are similar in gross form. The variability in scores is higher in Model III (Fig. 34) than in Models I and II. It is also apparent that model failures are not common from model to model. It follows that when the three models agree it is likely that the forecast will not fail and that when they differ fundamentally it is prudent to "believe" the two that are most similar.

Global deviation skill scores by model and season are given in Table VI. Most deviation skill scores fall between .19 and .23. While these skill scores are modest given the possible maximum score of 1.0 they indicate real magnitude forecast skill. These values are lower than the Heidke scores. In the deviation skill score the penalties are a function of the size of the forecast error. Large errors lower forecast skill more than small errors. The average deviation of the model is about 80% as large as the average deviation of the mean as the forecast.

There are no discernible patterns across seasons or between Models I, II and III. The hindcast and operational forecast period deviation skill scores are essentially the same.

TABLE VI

Global Deviation Skill Scores 1960-1980

SEASON	MODEL I	MODEL II	MODEL III
JANUARY-JUNE	.147(.158)	.213(.217)	.241(.248)
FEBRUARY-JULY	.217(.221)	.212(.214)	.215(.214)
MARCH-AUGUST	.224(.223)	.240(.235)	.210(.190)
APRIL-SEPTEMBER	.238(.224)	.190(.184)	.168(.150)
MAY-OCTOBER	.200(.195)	.206(.192)	.224(.215)
JUNE-NOVEMBER#	.202(.197)	.195(.193)	.241(.232)
JULY-DECEMBER#	.224(.219)	.204(.200)	.232(.218)
AUGUST-JANUARY	.207(.216)	.228(.230)	.212(.220)
SEPTEMBER-FEBRUARY	.231(.231)	.188(.185)	.204(.208)
OCTOBER-MARCH	.208(.209)	.213(.206)	.194(.198)
NOVEMBER-APRIL	.225(.222)	.200(.197)	.219(.222)
DECEMBER-MAY	.210(.216)	.193(.198)	.235(.241)
AVERAGE	.211(.211)	.207(.204)	.215(.213)

( ) = 1960-1983; # 1983 omitted

Quadratic Skill Score

Global quadratic skill scores by season and year for Models I, II and III are given in Figures 35, 36 and 37. Because the numeric departures of the forecasts from observations are squared and summed in this measure of skill, variability in skill scores is higher than observed for the deviation score. In this regard Model II is superior to Models I and III. Careful examination of Figures 35, 36 and 37 reveals that the upper bound of the curves differs little from model to model. Good forecasts are equally

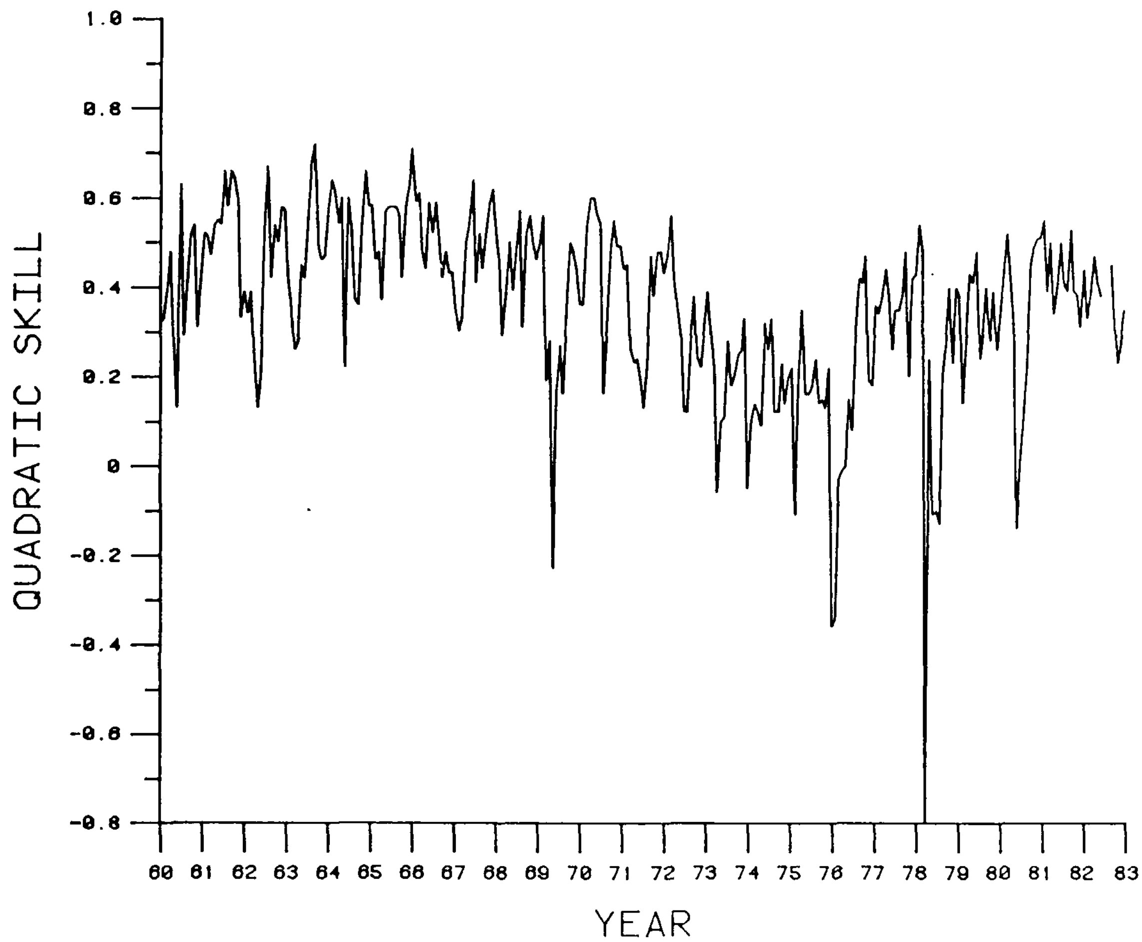


Fig. 35. Model I global quadratic skill scores (1960-1983) by season and year.

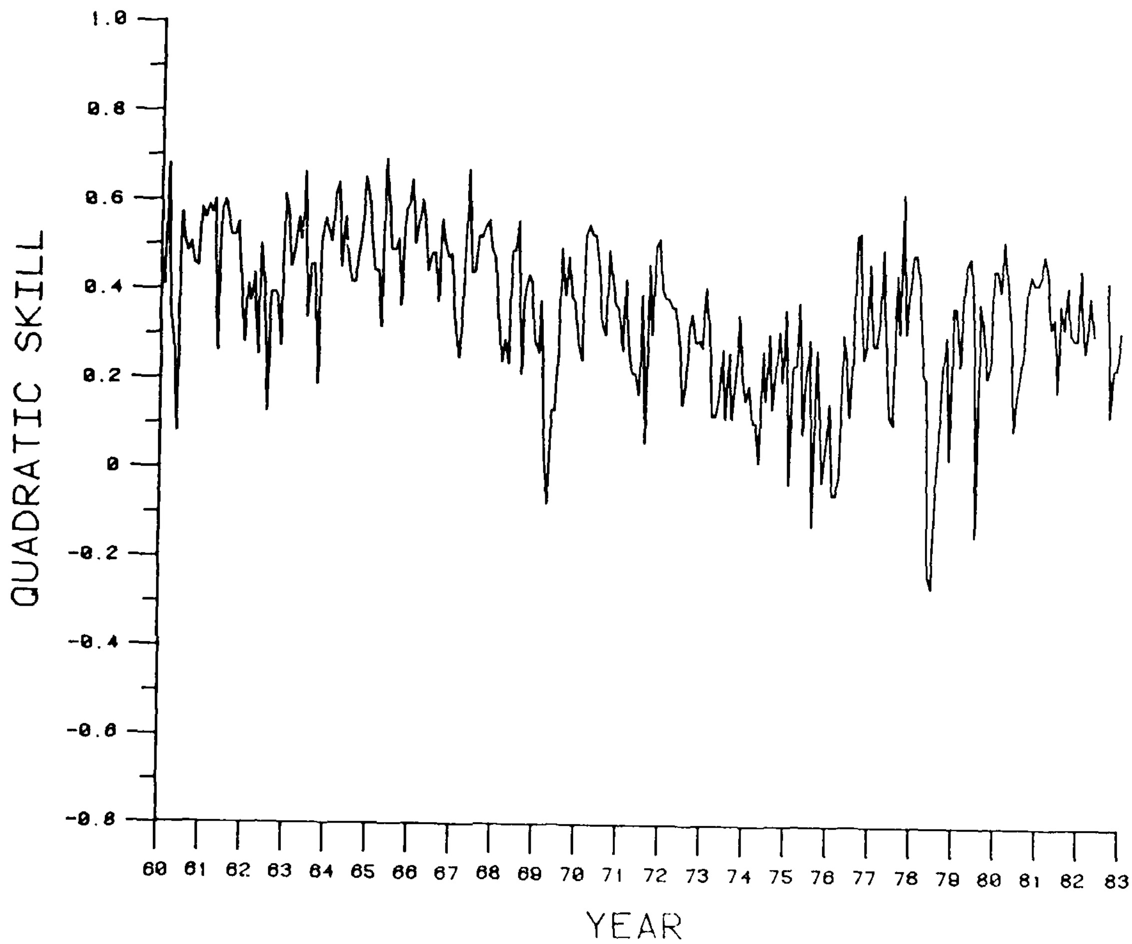


Fig. 36. Model II global quadratic skill scores (1960-1983) by season and year.

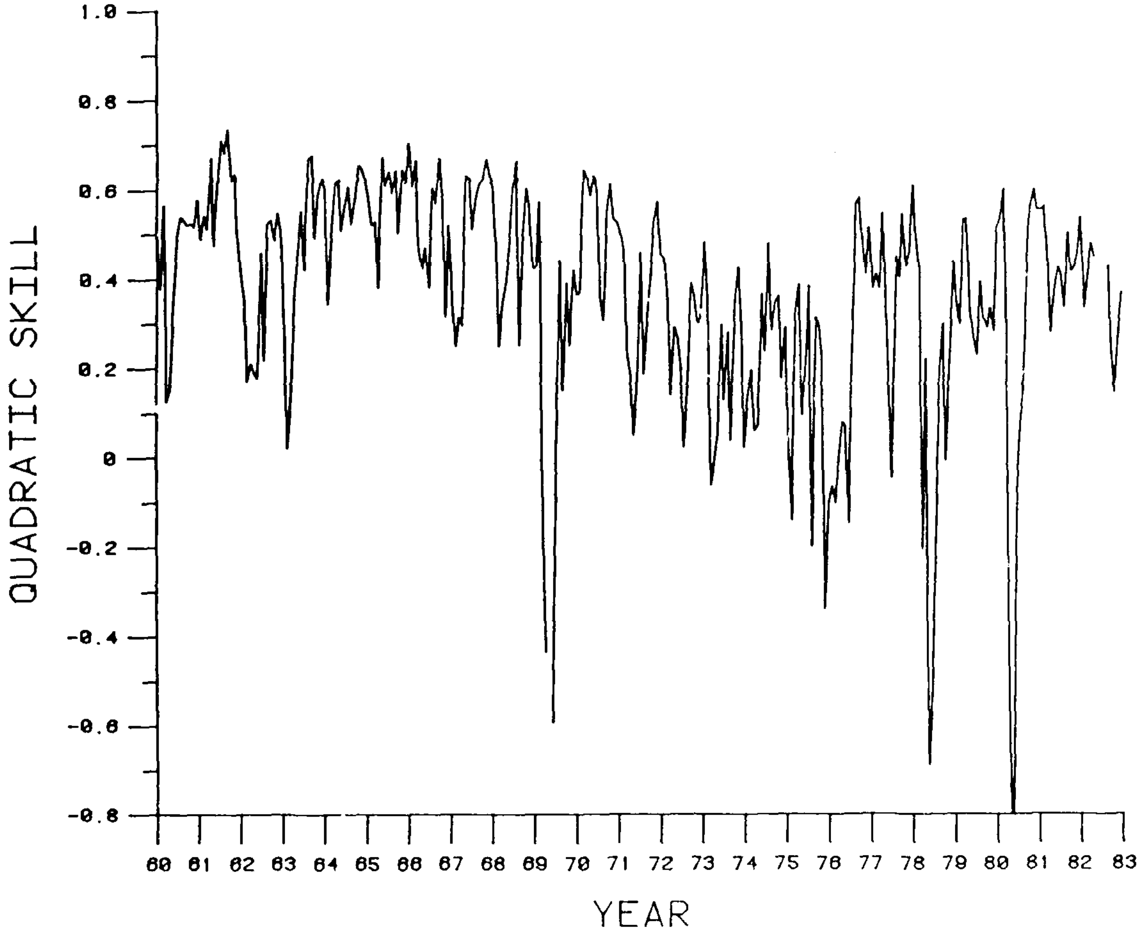


Fig. 37. Model III global quadratic skill scores (1960-1983) by season and year.

good from model to model. Poor forecasts, however, are poorer in Model I and Model III than in Model II. Fc failures in general do not occur in the same season and ye all three models. The differences among models confir wisdom of running all three types of models.

The gross trends in quadratic forecast skill are independent. The detail of the forecast failures vary from to model. Failures (J less than zero) are twice as com Model III as in Models I and II. However, forecast failur uncommon.

Three types of forecast failures can be defined: 1) ne categor/ forecast skill and positive numerical skill positive category skill and positive numerical skill; ar negative category skill and negative numerical skill. Wh category is correctly forecast but no numerical skill is pr large anomalies are usually present and the sign of the a is correct but the anomaly is so large that a large qu forecast error results. This is not a very serious erro most serious error occurs when the sign is incorrectly f and the quadratic skill is large and negative. This is t serious type of error. When the sign is poorly forecast quadratic skill score is high and positive it indicate small anomalies were forecast and small anomalies occur the sign was wrong. This type of error tends to happen w mean would have served as an excellent forecast.

Quadratic skill scores penalize the forecaster in proportion to the square of the error of the forecast. This is a severe penalty. The quadratic skill scores for Models I, II and III (Table IV) are uniformly higher than the deviation skill scores. This result is only possible if there is a preponderance of forecast errors between zero and unity. In this range, squaring results in a lower penalty value. Forecasts with errors less than 1 are rewarded. While this is a severe penalty, the quadratic skill scores for Models I, II and III (Table VII) are uniformly higher than the deviation skill scores. This result is only possible if there is a preponderance of forecast errors between zero and unity. In this range, squaring results in a lower penalty value. Forecasts with errors less than 1 are rewarded while errors greater than 1 are penalized. The largest quadratic skill score possible is 1.0. The minimum skill is technically minus infinity. The quadratic skill scores reported indicate that the models have real magnitude forecast skill. Quadratic skill scores, because the errors are squared, do not evaluate the sign of the forecast. Accordingly, quadratic skill scores should be used in conjunction with percent or Heidke skill scores. Quadratic skill scores reported in Table VII average 0.36. Values less than three and greater than four are uncommon. There is no seasonal cycle of quadratic forecast skill and there is no discernible difference between Models I, II and III.

Addition of the forecasts from the operational period resulted in no deterioration of forecast skill.

TABLE VII

Global Quadratic Skill Scores 1960-1980

SEASON	MODEL I	MODEL II	MODEL III
JANUARY-JUNE	.253(.273)	.373(.379)	.406(.418)
FEBRUARY-JULY	.378(.386)	.366(.371)	.368(.368)
MARCH-AUGUST	.386(.385)	.407(.401)	.347(.310)
APRIL-SEPTEMBER	.407(.384)	.328(.318)	.279(.239)
MAY-OCTOBER	.348(.342)	.357(.335)	.384(.368)
JUNE-NOVEMBER#	.355(.347)	.343(.340)	.417(.401)
JULY-DECEMBER#	.391(.383)	.356(.351)	.396(.383)
AUGUST-JANUARY	.358(.373)	.391(.395)	.355(.370)
SEPTEMBER-FEBRUARY	.390(.392)	.325(.321)	.343(.351)
OCTOBER-MARCH	.348(.351)	.362(.353)	.296(.307)
NOVEMBER-APRIL	.381(.379)	.351(.346)	.360(.368)
DECEMBER-MAY	.364(.375)	.340(.348)	.396(.407)
AVERAGE	.363(.364)	.358(.354)	.357(.357)

( ) = 1960-1983; # 1983 omitted

Average Absolute Errors

Figures 38, 39 and 40 show the global average absolute error of each of the three models by season and year. A seasonal cycle in error size is clearly present. Errors are larger in winter than in summer. This cycle is also present in the charts of the mean as a forecast (Fig. 41). This cycle is not due to the nature

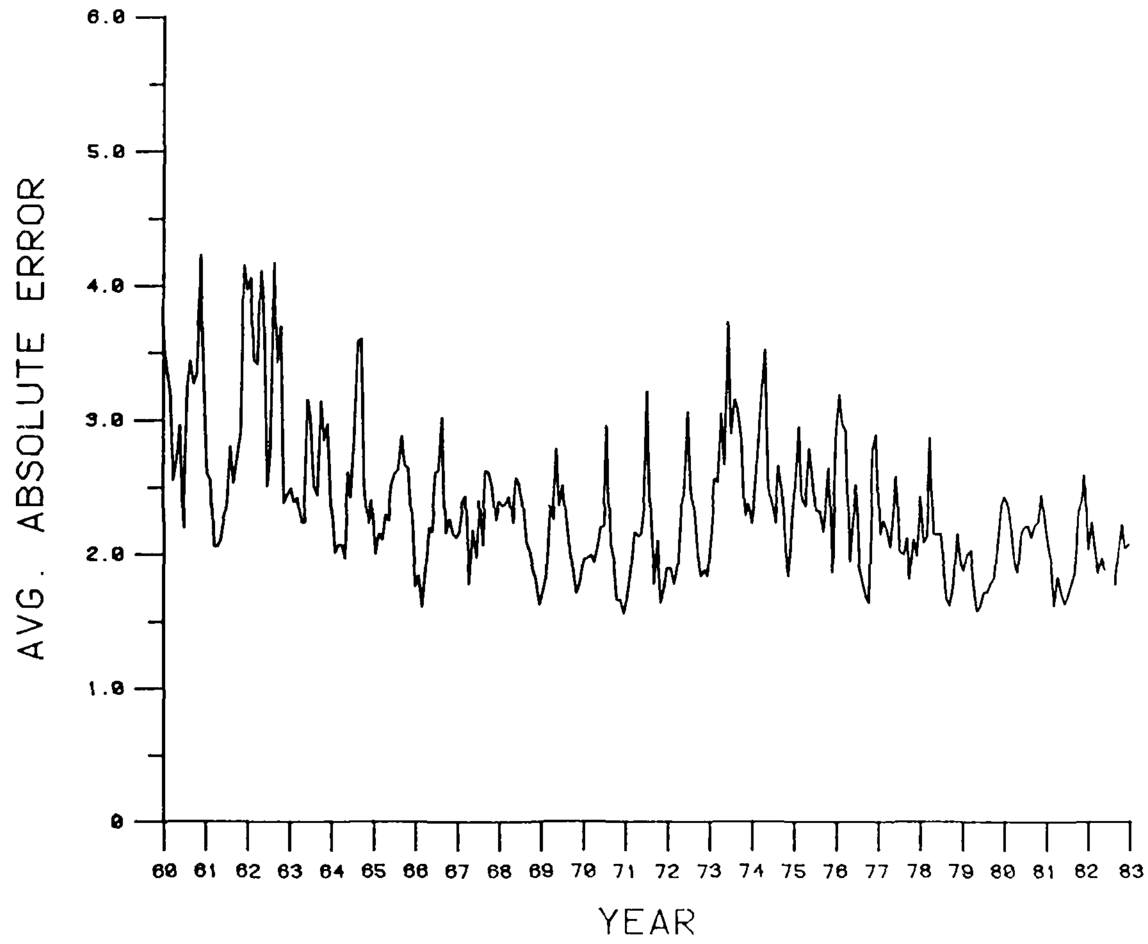


Fig. 38. Model I global average absolute errors (1960-1983) by season and year.

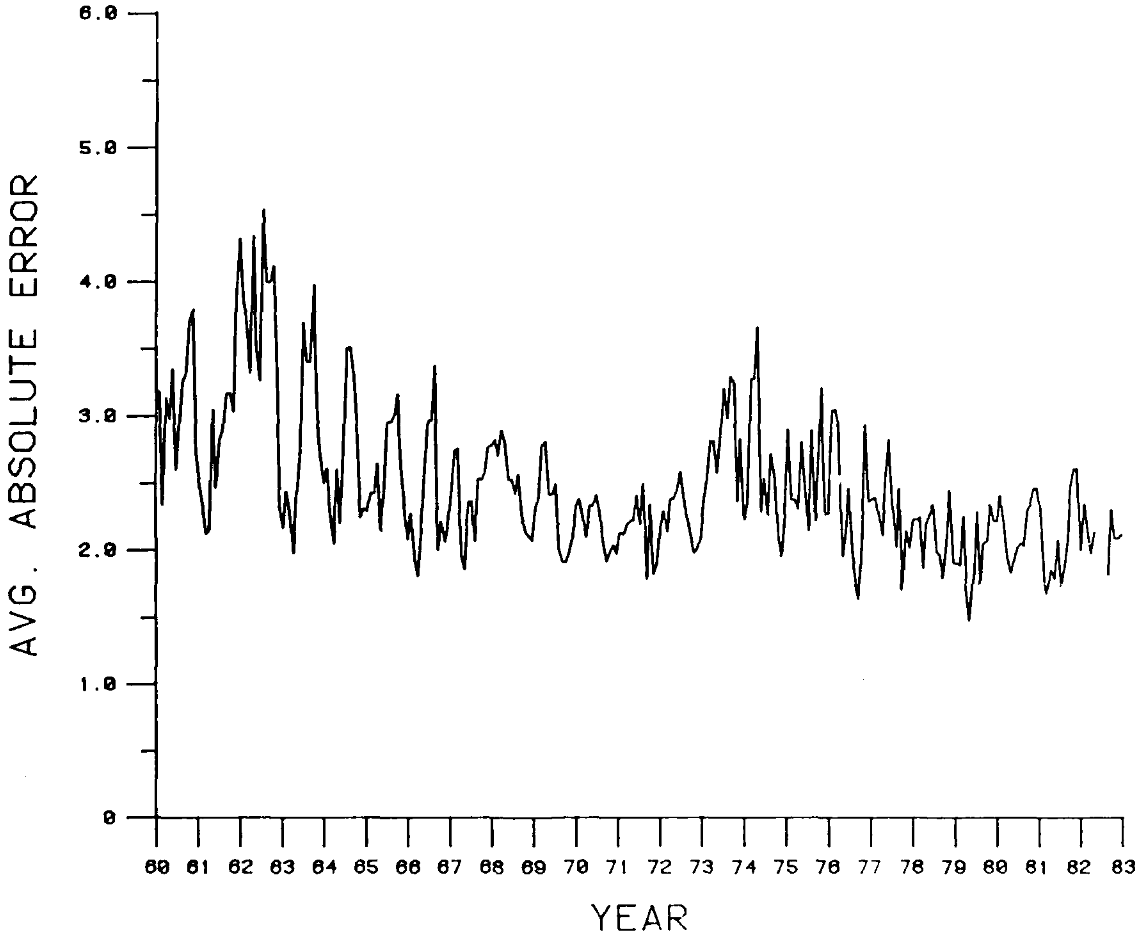


Fig. 39. Model II global average absolute errors (1960-1983) by season and year.

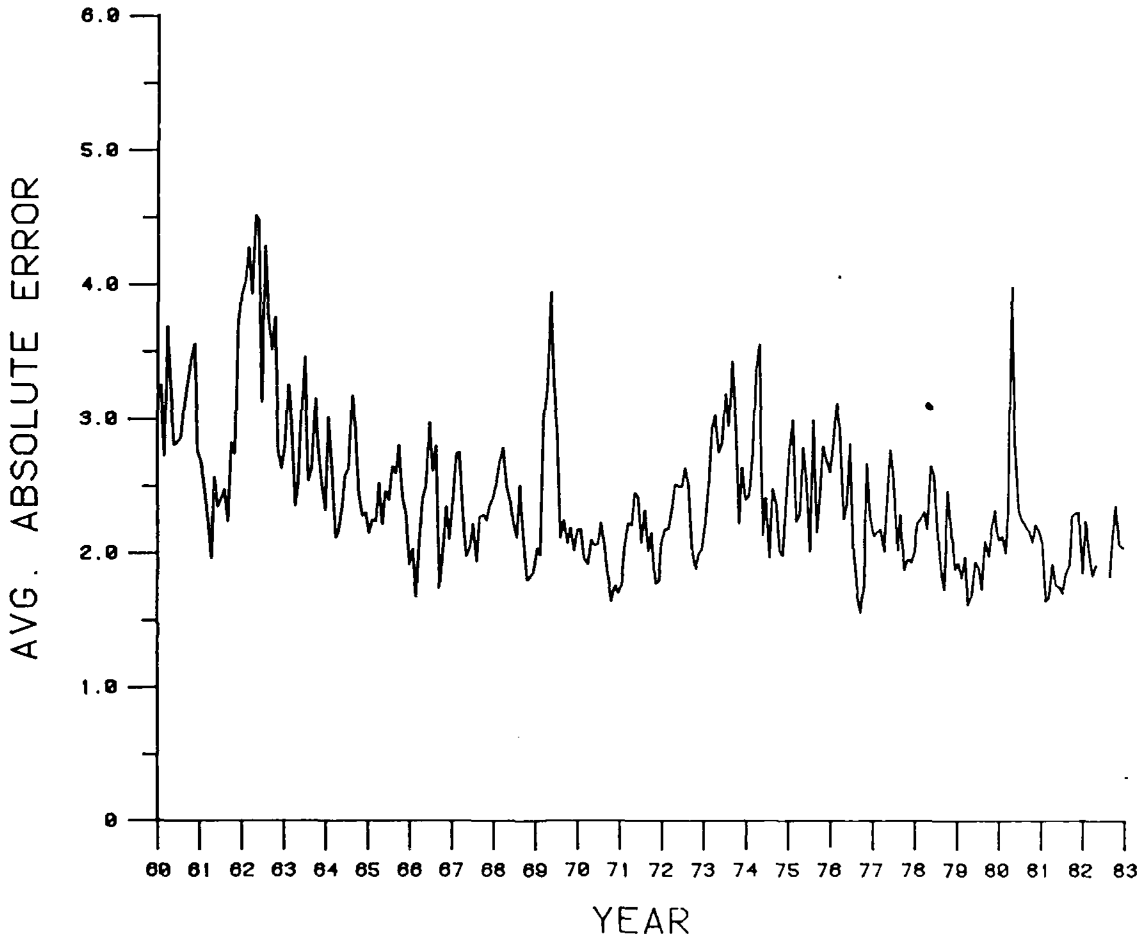


Fig. 40. Model III global average absolute errors (1960-1983) by season and year.

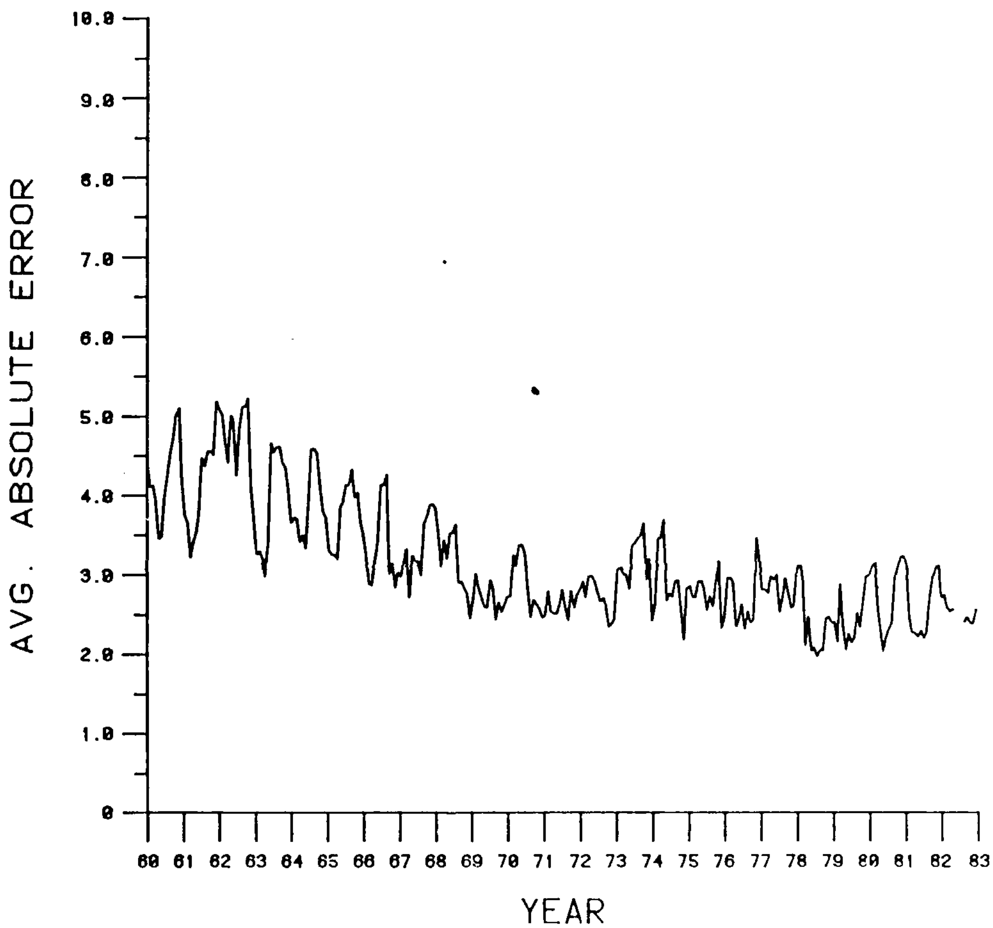


Fig. 41. Global average absolute errors 1(1960-1983) by season and year for the mean as the forecast.

of the models but rather it is due to the annual variation in cyclone numbers. When cyclone numbers are small the size of the possible error is also small, and when large the errors can also be large. Note that the scaling on Figure 41 differs from that used in Figures 38, 39 and 40.

Differences from model to model are few in number. The largest errors occurred in the 1960s but this is because cyclone numbers were larger in those years than in subsequent years. This is also apparent in Figure 41.

Table VII gives the average absolute errors for each model, for each season, and for both the hindcast and operational forecast periods. Average absolute errors as a measure of forecast skill must be viewed from the perspective of the mean as forecast. Accordingly, the percent average absolute error reduction over the mean as a forecast was calculated and is summarized in Table IX. Average absolute errors show a general seasonal cycle with the smallest errors in those forecast seasons which include summer months and the largest forecast error in winter. This cycle results from the occurrence of the annual variation in cyclone frequencies which is high in winter and low in summer and thus higher forecast errors are possible in winter. The three models are little different in terms of average absolute errors and the addition of the operational forecast results to the hindcast period did not result in a lowering of forecast skills.

The percent error reduction over the mean as a forecast (Table IX) averaged 22%. Error reductions for the period that included the operational forecasts improved slightly. While the differences are probably not significant they are not worse as might be expected from the observed reduction in the errors of the mean as a forecast over the 1960-1982 period (Fig. 12). There is no seasonal variation in error reduction and the differences between models are modest except for the January to June forecast and April to September periods where large differences are observed. The 22% error reduction indicates real predictability.

TABLE VIII

Global Average Absolute Errors 1960-1980

SEASON	MODEL I	MODEL II	MODEL III
JANUARY-JUNE	2.74(2.66)	2.60(2.54)	2.48(2.41)
FEBRUARY-JULY	2.49(2.41)	2.50(2.43)	2.47(2.41)
MARCH-AUGUST	2.39(2.32)	2.32(2.27)	2.40(2.42)
APRIL-SEPTEMBER	2.17(2.14)	2.44(2.38)	2.52(2.47)
MAY-OCTOBER	2.12(2.09)	2.33(2.29)	2.29(2.25)
JUNE-NOVEMBER#	2.13(2.11)	2.38(2.35)	2.25(2.22)
JULY-DECEMBER#	2.11(2.10)	2.39(2.35)	2.34(2.31)
AUGUST-JANUARY	2.27(2.23)	2.41(2.37)	2.45(2.39)
SEPTEMBER-FEBRUARY	2.43(2.40)	2.63(2.60)	2.56(2.51)
OCTOBER-MARCH	2.59(2.56)	2.58(2.57)	2.60(2.56)
NOVEMBER-APRIL	2.62(2.58)	2.77(2.72)	2.65(2.59)
DECEMBER-MAY	2.74(2.65)	2.81(1.72)	2.63(2.55)
AVERAGE	2.40(2.35)	2.39(2.51)	2.47(2.42)

( ) = 1960-1983; # 1983 omitted

TABLE IX

Percent Reduction in Global AAE 1960-1980

SEASON	MODEL I	MODEL II	MODEL III
JANUARY-JUNE	16.2(17.1)	22.2(22.5)	26.0(26.4)
FEBRUARY-JULY	22.8(23.0)	22.4(22.4)	23.1(22.9)
MARCH-AUGUST	23.4(23.2)	25.5(25.0)	22.9(20.7)
APRIL-SEPTEMBER	24.5(23.4)	20.2(19.6)	17.7(16.4)
MAY-OCTOBER	20.6(20.2)	21.7(20.5)	22.9(22.9)
JUNE-NOVEMBER#	20.1(19.6)	19.4(19.2)	24.4(23.2)
JULY-DECEMBER#	22.8(22.3)	21.4(21.0)	22.9(22.3)
AUGUST-JANUARY	21.7(22.4)	24.6(24.5)	23.6(24.0)
SEPTEMBER-FEBRUARY	25.0(24.8)	19.9(19.6)	22.1(22.3)
OCTOBER-MARCH	23.2(23.1)	23.3(22.6)	22.8(23.0)
NOVEMBER-APRIL	24.5(24.1)	20.6(20.8)	24.5(24.5)
DECEMBER-MAY	22.2(22.6)	20.2(20.5)	25.2(25.6)
AVERAGE	22.3(22.2)	21.8(21.5)	23.2(22.9)

( ) = 1860-1983; # 1983 omitted

Root Mean Square Error

Global RMSEs by season and year for Models I, II and III are given in Figures 42, 43 and 44. Figure 45 gives the global RMSE of the long-term mean as a forecast by season and year (note scale difference). Figure 45 clearly shows the improvement of the 1885-1960 mean as a forecast in the years 1960-1983. During this period the total number of cyclones declined and the variability also declined. The seasonality of RMSE is also

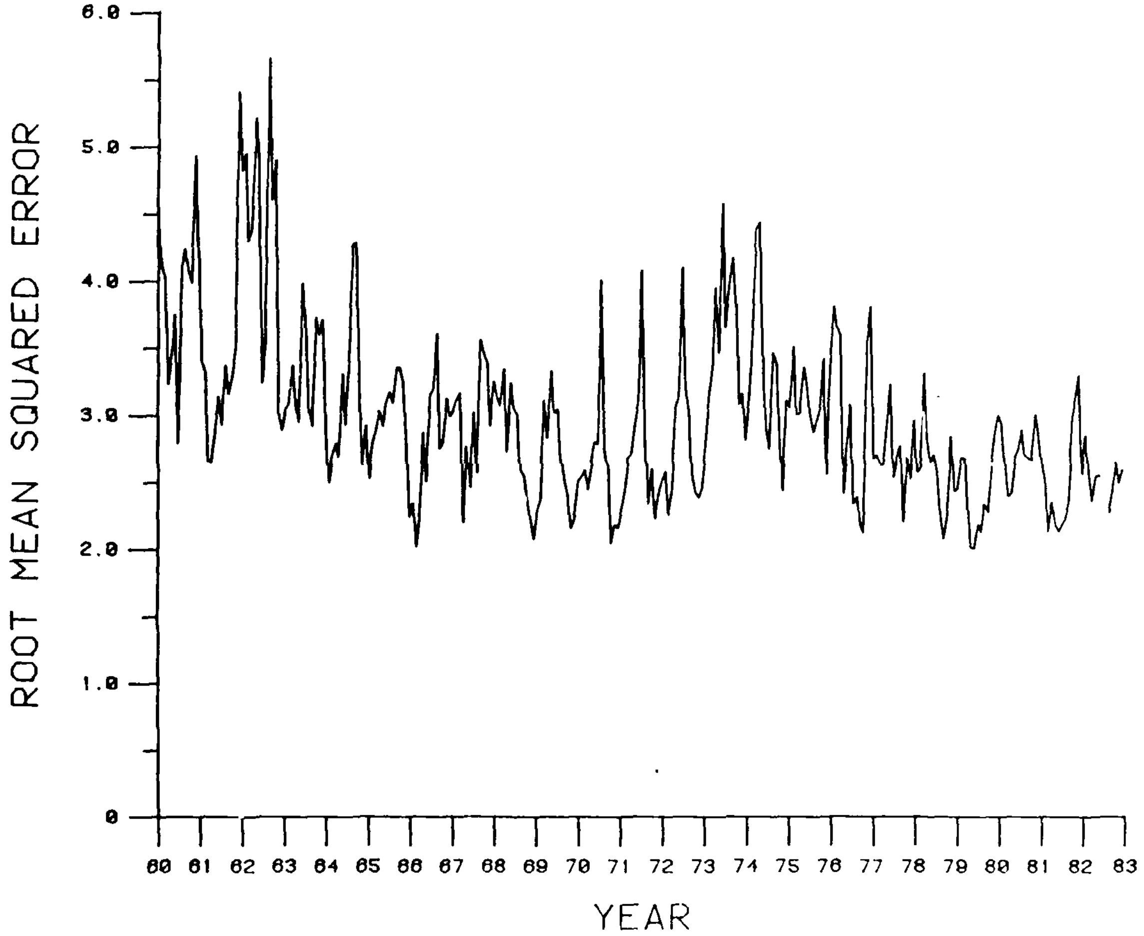


Fig. 42. Model I global RMSE (1960-1983) by season and year.

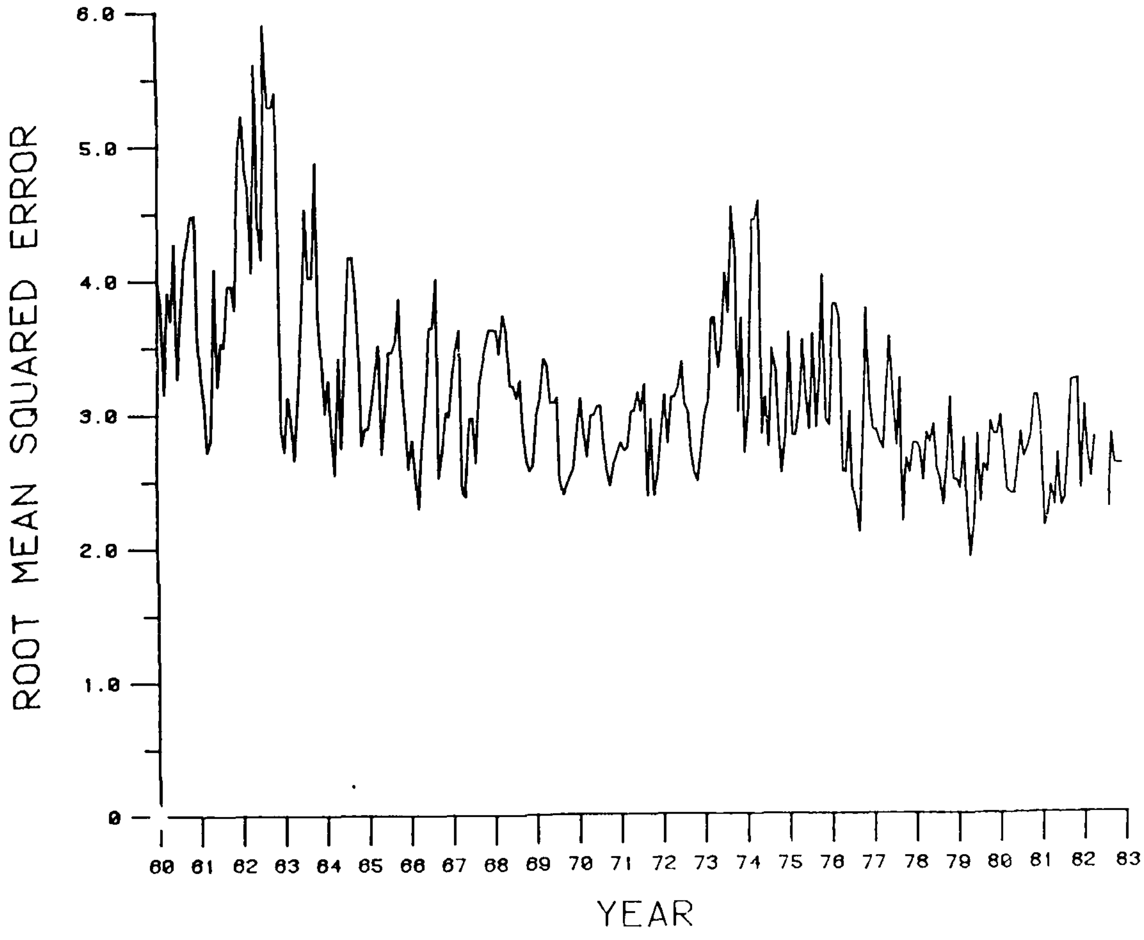


Fig. 43. Model II global RMSE (1960-1983) by season and year.

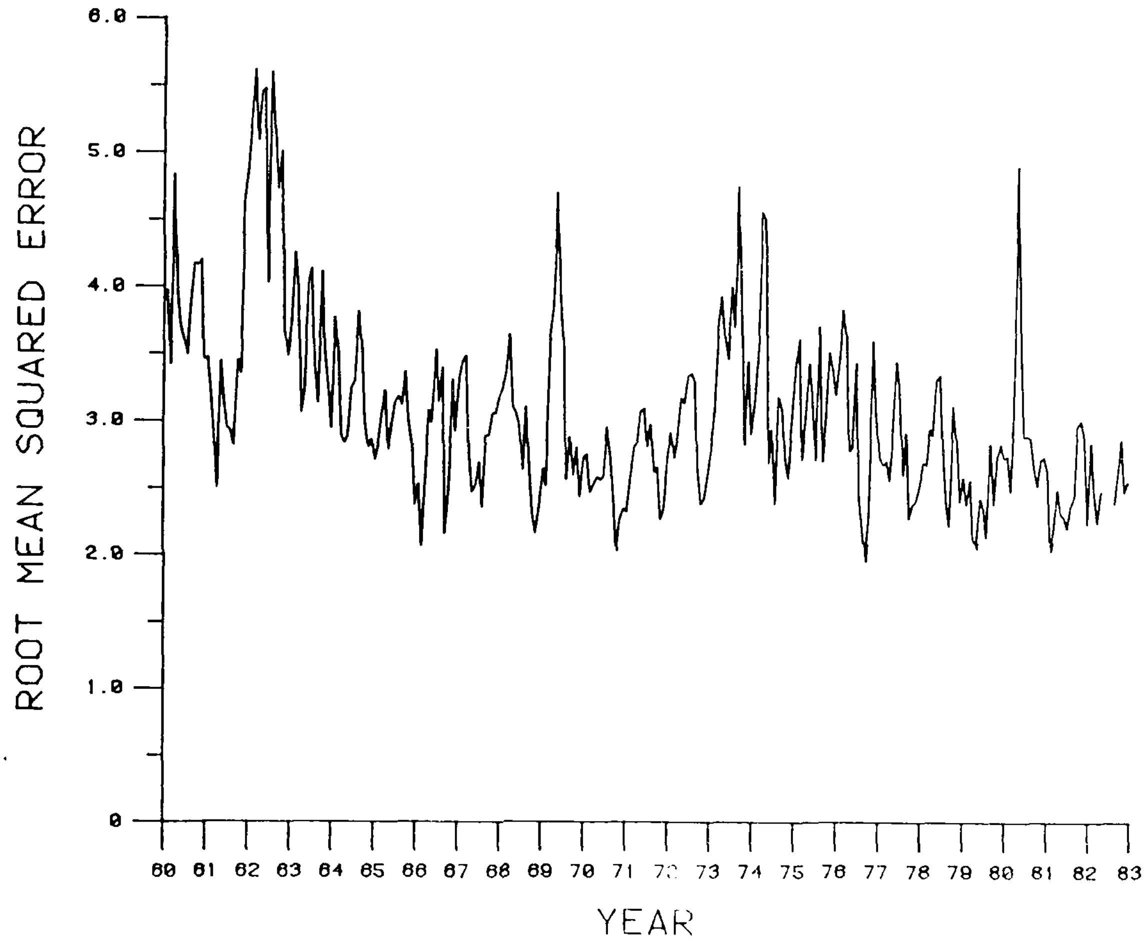


Fig. 44. Model III global RMSE (1960-1983) by season and year.

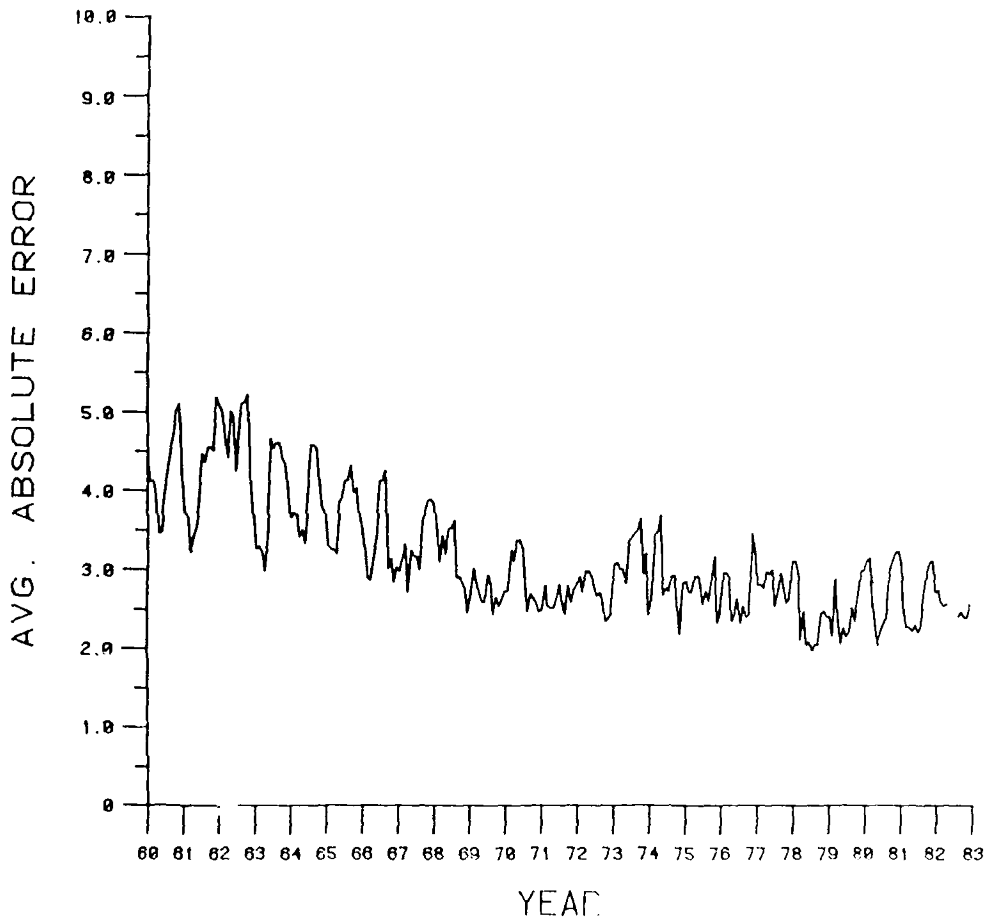
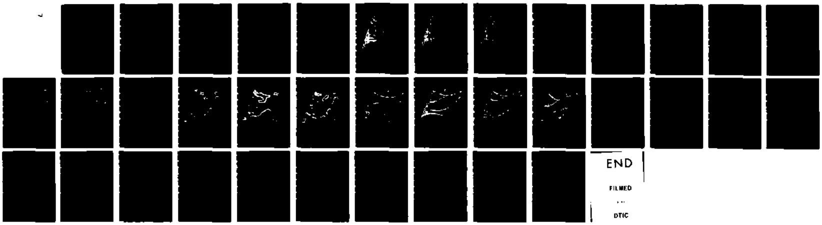
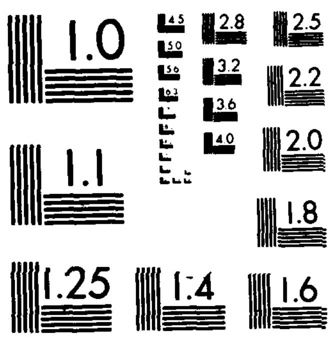


Fig. 45. Global RMSE (1960-1983) by season and year for the mean as the forecast.

AD-A143 576 CLIMATE PREDICTION PART 1 CYCLONE FREQUENCY(U) VIRGINIA UNIV CHARLOTTESVILLE DEPT OF ENVIRONMENTAL SCIENNES B P HAYDEN JUL 84 TR-30 N00014-81-K-0033 2/2

UNCLASSIFIED F/G 4/2 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

evident and reflects the seasonality in total number of cyclones. The higher errors in the 1973-1976 period are due to model failures. Models I, II and III have an average 22% reduction in the error over the mean as a forecast.

Root mean square errors and error reductions over the mean as a forecast are given in Tables X and XI. Because the average error is greater than 1, the root mean square errors are larger than the average absolute errors discussed in the previous section. Like the average absolute errors there is a seasonal cycle in root mean square errors and like the average absolute errors there is no seasonal cycle in error reductions. In addition, there is no degradation of forecast skills when the operational period is added. The average reduction of root mean square errors over the mean as a forecast is 22%. There are few differences in model skill between models or between seasons.

TABLE X

Global RMSE 1960-1980

SEASON	MODEL I	MODEL II	MODEL III
JANUARY-JUNE	3.50(3.39)	3.30(3.21)	3.13(3.04)
FEBRUARY-JULY	3.12(3.02)	3.14(3.06)	3.12(3.05)
MARCH-AUGUST	3.04(2.95)	2.98(2.91)	3.08(3.10)
APRIL-SEPTEMBER	2.76(2.73)	3.10(3.03)	3.23(3.17)
MAY-OCTOBER	2.81(2.77)	2.98(2.94)	2.94(2.88)
JUNE-NOVEMBER#	2.80(2.78)	3.10(3.05)	2.89(2.86)
JULY-DECEMBER#	2.76(2.73)	3.10(3.04)	2.99(2.96)
AUGUST-JANUARY	2.93(2.87)	3.08(3.01)	3.08(3.00)
SEPTEMBER-FEBRUARY	3.03(2.98)	3.30(3.26)	3.19(3.12)
OCTOBER-MARCH	3.23(3.19)	3.25(3.24)	3.28(3.23)
NOVEMBER-APRIL	3.28(3.23	3.48(3.42)	3.13(3.24)
DECEMBER-MAY	3.39(3.29)	3.51(3.40)	3.30(3.19)
AVERAGE	3.05(2.99)	3.19(3.13)	3.13(3.07)

( ) = 1960-1983; # 1983 omitted

TABLE XI

Percent Reduction in Global RMSE 1960-1980

SEASON	MODEL I	MODEL II	MODEL III
JANUARY-JUNE	17.3(17.8)	22.5(21.3)	25.3(25.5)
FEBRUARY-JULY	23.0(23.0)	22.4(22.1)	22.9(22.3)
MARCH-AUGUST	23.9(23.6)	25.2(24.5)	22.8(20.8)
APRIL-SEPTEMBER	24.6(23.2)	20.5(20.0)	17.3(16.2)
MAY-OCTOBER	21.1(20.8)	21.6(10.6)	22.9(22.1)
JUNE-NOVEMBER#	21.6(21.4)	19.1(19.0)	24.6(24.1)
JULY-DECEMBER#	24.7(24.3)	21.1(20.9)	23.6(23.0)
AUGUST-JANUARY	23.0(23.2)	24.4(24.0)	24.3(24.3)
SEPTEMBER-FEBRUARY	25.2(25.0)	19.3(18.8)	22.2(22.2)
OCTOBER-MARCH	22.9(22.6)	22.3(21.5)	21.7(21.7)
NOVEMBER-APRIL	24.0(23.5)	20.3(19.9)	24.1(24.2)
DECEMBER-MAY	22.7(23.2)	20.0(20.1)	24.9(25.1)
AVERAGE	22.8(22.6)	21.6(21.1)	23.1(22.6)

( ) = 1960-1983; # 1983 omitted

OPERATIONAL FORECASTS

Models I, II and III were used in operational trials beginning in January 1981. Three years of trials have now been completed. Two forecasts were inadvertently not verified as of this writing (June-November and July-December 1983). A total of 34 forecasts were made with each model version. Eighty-seven grid cell locations were forecast. In all 2958 forecasts were made using each model. This is a sufficiently large sample such that the global scores from this period can be reasonably compared with those of the hindcast period (1960-1980). In earlier sections of this report data from the operational period were merged with the hindcast period and so some comparisons have already been made. In this section a specific assessment of the performance of the models in real time forecasting is presented.

Average Local Skill

Local skill scores are usually averaged only over time, however, in this case only three forecasts were made at each grid cell for each season. This sample is too small to be meaningful so we have averaged across all seasons. The sample size in each grid cell is now 34 and a reasonable estimate of local skill in the operational period can be made.

Figures 46, 47 and 48 show the season averaged local skills for Models I, II and III. The regions of high skill and regions of low skill during the operational period are essentially the same as found for the hindcast period (Figs. 14-25). Perfect forecasts (34 correct in 34 trials) were made for 10 grid cells in Model I, 6 in Model II and 8 in Model III. The locations of these perfect forecasts were like those that occurred in the hindcast period. The local skills differed little between Models I, II and III. We conclude that the models are stable in a spatial sense relative to the hindcast period and because the skills high we assume also that the stability extends back into the dependent data period (1885-1980).

Global Skill

Global skill scores by model and season are reported in Tables XII, XIII and XIV. Percent correct, Heidke, deviation, and quadratic skill scores are given as are the average absolute errors, root mean square errors, and their error reductions over the errors of the long term means as forecasts.

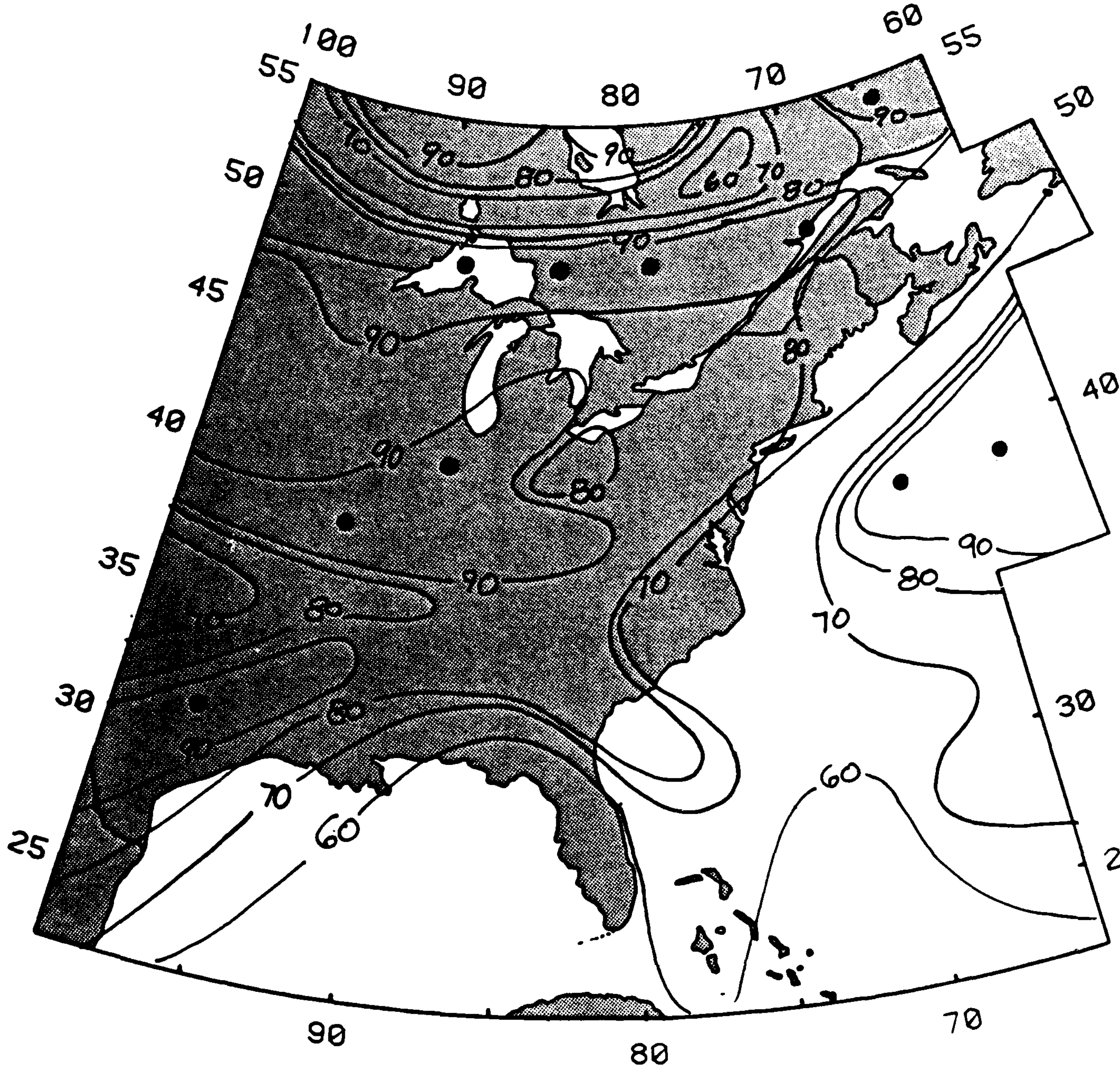


Fig. 46. Model I local skill scores averaged across all seasons for the operational forecast period. The units are percent correct in 34 forecasts. Solid black circles indicate grid cells where 34 correct forecasts were made in 34 trials.

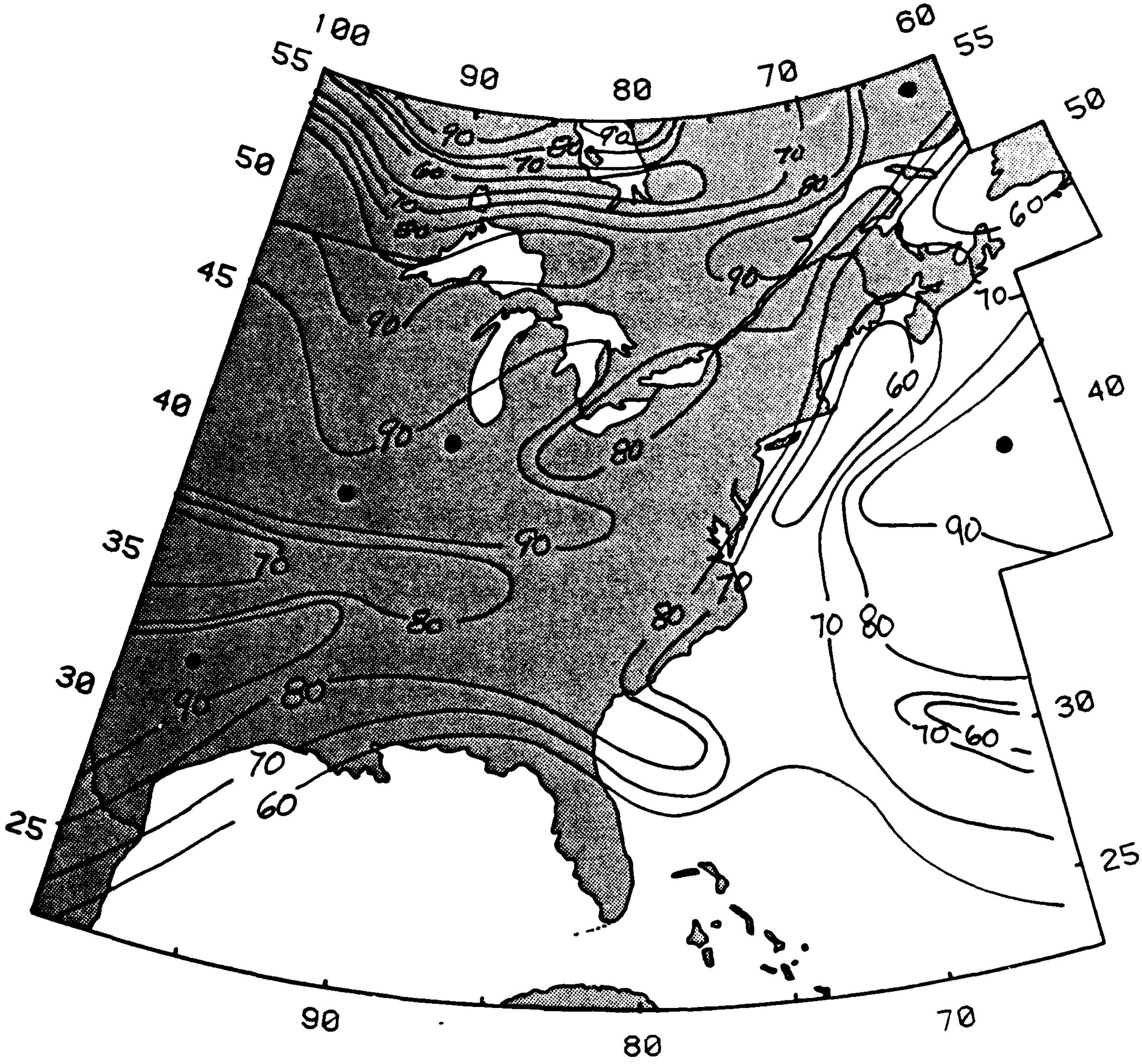


Fig. 47. Model II local skill scores averaged across all seasons for the operational forecast period. The units are percent correct in 34 forecasts. Solid black circles indicate grid cells where 34 correct forecasts were made in 34 trials.

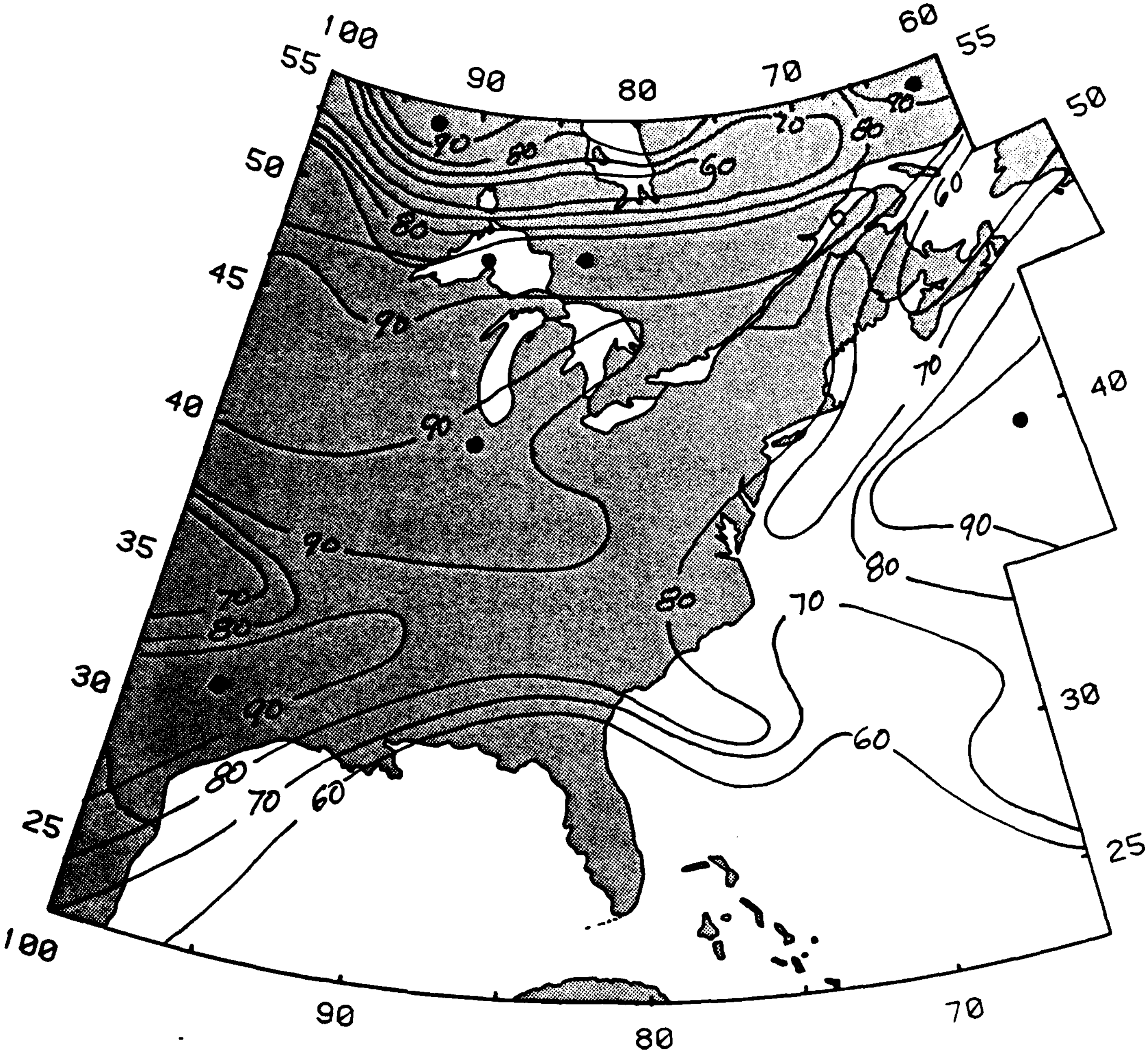


Fig. 48. Model III local skill scores averaged across all seasons for the operational forecast period. The units are percent correct in 34 forecasts. Solid black circles indicate grid cells where 34 correct forecasts were made in 34 trials.

Global percent correct scores averaged across all seasons for all three models in operational forecasts (76.5%, 75.8% and 76.8%) out-performed the models in the hindcast period (73.6%, 74.6% and 75.9%). Heidke skill scores followed suit. Deviation and quadratic skill scores were slightly lower in the operational trials compared to those of the hindcast period. AAE and RMSE were smaller during the operational period than in the hindcast period but the error reductions were also smaller. This circumstance results from the fact that there has been a decline in the size of the observed cyclone frequency departures from the long term means (see Figs. 11 and 12).

Overall there was no degradation of the models when applied on a real-time forecasting basis. This is extremely encouraging as it weighs well regarding reliability of the models tested.

Table XII

Model I Skill Scores by Season for the Operational Period

	%	H	D	Q	AAE(%)	RMSE(%)
JAN-JUNE	78.2	.563	.237	.413	2.11(24.1)	2.64(21.3)
FEB-JULY	74.7	.493	.253	.440	1.87(25.0)	2.36(23.1)
MAR-AUG	75.5	.507	.213	.380	1.85(21.6)	2.37(20.8)
APR-SEPT	75.1	.503	.130	.130	1.95(14.0)	2.47(16.1)
MAY-OCT	72.0	.440	.167	.297	1.90(17.3)	2.47(18.5)
JUNE-NOV*	73.6	.470	.140	.255	1.95(13.9)	2.54(18.2)
JULY-DEC*	73.6	.475	.165	.300	1.95(16.5)	2.47(19.6)
AUG-JAN	78.2	.567	.277	.477	1.95(27.7)	2.44(25.1)
SEPT-FEB	80.1	.600	.230	.407	2.19(23.3)	2.68(22.9)
OCT-MAR	79.7	.593	.213	.373	2.35(22.3)	2.93(19.9)
NOV-APR	76.6	.530	.203	.363	2.30(21.0)	2.88(19.3)
DEC-MAY	80.1	.603	.257	.447	2.05(26.2)	2.59(22.7)
AVERAGE	76.5	.529	.207	.357	2.04(21.1)	2.57(20.6)

* only 1981 and 1982 included

TABLE XIII

MODEL II Skill Scores by Season for the Operational Period

	%	H	D	Q	AAE(%)	RMSE(%)
JAN-JUNE	80.5	.609	.246	.427	2.09(24.8)	2.62(22.2)
FEB-JULY	74.7	.494	.228	.402	1.93(22.6)	2.44(20.4)
MAR-AUG	73.2	.463	.201	.361	1.88(20.2)	2.44(18.4)
APR-SEPT	74.3	.486	.137	.250	1.95(14.1)	2.49(15.5)
MAY-OCT	64.8	.295	.093	.178	2.06(9.3)	2.66(9.6)
JUNE-NOV*	75.3	.506	.165	.301	1.89(16.4)	2.56(17.3)
JULY-DEC*	76.1	.521	.164	.300	1.95(16.3)	2.50(18.3)
AUG-JAN	77.4	.548	.238	.419	2.06(23.8)	2.57(21.1)
SEPT-FEB	78.9	.578	.163	.294	2.37(17.1)	2.96(14.9)
OCT-MAR	78.9	.578	.160	.285	2.50(17.3)	3.12(14.7)
NOV-APR	77.4	.548	.175	.318	2.83(18.0)	2.99(16.5)
DEC-MAY	78.2	.563	.231	.406	2.14(23.2)	2.64(21.4)
AVERAGE	75.8	.516	.184	.328	1.94(18.9)	2.67(17.5)

* only 1981 and 1982 included

TABLE XIV

MODEL III Skill Scores by Season for the Operational Period

	%	H	D	Q	AAE(%)	RMSE(%)
JAN-JUNE	81.6	.632	.296	.498	1.95(29.8)	2.44(27.4)
FEB-JULY	77.0	.540	.206	.366	1.99(20.3)	2.54(17.3)
MAR-AUG	71.2	.425	.054	.051	2.58(2.6)	3.21(4.1)
APR-SEPT	73.6	.471	.024	-.040	2.16(4.8)	2.77(6.0)
MAY-OCT	69.0	.379	.147	.255	1.93(14.9)	2.47(16.2)
JUNE-NOV*	73.6	.471	.134	.255	1.96(13.1)	2.53(18.2)
JULY-DEC*	73.1	.462	.130	.240	2.03(12.8)	2.06114.8)
AUG-JAN	79.7	.594	.272	.470	1.92(27.4)	2.48(23.9)
SEPT-FEB	80.8	.617	.235	.408	2.16(24.3)	2.69(40.0)
OCT-MAR	80.1	.601	.228	.389	2.28(24.3)	2.86(21.9)
NOV-APR	80.1	.601	.243	.420	2.17(25.3)	2.69(24.6)
DEC-MAY	81.2	.624	.286	.486	1.98(28.9)	2.46(26.6)
AVERAGE	76.8	.536	.188	.316	2.09(19.0)	2.60(20.1)

* only 1981 and 1982 included

FORECAST COMPARISONS

In this section a forecast made during the period of operational forecasting using all three of the models is examined in detail. The July to December 1982 season was selected for this comparison. This season was selected because it was a forecast that was as successful in about the same measure as the average forecast made in forecast trials. The purpose is to study the similarities and differences between the three model versions for an individual forecast. Figures 49, 50 and 51 show the forecasted anomalies for the July-December 1982 season predicted by Models I, II and III. It is clear that all three models give essentially the same forecast. As noted elsewhere when all three models predict essentially the same forecast a bust is unlikely. This was a successful forecast. While similarities are great there are differences between the three forecasted anomaly fields. The range of forecasted anomalies was seven cyclones per grid cell in Model I, five cyclones per grid cell in Model II and nine cyclones per grid cell in Model III. The axes of maximum and minimum values in the forecasted anomaly fields are quite similar except that Model II indicates the Atlantic coast track as having its origin in the vicinity of New Orleans while Models I and III show the track starting in the central part of

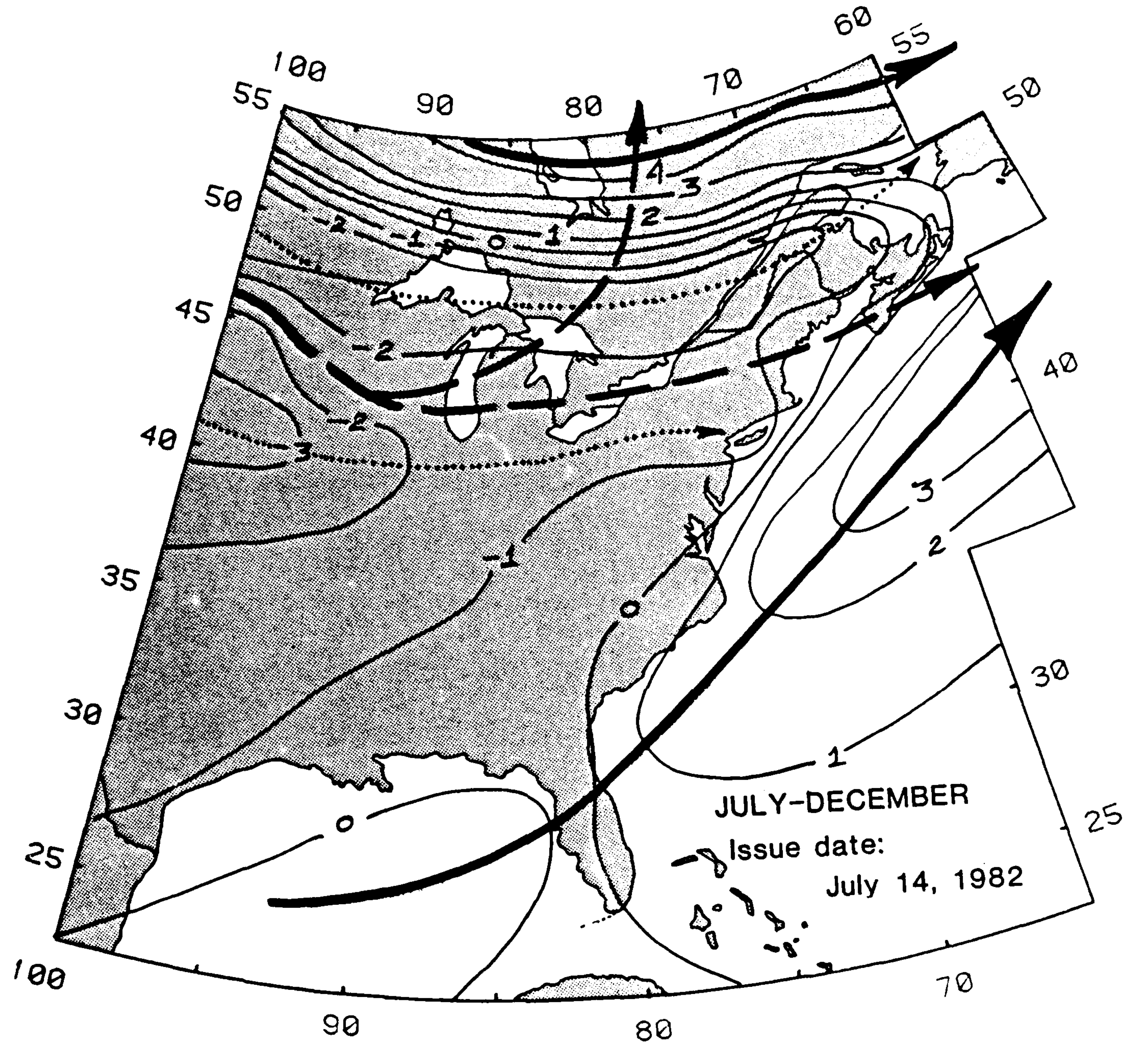


Fig. 49. Model I predicted cyclone frequency anomalies for the July-December 1982 season. Solid arrows indicate axes of maximum positive anomaly. Dotted arrows indicate axes of maximum negative anomaly. Dashed arrows indicate local maxima in a region of negative anomalies.

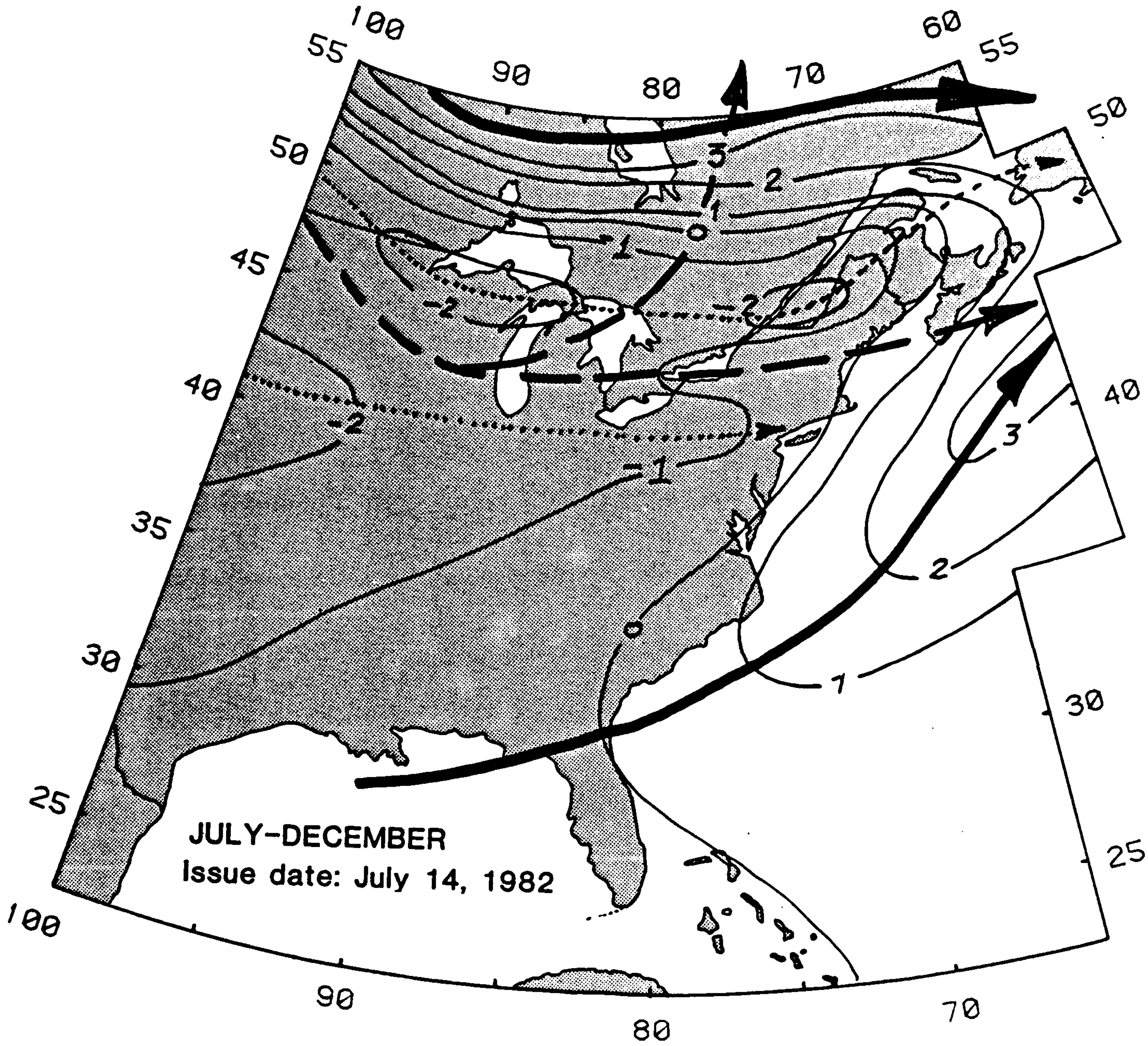


Fig. 50. Model II predicted cyclone frequency anomalies for the July-December 1982 season. Solid arrows indicate axes of maximum positive anomaly. Dotted arrows indicate axes of maximum negative anomaly. Dashed arrows indicate local maxima in a region of negative anomalies.

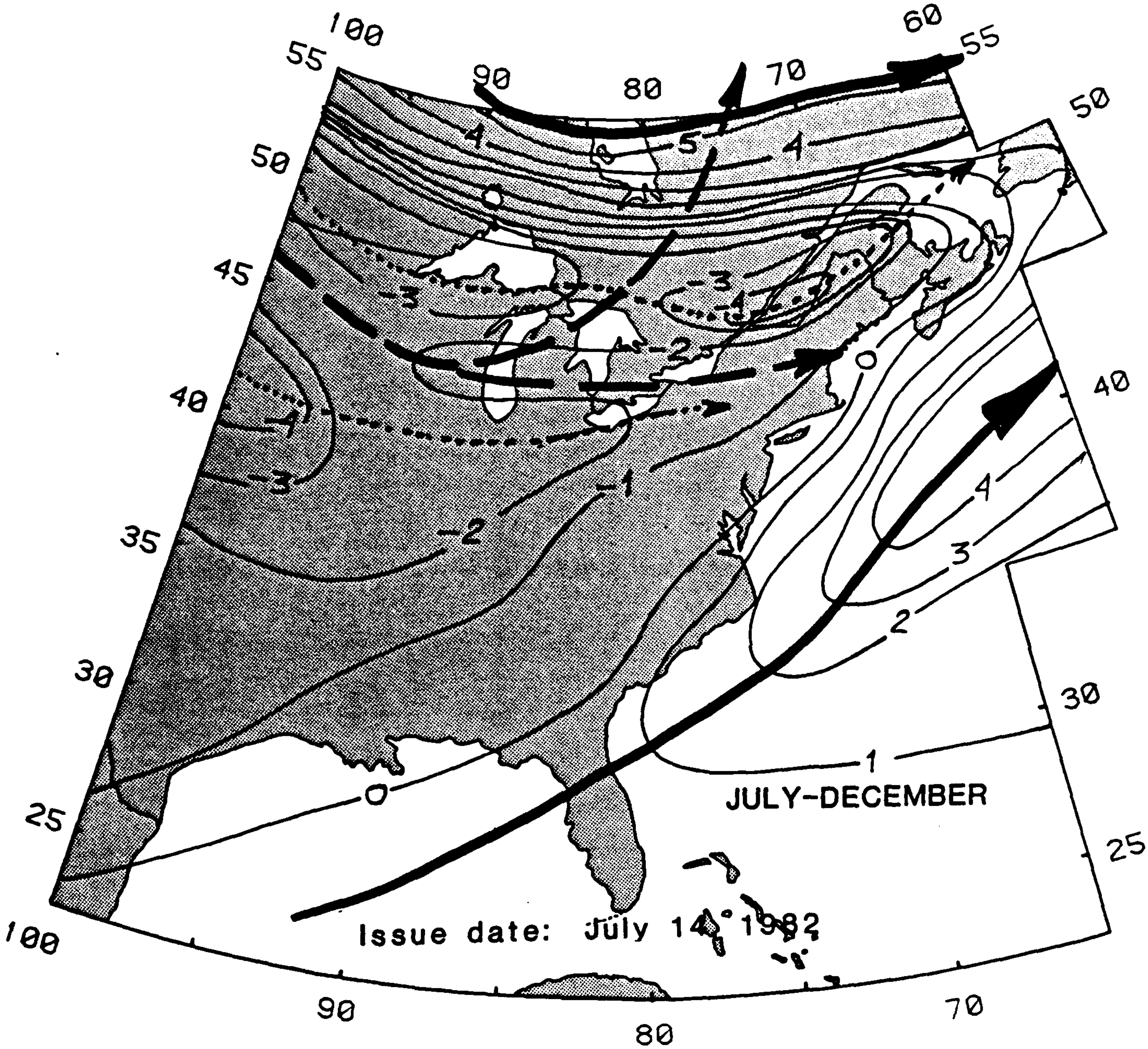


Fig. 51. Model III predicted cyclone frequency anomalies for the July-December 1982 season. Solid arrows indicate axes of maximum positive anomaly. Dotted arrows indicate axes of maximum negative anomaly. Dashed arrows indicate local maxima in a region of negative anomalies.

the Gulf of Mexico. Because most of the prediction errors tend to occur where the forecasted anomalies are between +1 and -1 cyclones per grid cell, there is value in examining which model version has the smallest area between +1 and -1 cyclones per grid cell. Model III is the best in this regard. This relationship between forecast skill and forecasted anomaly magnitude can be verified by examining charts of skill scores for each of the three models (Fig. 52, 53 and 54).

Figures 55, 56, and 57 show the forecasted cyclone frequencies for the July-December period, i.e., the frequency anomalies plus the long term mean frequencies. The arrows indicate the "ridge lines" of maximum forecasted cyclone frequencies. The major differences between Models I, II and III regarding the forecasted tracks are found in the southeastern U.S. Analyses of Model II forecasted frequencies indicated a double track across the Gulf states with both tracks further north than the single tracks indicated in Models I and III. The results of analyses of the actual occurring cyclones in July-December 1982 are shown in Figure 57. The double track indicated by Model II is evident in the observations. While the field of observed cyclone frequencies is more complex than the forecasted fields, most of the features of the forecasted fields are evident in the observations. Global percent skill was 77.0% for Model I, 75.9% for Model II and 75.9% for Model III. While Model II did well in predicting the tracks across the south, the overall skill for

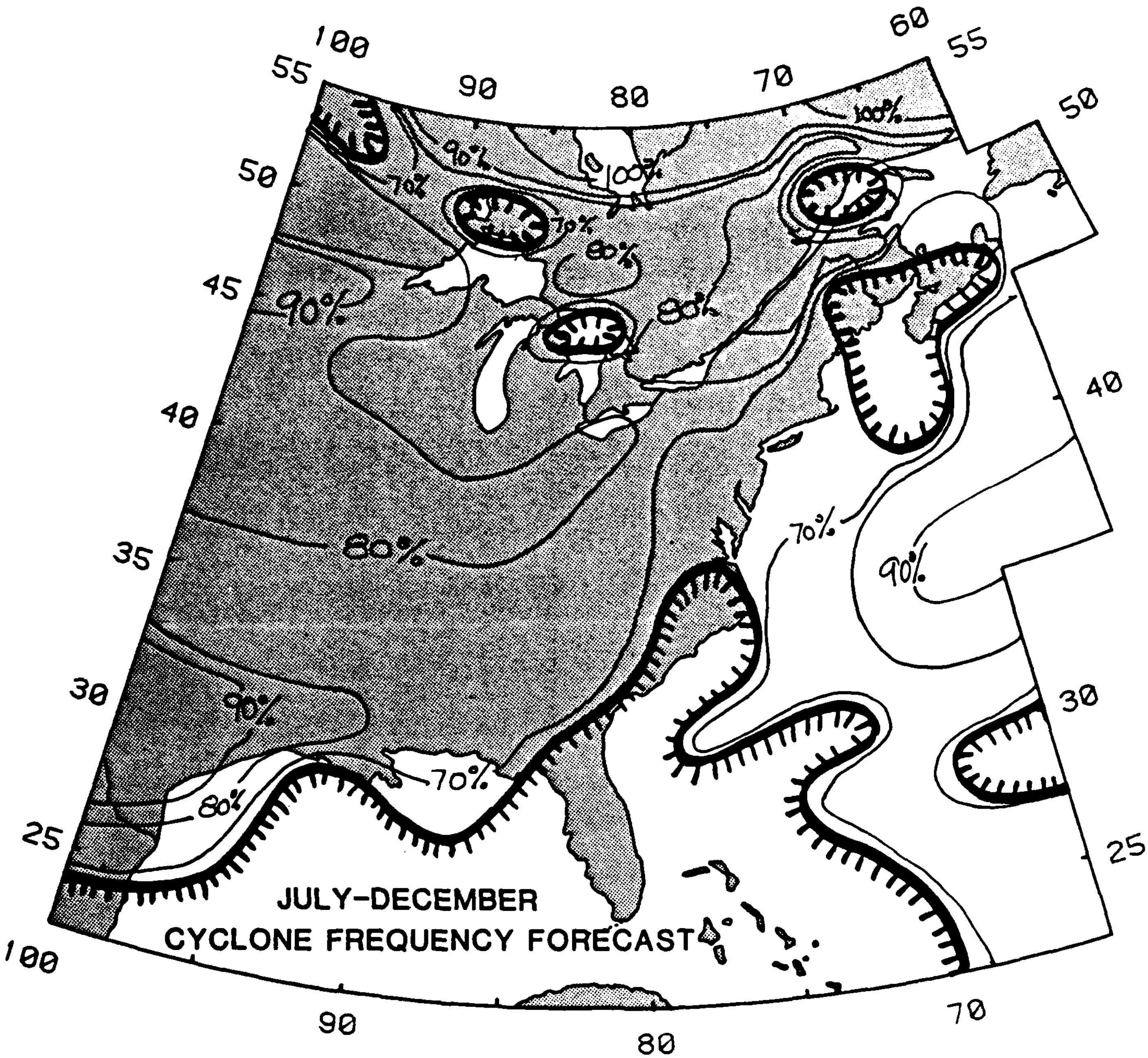


Fig. 52. Model I local percent correct skill scores (1960-1980) for the July-December forecast season. Heavy contours indicate skill scores equal to or less than 67% correct. Grid cells with 100% scores are not shown.

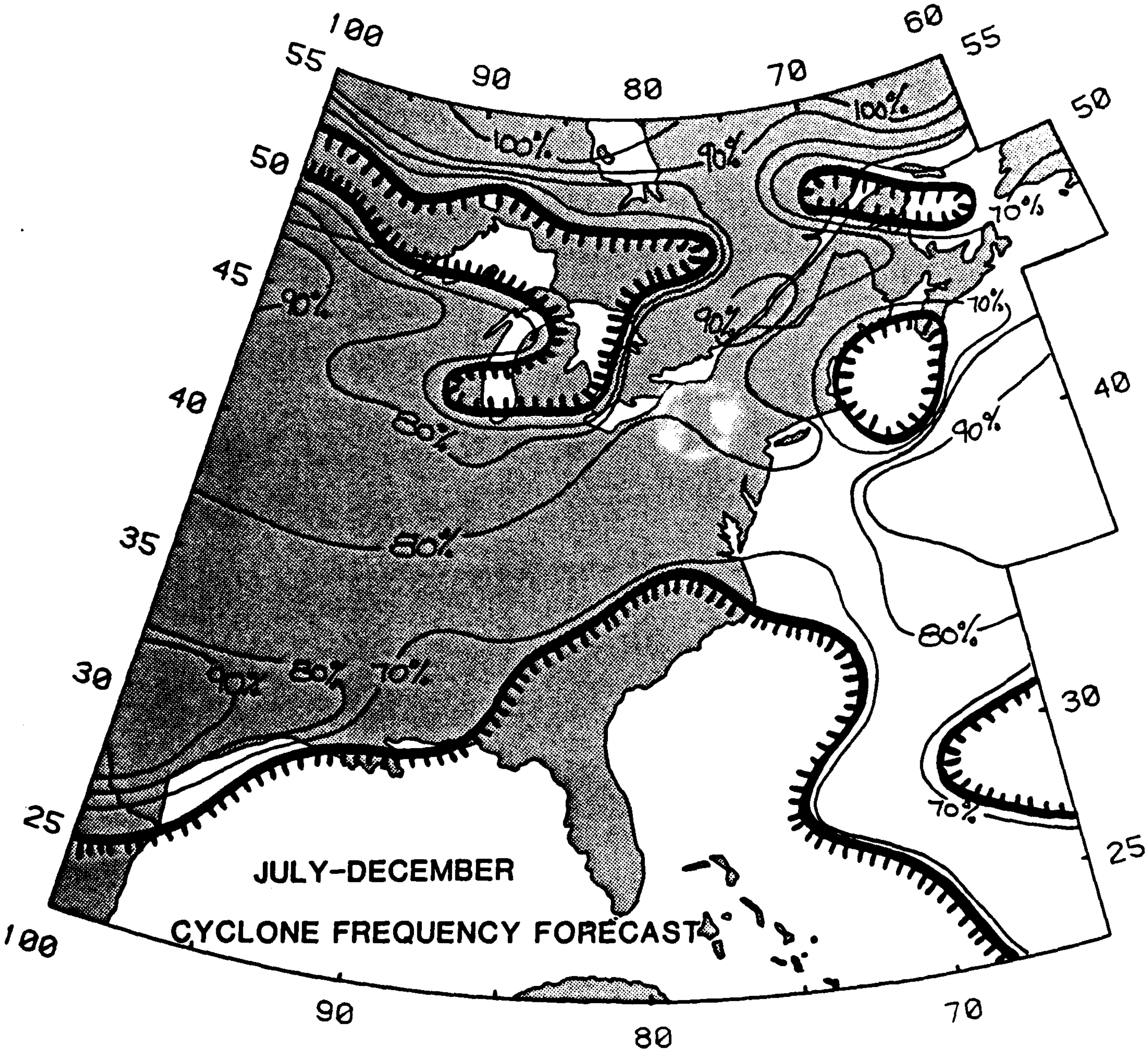


Fig. 53. Model II local percent correct skill scores (1960-1980) for the July-December forecast season. Heavy contours indicate skill scores equal to or less than 67% correct. Grid cells with 100% scores are not shown.

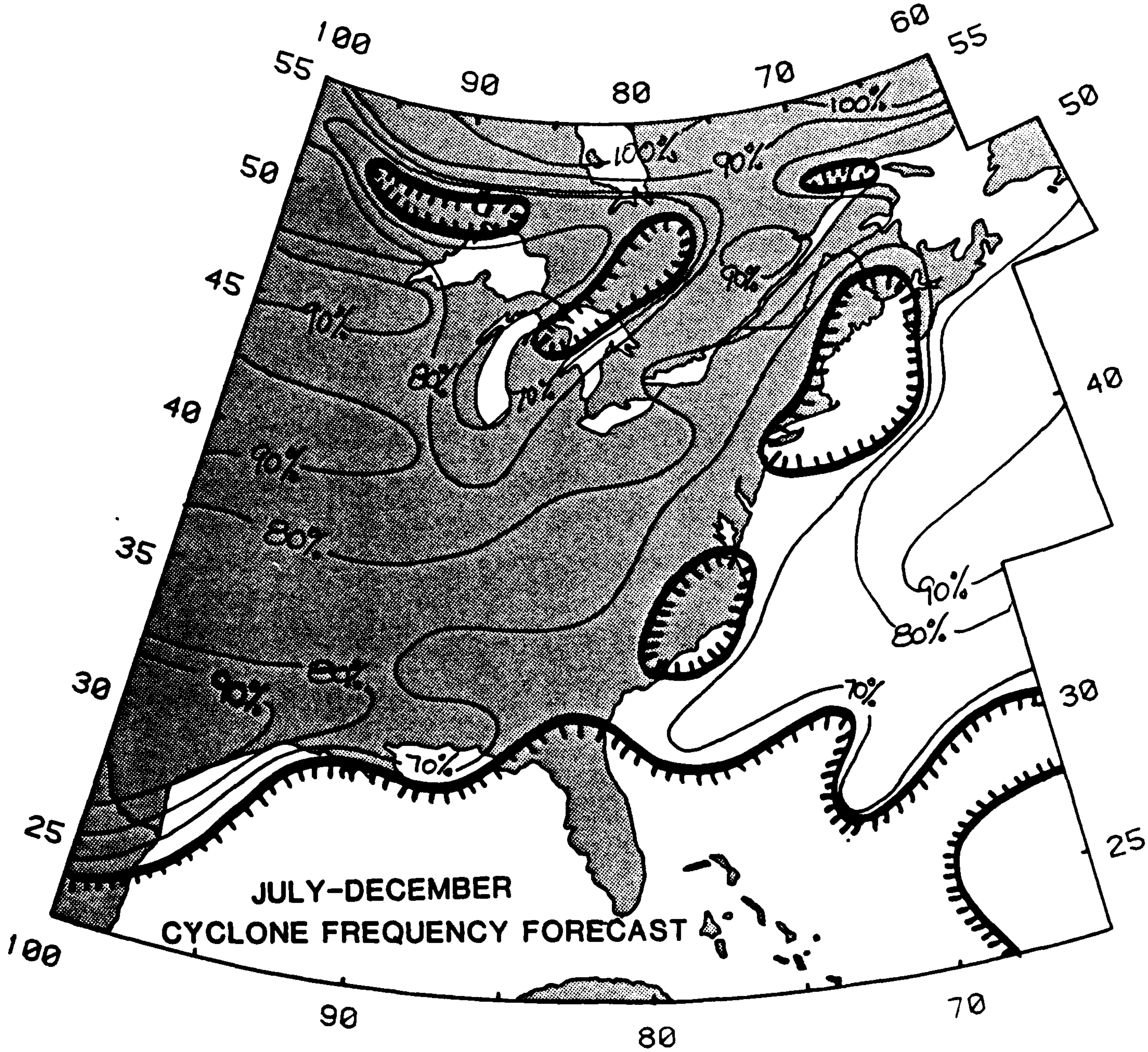


Fig. 54. Model III local percent correct skill scores (1960-1980) for the July-December forecast season. Heavy contours indicate skill scores equal to or less than 67% correct. Grid cells with 100% scores are not shown.

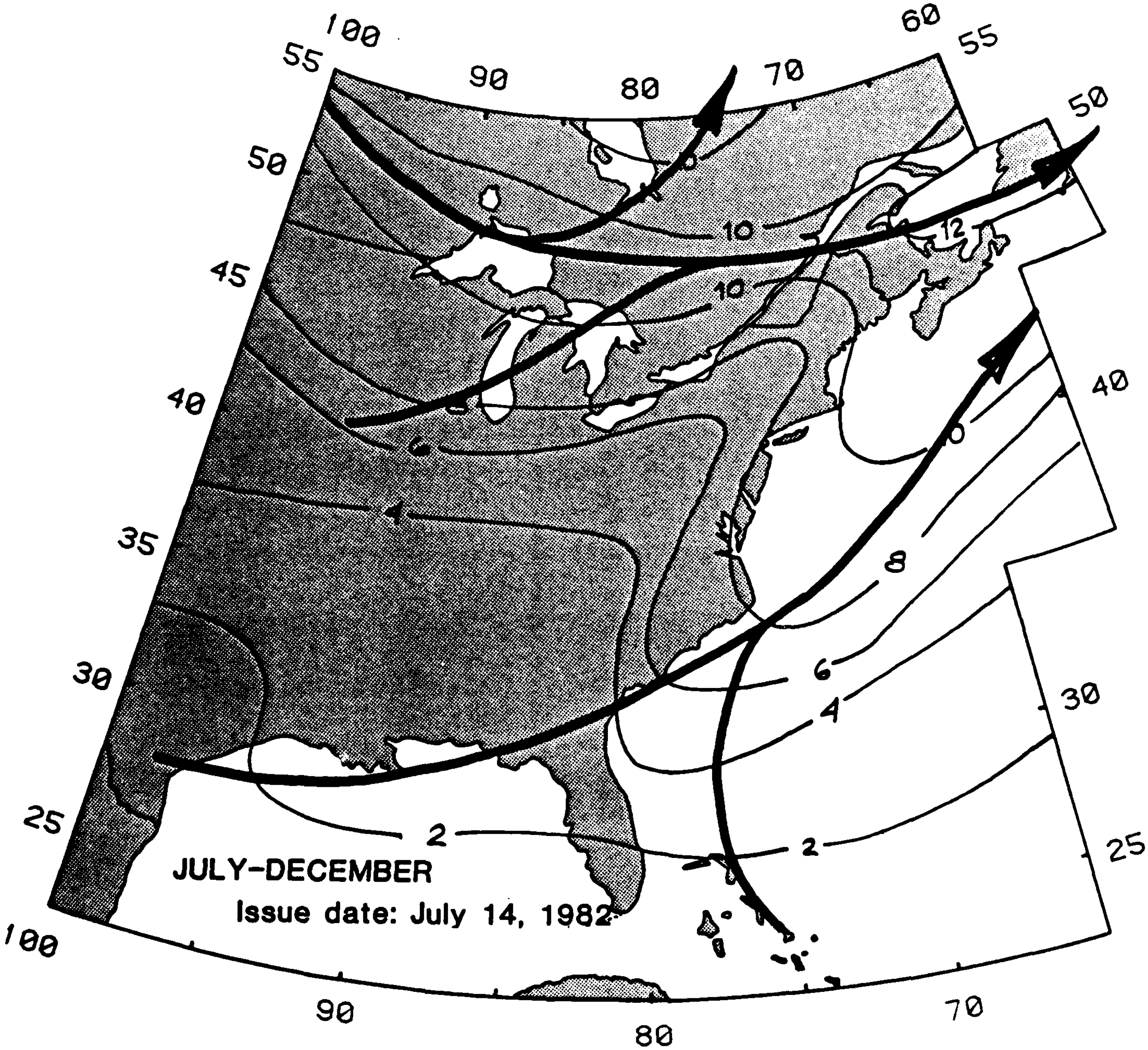


Fig. 55. Model I predicted cyclone frequencies for the July-December 1982 season. Arrows indicate the axes of maximum predicted cyclones per grid cell.

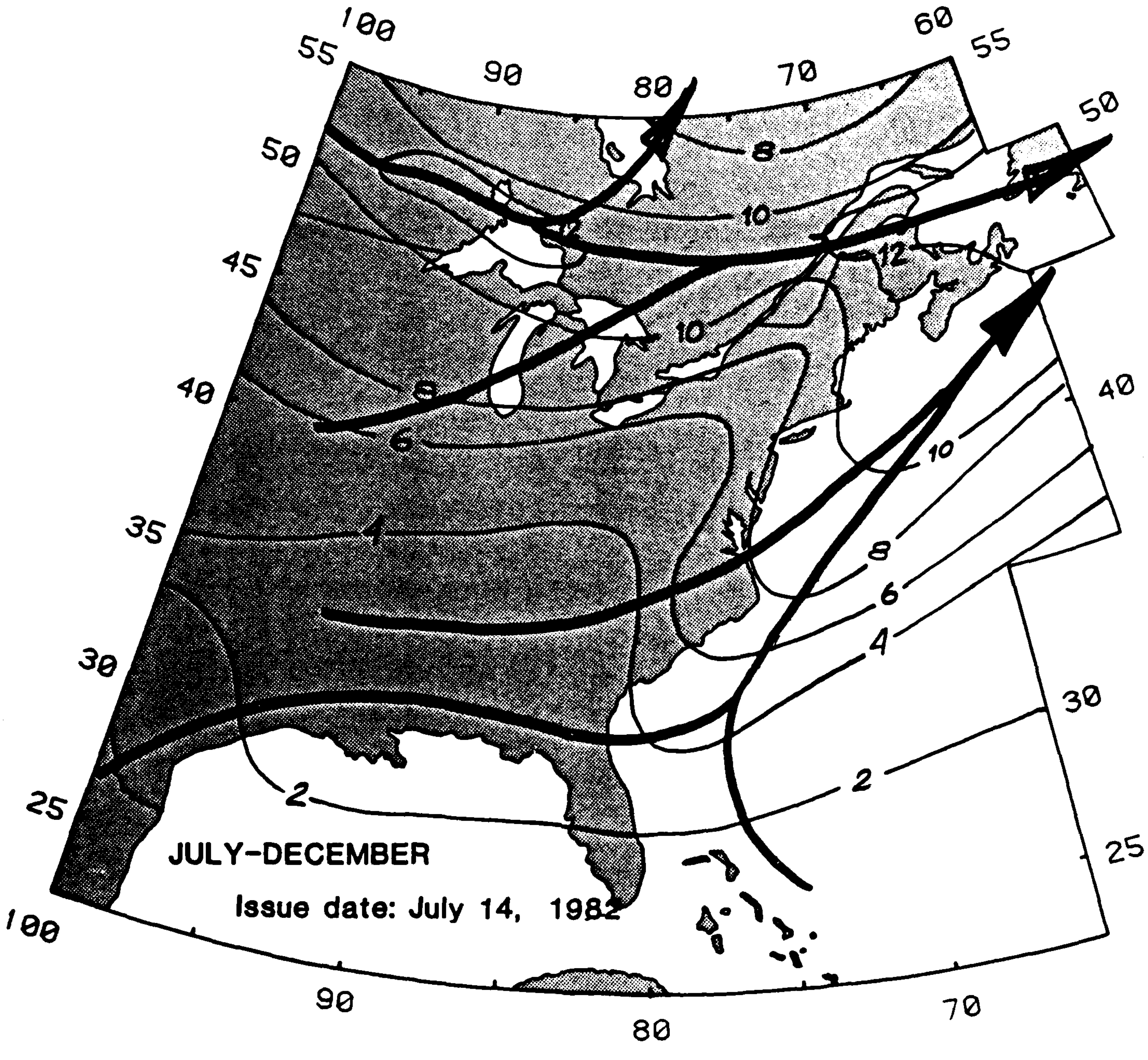


Fig. 56. Model II predicted cyclone frequencies for the July-December 1982 season. Arrows indicate the axes of maximum predicted cyclones per grid cell.

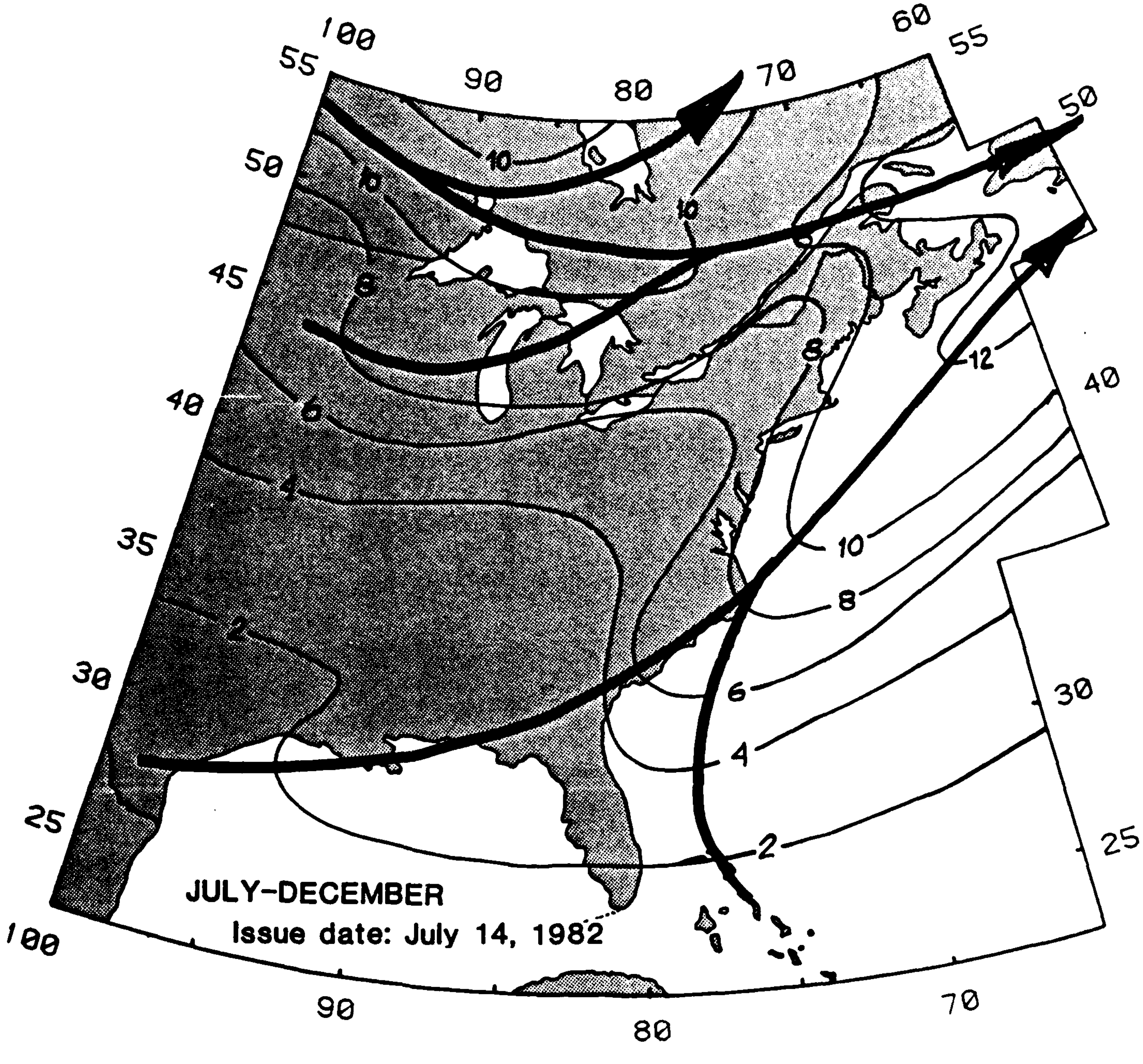


Fig. 57. Model III predicted cyclone frequencies for the July-December 1982 season. Arrows indicate the axes of maximum predicted cyclones per grid cell.

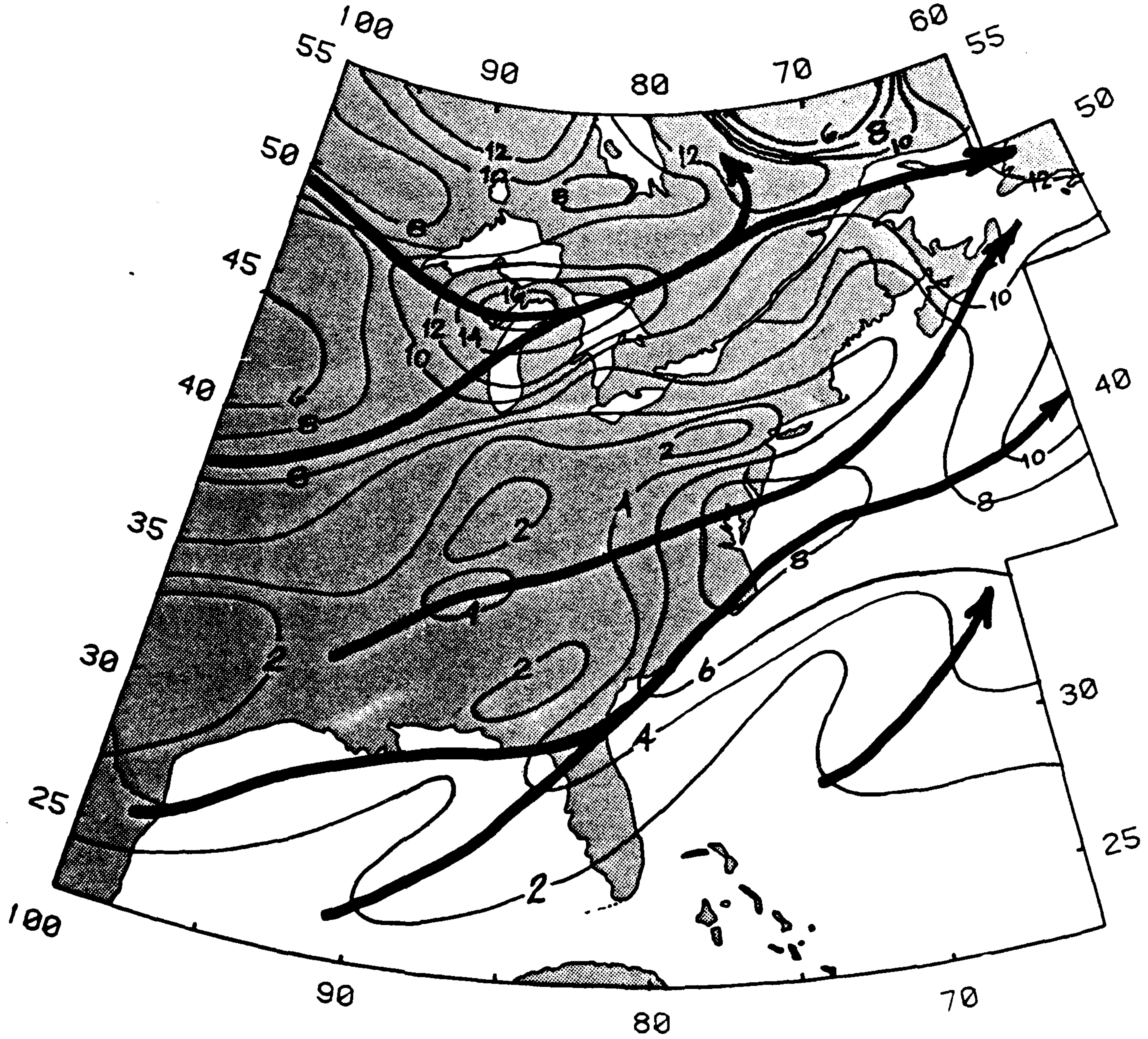


Fig. 58. Observed cyclone frequencies for the July-December 1982 season. Arrows indicate axes of maximum frequencies. The units are cyclones per grid cell.

Model III was not different from the skills for Models I and III. In general, we find that global skill rarely differs between models except when there is a persistence and forecast failure. There are frequently differences in the details of the forecast and there are differences in local skill between models. The three models are rarely contradictory and when they are the forecast that is fundamentally different is usually the forecast that fails.

CONCLUSIONS

Climate Predictability

Over the last two decades the predictability of climate has become a fundamental topic of research and a topic about which there exists fundamental differences among scientists. This circumstance prompted Lorenz (1973) to note that the predictability of climate will be established when someone shows that it can be done. Much of the recent work on climate predictability focuses on the partitioning of signal and noise in historical data. The spatial and temporal variations in the signal-to-noise ratio thus serves as a proxy of the attribute of predictability. Much of the work to date focuses on temperature, pressure and precipitation. Based on signal-to-noise ratios for monthly temperatures a general rule of thumb has emerged: climate predictability is highest along the coastal margins of the continents and decreases toward the interior of the continents. Based on our work we conclude that this rule of thumb does not apply to the prediction of cyclone frequencies. A different pattern of predictability emerges. We would then conclude that predictability will vary from parameter to parameter and according to season duration.

Given Lorenz's rather pragmatic approach to the question of predictability we conclude that such demonstration of predictability has been realized for a climatic parameter of fundamental synoptic significance. As such, new avenues are now open to a new approach to the prediction problem.

Categorical Forecast Skill

Most attempts to forecast climate take a categorical approach. Forecasts of above or below the long-term means are forwarded. On occasion terciles or quartiles are predicted. Both categorical and numerical forecasts have been prepared and evaluated in this study. Based on the results of the categorical 2-by-2 tests of forecast skill we place the level of forecast skill for each of the three models developed and tested at about 75%. This is a global skill that covers an 87-location forecast domain and a period of forecast trials on independent data that spans 25 years. This skill level meets the requirements of statistical significance ($p = 0.05$) at an individual location let alone as the average for 87 locations. The categorical skills achieved could not have occurred by chance alone. Cyclone frequencies relative to the long-term means are predictable quantities.

Forecast skill is high in baroclinic and low in barotropic areas. Also skill is generally low along the coastal margins and along the northern shores of the Great Lakes. Both of these areas are axes of maximum frequencies in the long-term means but are not axes of maximum standard deviations about the means. Magnitude modulation of the mean pattern is not predictable by the methods used in this study.

Categorical forecast skills are uniform from season to season and show no trends in levels over the period of forecast trials. When the mean for the period 1885-1960 is used to predict the conditions in the years that followed it turns out that the mean has become progressively better as a forecast. This is due to the general decline in variability in cyclone frequencies over the last two decades. A similar decline in forecast skill for the models is not observed even though the average departure (mean minus observed) has become smaller. The sign of these smaller anomalies remains as predictable as at the beginning of the test period when cyclone numbers were higher.

Numerical Forecast Skill

Numerical forecasts were made and evaluated for skill. The skill was measured using a penalty proportional to the size of error (deviation skill score) and also using a squaring of the penalty (quadratic skill score). Positive skill is found in 95% of the forecasts made. Since 286 forecasts were made (12 seasons times 25 years less 2 missing seasons) it is highly unlikely that this result is due to chance.

Numerical forecast skill was found to be linearly related to 2-by-2 categorical forecast skill. It is clear that models exhibit both categorical and numerical skill. It is interesting to note that numerical skill goes to zero as the categorical skill falls below 60%. This then may be a bottom level of skill for climate prediction models, i.e., when numerical skill cannot be demonstrated. In our work we have applied a considerably higher standard.

Forecast Failures

Forecast failures, i.e., categorical skill below 50% or numerical skill below 0 occurred only about 5% of the time. Poor forecast skill (60 to 65%) occurred and persisted for a few years in the mid-1970s. We conclude that the variability during this period was not contained within the statistical base used to construct the models. Earlier studies using jackknifed trials for the entire 95 year period revealed no other period with a comparable persistent period of failures. The type of statistical models employed cannot predict patterns not included in the training base. The three years beginning about February 1973 then become a special case that merits additional study.

The duration of forecast failure is interesting. Here a "forecast failure event" is defined as a 10% skill score fall and a 10% skill score rise (e.g. see Fig. 26). Of the 48 "events" 25 had a one forecast duration; 12 a two forecast duration; 8 a three forecast duration; and, 3 a four forecast duration. We infer this to indicate that when cyclone frequency climate changes and persistence fails that the model fails but recovers to correctly forecast the changed climate on the next or following forecast. While models are not instantaneously responsive to changes in cyclone tracks and numbers the response is less than 1/3 of the duration of the period forecast.

Forecast Models

Three versions of the forecast models were constructed and tested. They differed in regards to the attribute of rotation of principal component axes. The three models performed in a global sense essentially the same. There were slight differences in skill from place to place and from season to season. In general, the forecast failures found in one model were not the same as those found in the other models. Forecast successes were common among models. We conclude that running all three models is a positive utility and may provide a means of detecting poor forecasts at the time of issue.

Hindcast vs Operational Forecasts

Three years of operational forecasting have been completed. The results of these operational trials are indistinguishable from those made on independent data in a hindcast mode. We conclude that the prediction models are stable.

BIBLIOGRAPHY

Bergeron, T., 1930: Richtlinien einer dynamischen Klimatologie. Met. Zeit. 47, 246-262.

Brier, G. and Allen, R., 1951: Verification of Weather Forecasts. In Compendium of Meteorology, p. 841-848, Thomas F. Malone, Ed., American Meteorological Society, Boston, MA, Waverly Press, Baltimore, MD.

Hayden, B.P., 1981a: Season-to-season Cyclone Frequency Prediction. Technical Report 22, Office of Naval Research, Contract No. N00014-81-K-0033, Department of Environmental Sciences, University of Virginia, Charlottesville, VA, 34 p.

_____, 1981b: Secular Variation in Atlantic Coast Extratropical Cyclones. Mon. Wea. Rev. 109(1):159-167.

_____, 1981c: Cyclone Occurrence Mapping: Equal Area or Raw Frequencies. Mon. Wea. Rev. 109(1):168-172.

_____, and Smith, W., 1982: Season-to-Season Cyclone Frequency Prediction. Mon. Wea. Rev. 110(4):239-253.

Livezey, R.E.and Chen, W.Y., 1983: Statistical Field Significance and its Determination by Monte Carlo Techniques. Mon. Wea. Rev. 111(1):46-59.

Lorenz, E.N., 1973: On the Existence of Extended Range Predictability. J. Appl. Meteor. 12:543-546.

Madden, R.A. and Shea, D.J., 1978: Estimate of the Natural Variability of Time-Averaged Temperatures over the United States. Mon. Wea. Rev. 106:1695-1703.

Richman, M.B., 1981: Obliquely Rotated Principal Components: An Improved Meteorological Map Typing Technique? J. Appl. Meteor. 20:1145-1159.

_____, 1983a: Rotation of Principal Components in Climatological Research. Part I: Theoretical Considerations, Suitable Application and Advantages over Unrotated Solutions. In Proceedings of the 8th Conf. on Probability and Statistics in Atmos. Sciences, Chap. 3, p. 59-68. American Meteorological Society, Boston, MA.

_____, 1983b: Rotation of Principal Components in Climatological Research. Part II: How the Various Analytic Simple Structure Rotations on the Major Statistical Packages Compare on Different Types of Data. In Proceedings of the 8th Conf. on Probability and Statistics in Atmos. Sciences, Chap. 5, p. 115-124. American Meteorological Society, Boston, MA.

Vernon, E.M., 1953: A New Concept of Skill Scores for Rating Quantitative Forecasts. Mon. Wea. Rev. 81:326-329.

Whittaker, L.M. and Horn, L.H., 1981: Geographical and Seasonal Distribution of North Amercan Cyclogenesis 1958-1977. Mon. Wea. Rev., 109(11):2312-2322.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report 30
2. GOVT ACCESSION NO.: AD-A143 576
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): Climate Prediction Part I: Cyclone Frequency
5. TYPE OF REPORT & PERIOD COVERED: Technical Report
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Bruce P. Hayden
8. CONTRACT OR GRANT NUMBER(s): N00014-81-K-0033
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia 22903
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR389-170
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Coastal Sciences Program, Arlington, Virginia 22217
12. REPORT DATE: July 1984
13. NUMBER OF PAGES: 115
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report):
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release, distribution unlimited
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): climate prediction, forecast, cyclones, climate, storm climate principal components
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

Climate prediction models based on multivariate analyses of cyclone frequencies are constructed from historical data (1885-1960) and evaluated for forecast skill on independent data (1960-1983). Cyclone frequencies are predicted for six-month duration seasons at 87 locations over North America and the western North Atlantic from 27.5° to 55°. Three types of principal components are constructed and tested. Model I uses unrotated principal component axes, Model II uses rigid rotation of the component axes, and Model III uses oblique rotations of the component axes.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

OCT. 26, 1977 JOB 460-101

SYSTEMATIC GEOGRAPHY LISTING

```
0001  1  OFFICE OF NAVAL RESEARCH          2 COPIES
0001  2  GEOGRAPHY PROGRAMS
0001  3  CODE 462
0001  4  ARLINGTON, VIRGINIA 22217

0002  1  DEFENSE DOCUMENTATION CENTER  12 COPIES
0002  2  CAMERON STATION
0002  3  ALEXANDRIA, VIRGINIA 22314

0003  1  DIRECTOR, NAVAL RESEARCH LAB    6 COPIES
0003  2  ATTENTION TECHNICAL INFORMATION OFFICER
0003  3  WASHINGTON, D. C. 20375

0004  1  DIRECTOR
0004  2  OFFICE OF NAVAL RESEARCH BRANCH OFFICE
0004  3  1030 EAST GREEN STREET
0004  4  PASADENA, CALIFORNIA 91101

0005  1  DIRECTOR
0005  2  OFFICE OF NAVAL RESEARCH BRANCH OFFICE
0005  3  536 SOUTH CLARK STREET
0005  4  CHICAGO, ILLINOIS 60605

0006  1  DIRECTOR
0006  2  OFFICE OF NAVAL RESEARCH BRANCH OFFICE
0006  3  495 SUMMER STREET
0006  4  BOSTON, MASSACHUSETTS 02210

0007  1  COMMANDING OFFICER
0007  2  OFFICE OF NAVAL RESEARCH BRANCH OFFICE
0007  3  BOX 39
0007  4  FPO NEW YORK 09510

0008  1  CHIEF OF NAVAL RESEARCH
0008  2  ASST. FOR MARINE CORPS MATTERS
0008  3  CODE 100M
0008  4  OFFICE OF NAVAL RESEARCH
0008  5  ARLINGTON, VIRGINIA 22217

0009  1  OFFICE OF NAVAL RESEARCH
0009  2  CODE 480
0009  3  NATIONAL SPACE TECHNOLOGY LABORATORIES
0009  4  BAY ST. LOUIS, MISSISSIPPI 39520

0010  1  OFFICE OF NAVAL RESEARCH
0010  2  OPERATIONAL APPLICATIONS DIVISION
0010  3  CODE 200
0010  4  ARLINGTON, VIRGINIA 22217
```

OCT. 26, 1977   JOB 460-101

SYSTEMATIC GEOGRAPHY LISTING

0011	1	OFFICE OF NAVAL RESEARCH
0011	2	SCIENTIFIC LIAISON OFFICER
0011	3	SCRIPPS INSTITUTION OF OCEANOGRAPHY
0011	4	LA JOLLA, CALIFORNIA 92093
0012	1	DIRECTOR, NAVAL RESEARCH LABORATORY
0012	2	ATTN LIBRARY, CODE 2628
0012	3	WASHINGTON, D. C. 20375
0013	1	ONR SCIENTIFIC LIAISON GROUP
0013	2	AMERICAN EMBASSY - ROOM A-407
0013	3	APO SAN FRANCISCO 96503
0101	1	COMMANDER
0101	2	NAVAL OCEANOGRAPHIC OFFICE LIBRARY
0101.	3	~~ATTN LIBRARY CODE 1600~~ NAVOCEANO - NSTL STATION
0101	4	~~WASHINGTON, D. C. 20374~~ BAY ST. LOUIS, MISSISSIPPI 3952
0102	1	NAVAL OCEANOGRAPHIC OFFICE
0102	2	~~CODE 3001~~ NSTL STATION
0102	3	~~WASHINGTON, D. C. 20374~~ BAY ST. LOUIS, MISSISSIPPI 3952
0103	1	CHIEF OF NAVAL OPERATIONS
0103	2	OP 987P1
0103	3	DEPARTMENT OF THE NAVY
0103	4	WASHINGTON, D. C. 20350
0104	1	OCEANOGRAPHER OF THE NAVY
0104	2	HOFFMAN II BUILDING
0104	3	200 STOVALL STREET
0104	4	ALEXANDRIA, VIRGINIA 22322
0105	1	NAVAL ACADEMY LIBRARY
0105	2	U. S. NAVAL ACADEMY
0105	3	ANNAPOLIS, MARYLAND 21402
0106	1	COMMANDING OFFICER
0106	2	NAVAL COASTAL SYSTEMS LABORATORY
0106	3	PANAMA CITY, FLORIDA 32401
0107	1	LIBRARIAN
0107	2	NAVAL INTELLIGENCE
0107	3	SUPPORT CENTER
0107	4	4301 SUITLAND ROAD
0107	5	WASHINGTON, D. C. 20390

OCT. 26, 1977 JOB 460-101

SYSTEMATIC GEOGRAPHY LISTING

0112	1	OFFICER IN CHARGE
0112	2	ENVIRONMENTAL RESEARCH PRDCTN FCLTY
0112	3	NAVAL POSTGRADUATE SCHOOL
0112	4	MONTEREY, CALIFORNIA 93940
0201	1	COMMANDING GENERAL
0201	2	MARINE CORPS DEVELOPMENT AND
0201	3	EDUCATIONAL COMMAND
0201	4	QUANTICO, VIRGINIA 22134
0202	1	DR. A. L. SLAFKOSKY
0202	2	SCIENTIFIC ADVISOR
0202	3	COMMANDANT OF THE MARINE CORPS
0202	4	CODE MC-RD-1
0202	5	WASHINGTON, D. C. 20380
1001	1	DEFENSE INTELLIGENCE AGENCY
1001	2	CENTRAL REFERENCE DIVISION
1001	3	CODE RDS-3
1001	4	WASHINGTON, D. C. 20301
1002	1	DIRECTOR
1002	2	DEFENSE MAPPING TOPOGRAPHIC CENTER
1002	3	ATTN= CODE 50200
1002	4	WASHINGTON, D. C. 20315
1003	1	COMMANDING OFFICER
1003	2	U.S. ARMY ENGINEERING
1003	3	TOPOGRAPHIC LABORATORY
1003	4	ATTN= ETL-ST
1003	5	FORT BELVOIR, VIRGINIA 22060
1005	1	CHIEF, WAVE DYNAMICS DIVISION
1005	2	USAE-WES
1005	3	P. O. BOX 631
1005	4	VICKSBURG, MISSISSIPPI 39180
2003	1	NATIONAL OCEANOGRAPHIC DATA
2003	2	CENTER (D764
2003	3	ENVIRONMENTAL DATA SERVICES
2003	4	NOAA
2003	5	WASHINGTON, D. C. 20235
2005	1	CENTRAL INTELLIGENCE AGENCY
2005	2	ATTENTION OCR/DD-PUBLICATIONS
2005	3	WASHINGTON, D. C. 20505

OCT. 26, 1977 JOB 460-101

SYSTEMATIC GEOGRAPHY LISTING

2007	1	DR. MARK M. MACOMBER
2007	2	ADVANCED TECHNOLOGY DIVISION
2007	3	DEFENSE MAPPING AGENCY (DELETED)
2007	4	NAVAL OBSERVATORY
2007	5	WASHINGTON, D. C. 20390
5001	1	MINISTERIALDIREKTOR DR. F. WEVER
5001	2	RUE/FO
5001	3	BUNDESMINISTERIUM DER VERTEIDIGUNG
5001	4	HARDTHOEHE
5001	5	D-5300 BONN, WEST GERMANY
5002	1	OBERREGIERUNGSRAT DR. ULLRICH
5002	2	RUE/FO
5002	3	BUNDESMINISTERIUM DER VERTEIDIGUNG
5002	4	HARDTHOEHE
5002	5	D-5300 BONN, WEST GERMANY
5003	1	MR. TAGE STRARUP
5003	2	DEFENCE RESEARCH ESTABLISHMENT
5003	3	OSTERBROGADES KASERNE
5003	4	DK-2100 KOBENHAVN O, DENMARK
5302	1	IR. M. W. VAN BATENBERG
5302	2	PHYSISCH LABORATORIUM TNO
5302	3	OUDE WAALSDORPER WEG 63, DEN HAAG
5302	4	NETHERLANDS
8002	1	COASTAL STUDIES INSTITUTE
8002	2	LOUISIANA STATE UNIVERSITY
8002	3	BATON ROUGE, LOUISIANA 70803
8003	1	DR. CHOULE J. SONU
8003	2	TEKMARINE, INC.
8003	3	37 AUBURN AVENUE
8003	4	SIERRA MADRE, CA 91024
9001	1	DR. LESTER A. GERHARDT
9001	2	RENNSSELAER POLYTECHNIC INSTITUTE
9001	3	TROY, NEW YORK 12181
9002	1	MR. FRED THOMSON
9002	2	ENVIRONMENTAL RESEARCH INSTITUTE
9002	3	P.O. BOX 618
9002	4	ANN ARBOR, MICHIGAN 48107

OCT. 26, 1977 JOB 460-101

SYSTEMATIC GEOGRAPHY LISTING

9004	1	DR. THOMAS K. PEUCKER
9004	2	SIMON FRASER UNIVERSITY
9004	3	DEPARTMENT OF GEOGRAPHY
9004	4	BURNABY 2, B. C., CANADA
9005	1	DR. ROBERT DOLAN
9005	2	DEPARTMENT OF ENVIRONMENTAL SCIENCES
9005	3	UNIVERSITY OF VIRGINIA
9005	4	CHARLOTTESVILLE, VIRGINIA 22903