NASA/TM–2000–104606, Vol. 18

Technical Report Series on Global Modeling and Data Assimilation

Max J. Suarez, Editor

Volume 18

An Assessment of the Predictability of Northern Winter Seasonal Means with the NSIPP 1 AGCM

Philip J. Pegion, Siegfried D. Schubert, and Max J. Suarez

December 2000

The NASA STI Program Office ... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.
- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.
- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and mission, often concerned with subjects having substantial public interest.
- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at http://www.sti.nasa.gov/STI-homepage.html
- E-mail your question via the Internet to help@sti.nasa.gov
- Fax your question to the NASA Access Help Desk at (301) 621-0134
- Telephone the NASA Access Help Desk at (301) 621-0390
- Write to:
  NASA Access Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076–1320

NASA/TM–2000–104606, Vol. 18

Technical Report Series on Global Modeling and Data Assimilation

Max J. Suarez, Editor
Goddard Space Flight Center, Greenbelt, Maryland

Volume 18

An Assessment of the Predictability of Northern Winter Seasonal Means with the NSIPP 1 AGCM

Philip J. Pegion
General Sciences Corporation, Laurel, Maryland

Siegfried D. Schubert
Data Assimilation Office, Goddard Space Flight Center, Greenbelt, Maryland

Max J. Suarez
Climate and Radiation Branch
NASA Seasonal to Interannual Predication Project
Goddard Space Flight Center, Greenbelt, Maryland

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

December 2000

Available from:

NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076–1320
Price Code: A17

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Price Code: A10

Abstract

This atlas assesses the predictability of January-February-March (JFM) means using version 1 of the NASA Seasonal-to-Interannual Prediction Project Atmospheric General Circulation Model (the NSIPP 1 AGCM). The AGCM is part of the NSIPP coupled atmosphere-land-ocean model. For these results, the atmosphere was run uncoupled from the ocean, but coupled with an interactive land model. The results are based on 20 ensembles of nine JFM hindcasts for the period 1980-1999, with sea surface temperature (SST) and sea ice specified from observations. The model integrations were started from initial atmospheric conditions (taken from NCEP/NCAR reanalyses) centered on December 15.

The analysis focuses on 200 mb height, precipitation, surface temperature, and sea-level pressure. The results address issues of both predictability and forecast skill. Various signal-to-noise measures are computed to demonstrate the potential for skillful prediction on seasonal time scales under the assumption of a perfect model and perfectly known oceanic boundary forcings. The results clearly identify El Nino/Southern Oscillation (ENSO)-related anomalies as the dominant seasonal mean signal in the tropics and the western hemisphere extratropics. Various probabilistic verification measures that compare the model simulations with observations are employed to assess the veracity of the model's ENSO response. The results show that the model produces a realistic ENSO response in both the tropics and extratropics. A comparison of the two major warm events of this period (JFM 1983 and 1998), employing larger ensembles, indicates that the model produces realistic and potentially predictable differences in the details of the atmospheric response to warm events.

Contents

List of Figures vii

1 Introduction 1

2 Description of the Model and Integration 1

3 Definitions and Computations 2

3.1 Signal and noise 2

3.2 Correlations 3

3.3 Probability Density Function 4

3.4 Reliability 4

4 Discussion of Results 5

4.1 Summary statistics 5

4.2 Model comparison 9

4.3 Individual years 10

4.4 Large ensemble comparisons 11

4.5 Probabilistic Verification 13

5 Conclusions 16

6 References 17

List of Figures

SUMMARY STATISTICS 19

1 Variance of JFM 200mb geopotential height for all years. 20

2 JFM 200mb geopotential height correlations for all years. 21

3 Variance of JFM precipitation for all years. 22

4 JFM precipitation correlations for all years. 23

5 Variance of JFM sea-level pressure for all years. 24

6 JFM sea-level pressure correlations for all years. 25

7 Variance of JFM surface temperature for all years. 26

8 JFM surface temperature correlations for all years. 27

9 Variance of JFM 200mb geopotential height for ENSO years. 28

10 JFM 200mb geopotential height correlations for ENSO years. 29

11 Variance of JFM precipitation for ENSO years. 30

12 JFM precipitation correlations for ENSO years. 31

13 Variance of JFM sea-level pressure for ENSO years. 32

14 JFM sea-level pressure correlations for ENSO years. 33

15 Variance of JFM surface temperature for ENSO years. 34

16 JFM surface temperature correlations for ENSO years. 35

17 Variance of JFM 200mb geopotential height for non-ENSO years. 36

18 JFM 200mb geopotential height correlations for non-ENSO years. 37

19 Variance of JFM precipitation for non-ENSO years. 38

20 JFM precipitation correlations for non-ENSO years. 39

21 Variance of JFM sea-level pressure for non-ENSO years. 40

22 JFM sea-level pressure correlations for non-ENSO years. 41

23 Variance of JFM surface temperature for non-ENSO years. 42

24 JFM surface temperature correlations for non-ENSO years. 43

MODEL COMPARISON 44

25 Spatial anomaly correlation of 500 mb height over the Pacific–North American sector (25N–70N and 60W–150W). 45

26 Total Variance of JFM 200 mb geopotential height in the Pacific–North American sector for the four models that participated in the DSP comparison, for the NSIPP-1 GCM, and for the reanalysis. 46

27 "Signal-to-Noise" ratio of the JFM 200 mb geopotential height in the Pacific–North American sector as in figure 26. 47

28 "Signal" of the JFM 200 mb geopotential height in the Pacific–North American sector as in figure 26. 48

29 "Noise"of the JFM 200 mb geopotential height in the Pacific–North American sector as in figure 26. 49

30 Total variance of precipitation in the tropical Pacific as in figure 26. 50

31 "Signal-to-Noise" ratio of precipitation in the tropical Pacific as in figure 26. 51

32 "Signal" of precipitation in the tropical Pacific as in figure 26. 52

33 "Noise" of precipitation in the tropical Pacific as in figure 26. 53

INDIVIDUAL YEARS 55

34 JFM geopotential height at 200 mb for 1980, 1981, and 1982. 56

35 JFM geopotential height at 200 mb for 1983, 1984, and 1985. 57

36 JFM geopotential height at 200 mb for 1986, 1987, and 1988. 58

37 JFM geopotential height at 200 mb for 1989, 1990, and 1991. 59

38 JFM geopotential height at 200 mb for 1992, 1993, and 1994. 60

39 JFM geopotential height at 200 mb for 1995, 1996, and 1997. 61

40 JFM geopotential height at 200 mb for 1998 and 1999. 62

41 JFM precipitation for 1980, 1981, and 1982. 63

42 JFM precipitation for 1983, 1984, and 1985. 64

43 JFM precipitation for 1986, 1987, and 1988. 65

44 JFM precipitation for 1989, 1990, and 1991. 66

45 JFM precipitation for 1992, 1993, and 1994. 67

46 JFM precipitation for 1995, 1996, and 1997. 68

47 JFM precipitation for 1998 and 1999. 69

48 JFM surface temperature for 1980, 1981, and 1982. 70

49 JFM surface temperature for 1983, 1984, and 1985. 71

50 JFM surface temperature for 1986, 1987, and 1988. 72

51 JFM surface temperature for 1989, 1990, and 1991. 73

52 JFM surface temperature for 1992, 1993, and 1994. 74

53 JFM surface temperature for 1995, 1996, and 1997. 75

54 JFM surface temperature for 1998 and 1999. 76

LARGE ENSEMBLE COMPARISONS 77

55 Model results for 1998 and 1983 for the JFM geopotential height at 200 mb. 78

56 Difference in geopotential height at 200 mb (1998 minus 1983). 79

57 Model results for 1998 and 1983 for the JFM precipitation 80

58 Difference in precipitation (1998 minus 1983). 81

59 Model results for 1998 and 1983 for the JFM sea-level pressure 82

60 Difference in sea-level pressure (1998 minus 1983). 83

61 Model results for 1998 and 1983 for the JFM surface temperature 84

62 Difference in surface temperature (1998 minus 1983). 85

63 Ratio of the intra-ensemble variance of seasonal means of 200mb geopotential height and precipitation (1998 / 1983). 86

64 Ratio of the intra-ensemble variance of seasonal means of surface temperature and sea-level pressure (1998 / 1983). 87

65 Model results for 1982 with 18 ensemble members for JFM-mean 200 mb height and precipitation. 88

66 Model results for 1989 with 18 ensemble members for JFM-mean 200 mb height and precipitation. 89

67 Difference in geopotential height at 200 mb (1983 minus 1989). 90

68 Difference in precipitation (1983 minus 1989). 91

69 Ratio of the intra-ensemble variance of seasonal means of 200mb geopotential height and precipitation (1989 / 1983). 92

PROBABILISTIC VERIFICATION **93**

70 Probability distribution function of the variance of the PNA index as defined on page 4 94

71 Probability distribution function of the variance of the area-mean precipitation in the tropical Pacific. 95

72 Scatter plot of precipitation and PNA index variance using sub-sampled synthetic realizations, as in the pdfs shown in the two previous figures. 96

73 Reliability diagram over the two hemispheres for geopotential height at 200 mb. 98

74 Reliability diagram over North America and Europe for geopotential height at 200 mb. 99

75 Reliability diagram over the two hemispheres for precipitation. 100

76 Reliability diagram over North America and tropical Pacific for precipitation. 101

77 Reliability diagram over the two hemispheres for sea-level pressure. 102

78 Reliability diagram over North America and Europe for sea-level pressure. 103

79 The 5th and 95th percentiles of the local distribution of the difference of geopotential height at 200 mb (1998 minus 1983). 104

80 The 5th and 95th percentiles of the local distribution of the difference of precipitation (1998 minus 1983). 105

81 The 5th and 95th percentiles of the local distribution of the difference of surface temperature (1998 minus 1983). 106

82 The 5th and 95th percentiles of the local distribution of the difference of sea-level pressure (1998 minus 1983). 107

1 Introduction

The mission of the NASA Seasonal-to-Interannual Prediction Project (NSIPP) is to use remotely-sensed observations to enhance the predictability of El Niño/Southern Oscillation (ENSO) and other major seasonal-to- interannual signals and their global teleconnections. Fullfilling this mission requires state-of-the-art general circulation models of the coupled ocean-atmosphere-land system that can be used to assimilate observations and to demonstrate the utility of those observations through experimental prediction.

This report presents an assessment of the predictability of seasonal means (January-February-March, JFM) using version 1 of the NSIPP Atmospheric General Circulation Model (the NSIPP 1 AGCM). This model, which is the atmosphere/land component of the full coupled atmosphere-land-ocean model, is run here uncoupled from the ocean, but coupled with an interactive land model. The NSIPP AGCM was developed at Goddard. NSIPP 1 is a production version of the development cycle Aries 1_1/Patch 4. The climate characteristics of this model, and a description of the model components are presented in a previous volume of this report series (Bacmeister et al. 2000).

The results presented here are from 20 ensembles of nine hindcasts[1] made with NSIPP 1 AGCM for the period 1980-1999, with sea surface temperature (SST) and sea ice specified from observations. For selected cases the ensemble size was increased to help assess the stability of the statistics and improve statistical significance of the results. The model integrations begin in mid-December with the various ensemble members starting from different atmospheric conditions centered on December 15. The atmospheric initial conditions and verification data are from the reanalysis performed by the National Centers for Environmental Prediction (the NCEP/NCAR Reanalysis, Kalnay et al., 1996). Precipitation is verified against the combined satellite and gauge-based estimates of Xie and Arkin (1997).

The results address issues of both predictability and forecast skill. Various signal-to-noise measures are computed to demonstrate the potential for skillful prediction on seasonal time scales under the assumption of a perfect model and perfectly known boundary forcings. Various probablistic verification measures that compare the model simulations with observations are employed to assess the veracity of the model's response to the boundary forcing.

Section 2 gives a brief overview of the model and the model integrations. Various definitions and details of the computations are given in section 3. The results are discussed in section 4, and the conclusions are given in section 5.

2 Description of the Model and Integration

For the runs described here, the NSIPP 1 model is run uncoupled from the ocean, but coupled with an interactive land surface model (the Mosaic LSM of Koster and Suarez (1992, 1996). Details of the NSIPP 1 atmospheric and land models are given in Bacmeister

[1]Strictly speaking these are not hindcasts since, by using observed SSTs, these could not be done in real time

et al. (2000) and references therein. We briefly discuss below some of the recent changes to the model and the most relevant parameterizations.

Compared with earlier versions of the model, the current version (development cycle Aries 1_1/Patch 4) has much improved stationary waves and sub-monthly variability (Bacmeister et al. 2000). The changes in Patch 4 include an increase in vertical resolution from 22 to 34 levels, with all new levels added near the surface; a modified version of the convection parameterization (see below); a modified version of the turbulence scheme, together with the elimination of dry convective adjustment; the use of filtered topography; and some minor modifications to the cloud disgnostic scheme.

The model uses the Relaxed Arakawa-Schubert (RAS) convection scheme (Moorthi and Suarez, 1992), which is a simple and efficient implementation of the Arakawa-Schubert scheme. The version described in Moorthi and Suarez, RAS-1, is the standard parameterization used at Goddard. It has also been tested at NCEP, NCAR, and COLA, and has performed particularly well in simulating the atmospheric response to tropical SST anomalies—a crucial aspect of the ENSO prediction problem. We have recently updated it by including a more detailed condensate budget in the updraft. This version, which we refer to as RAS 1.5, is the one used in the NSIPP 1 AGCM.

For these runs, the model was integrated at a resolution of 2° latitude by 2.5° longitude.

The basic results are from 20 ensembles of nine JFM hindcasts for the period 1980-1999 using specified SST and sea ice fractions based on the monthly Reynolds O-I dataset (Reynolds and Smith 1994). The initial atmospheric conditions are taken from the NCEP/NCAR reanalyses. The nine ensemble members for each year differ only in the atmospheric initial conditions. Members are started 12 hours apart centered on December 15. The initial soil conditions are the same for each year, and are taken from the December 1 state of a long model simulation. For selected years, the ensemble size was increased to 18 or 36 members and this is noted in the figures. In those cases the initial atmospheric conditions are taken 6 hours apart and again are centered on December 15.

3 Definitions and Computations

The following subsections provide definitions and details of the calculations of the quantities displayed in the atlas. These include signal-to-noise ratios, various correlations, and measures of reliability.

3.1 Signal and noise

Let x be a January-February-March (JFM) average of a particular quantity (e.g., 200mb height). We denote the ensemble mean of x by an overbar, and the 20 year mean of x by a square bracket. From the model runs, we define an unbiased estimate of the noise or intra-ensemble variance of x as

$$\sigma^2_{\text{noise}} = \frac{m}{m-1}\left[\overline{(x-\overline{x})^2}\right], \tag{1}$$

where m is the number of ensemble members. Following Rowell et al (1995), an unbiased estimate of the signal or inter-ensemble variance of x is

$$\sigma^2_{\text{signal}} = \frac{n}{n-1}\left[\overline{(\overline{x} - [\overline{x}])^2}\right] - \frac{1}{m}\sigma^2_{\text{noise}}, \tag{2}$$

where n is the number of years (in this case 20). The signal-to- noise ratio (SNR) is simply (2) divided by (1). The total variance of x is defined as the sum of the signal (2) and noise (1)

$$\sigma^2_{\text{totmod}} = \sigma^2_{\text{signal}} + \sigma^2_{\text{noise}}. \tag{3}$$

For the observations we have only one realization, and so the total variance of x is define as

$$\sigma^2_{\text{totobs}} = \frac{n}{n-1}\left[(x - [x])^2\right]. \tag{4}$$

3.2 Correlations

We compute the temporal correlation between the ensemble mean of the model runs $\overline{x}$ and the corresponding observation y as

$$r_{xy} = \frac{[(\overline{x} - [\overline{x}])(y - [y])]}{\sqrt{[(\overline{x} - [\overline{x}])^2]\,[(y - [y])^2]}}, \tag{5}$$

where the overbar and brackets are as defined above.

A difficulty with interpreting the correlation (5) is that it is not clear how large the values should be. For example, we would not expect them to be equal to 1 even for a perfect model with perfect boundary conditions, since small errors in the intial conditions limit the predictability of the atmosphere (Lorenz 1969). One approach to assessing the veracity of these correlations is to compare them with the results from a perfect model. To do this we remove one of the nine ensemble members and treat it as if it were an observation. We take this sythnetic observation and compute the correlation (5) between it and the ensemble mean of the 8 remaining ensemble members. This is repeated a number of times (100) by withholding different combinations of ensemble members from the twenty years. These are then averaged to obtain a "perfect model" correlation. To be consistent with these perfect model correlations, we compute the correlations with the true observations using only 8 ensemble members. This allows us to also generate different estimates of these correlations by withholding a different ensemble member each time in the computation of the 8-member ensemble mean—though since each ensemble mean differs in only a single ensemble member these tend to be very similar. These "imperfect model" estimates of the correlation are then averaged and compared with the "perfect model" results.

The correlations presented as a bargraph in Figure 25 are the spatial corelations between the ensemble mean of x ($\overline{x}$) and the observation y. These are defined as

$$s_{xy} = \frac{\langle(\overline{x} - \langle\overline{x}\rangle)(y - \langle y\rangle)\rangle}{\langle(\overline{x} - \langle\overline{x}\rangle)^2\rangle^{0.5}\langle(y - \langle y\rangle)^2\rangle^{0.5}}. \tag{6}$$

Here angle brackets denote a spatial mean. In Figure 25 the correlation is computed over the North American region (25N–70N, 60W–150W).

3.3 Probability Density Function

In pages 94 and 95 we show estimates of the probability density function (PDF) of the variance of the PNA index and the area -averaged tropical Pacific precipitation for various models. Following Wallace and Gutzler (1981), the PNA index is defined at the following grid points in terms of the 500mb JFM height anomaly ($Z' = Z - [Z]$) as

$$\text{PNA} = \frac{1}{4}(Z'_{20°N,160°W} - Z'_{45°N,165°W} + Z'_{55°N,115°W} - Z'_{30°N,85°W}). \tag{7}$$

The variance of PNA is

$$\sigma^2_{\text{PNA}} = \frac{n}{n-1}\left[(\text{PNA} - [\text{PNA}])^2\right], \tag{8}$$

with an analogous expression for the variance of the precipitation. Since we are comparing several models in pages 94 – 96, we limit the variance calculation to those years for which we also have data from the other models (1982-96). The PDF is determined by generating a large number of different estimates of the variance by taking different combinations of ensemble members. For the results shown in pages 94 – 96, we generated 100,000 different variances, and the values are binned over equal intervals in variance. The ensemble members composing each "realization" of 1982-96 were chosen at random for each year from a uniform distribution. By seeing where the variance that actually occurred in nature (computed from the observations) falls within the model's distribution of the variance, we can get a measure of the reliability of the model statistics. For precipitation we present results for two regions: (0-8° N, 160-200° E) and (0–8° S, 160–200° E). We also estimate the joint PDFs of the variance of the PNA index and the variance of the precipitation in the two regions. The estimates are obtained by requiring that the variances of the PNA index and precipitation are computed from the same permutations. The results are presented on page 96 as scatter plots. Other measures of the reliability of probabilistic forecasts are described below.

3.4 Reliability

A standard display of the performance of probability forecasts is the reliability diagram (Wilks 1995). The reliability diagram displays the frequency of occurrence of dichotomous events as a function of the forecast probability. The diagram also includes a histogram of the relative frequency of use of each of the probability forecast values. A perfectly reliable forecast would lie along the 45° diagonal (i.e., the forecast probability would be equal to the frequency of occurrence). The observed probability of an event is computed as the fraction of the ensemble members for which the event occurred. The observation is assigned a probability of 1 if the event occurred, and 0 otherwise. For the results shown in pages 98 through 103, events are defined as the occurrence of a positive (negative) anomaly that falls outside one standard deviation of the interannual variation of the quantity. The frequencies are computed from all 20 years and from every grid point making up the region in question.

We next describe our approach to assessing the reliability of the differences between the 1983 and 1998 simulations (pages 104 through 107). For this comparison we increased the

ensemble size to 36. Our purpose here is to establish whether the differences in model response are consistent with the differences actually observed, or whether they reflect deficiencies in the model or boundary forcing. At each grid point we generate $36 \times 36 = 1296$ differences by taking all combinations of the two sets of 36 ensemble members. We then determine the 95^{th} percentile by ordering the 1296 values from smallest to largest and taking the $(0.95 \times 1296 =)$ 1231st value. Similarly for the 5^{th} percentile. For example, the contoured field in the upper panel on page 104 is the 95^{th} percentile, based on a separate calculation at each grid point. The shading indicates those regions where the observations fall outside the 95^{th} percentile. In these regions, we conclude that the observed difference is larger than would be expected (at the 5% level) from the model's statistics.

We establish the field significance (Livezey and Chen 1983) of the results by determining how much of the globe would satisfy the criterion by chance if the observations indeed came from the same distribution as that of the model. For an inifinite number of independent gridpoints this would be simply 5% of the points. However, for a finite number of spatially correlated fields, this is generally not true. The field significance test uses a Monte Carlo approach to estimate the PDF of the significant area of the globe. We do this by generating 10000 realizations of this area (the 5% level in our case) using synthetic data. The synthetic data are constructed to have the same covariance structure as the model difference fields as follows:

$$x_{\text{syn}} = \sum_{i=1}^{N} \sqrt{\lambda_i}\alpha_i E_i, \tag{9}$$

where the α_is are unit-normal random variables with zero mean, and λ_i and E_i are the ith eigenvalue and eigenvector, respectively, of the 36×36 model difference fields.

4 Discussion of Results

4.1 Summary statistics

This section presents various statistics that summarize, over the 20 years (1980-1999), the signal and noise characteristics of the 9- member ensemble hindcasts. In addition, we present the correlations of the ensemble mean hindcast with observations (referred to as "imperfect model" correlations), and compare them with the correlations computed for an analogous "perfect model" hindcast (see section 3.2). The signal and noise are defined in section 3.1. The analysis is carried out for all years, ENSO years (1983, 85, 87, 89, 92, 98, 99), and non-ENSO years. The quantities presented are the 200mb height, precipitation, sea level pressure, and surface temperature. All results are for the January, February, March (JFM) mean fields.

The results based on all 20 years are presented in Figures 1-8. The model reproduces the total variance in the observed 200mb height field very well (Figure 1). The model does, however, show a general tendency to underestimate the height variance. The tropical height variance, though relatively small, is comprised almost entirely of signal. For example, over the eastern Pacific, where the tropical variance is largest, the signal-to-noise ratio (SNR) exceeds 30. Poleward of about 30N the variance is mainly noise, although the North Pacific

and the southern United States have areas with SNRs that exceed 2. As will become clearer in Figures 9-16, these areas correspond to centers of the wave-like ENSO response emanating from the central and eastern tropical Pacific.

Figure 2 (top panel) shows the correlations between the ensemble mean and the observed 200mb height. These "imperfect model" correlations (see section 3.2) exceed 0.6 over much of the tropics, the Pacific Ocean and parts of North America. We note that correlations less than 0.38 are not significantly different from zero at the 5% level, based on a Fischer's z-transform statistic (e.g. Stuart and Ord, 1994). For comparison, the bottom panel of Figure 2 shows remarkably similar correlations for a so-called "perfect model" calculation (see section 3.2). The main difference is a somewhat greater extension of the high tropical correlations into the subtropics and middle latitudes. Nevertheless, this comparison suggests that, for this field, the observations are nearly statistically indistinguishable from individual model ensemble members.

Results for precipitation are shown in Figures 3 and 4. The basic pattern of variability (largest over the central tropical Pacific) is reproduced well by the model. The magnitude is, however, less than observed throughout most of the tropics. The model has a local maximum in variability off the west coast of central America that is not found in the observations. The tropical precipitation variability is dominated by signal, with SNRs exceeding 5 over the central and eastern Pacfic. The signal is largely confined to the oceanic regions. SNRs drop rapidly outside the deep tropics, and are less than one everywhere poleward of about 20. The region of relatively high precipitation variance extending poleward into the North Pacific is comprised mainly of noise. The imperfect and perfect model correlations (Figure 4) both show high tropical correlations (greater than 0.8), with the largest values occurring over the tropical Pacific, and the lowest tropical correlations occurring over land and the Indian Ocean. The perfect model correlations, however, remain high over a broader region of the tropics and subtropics compared with the imperfect model correlations.

Results for the sea level pressure (Figures 5 and 6) are very similar to those for the 200mb height, with the model reproducing the total variance very well. In contrast to the results for the 200mb height the model, however, tends to produce more total variance in SLP than observed, especially at high latitudes and the North Pacific and North Atlantic oceans. The SNR is again largest in the tropics, though here the maximum SNRs occur over the western tropical Pacific. In the extratropics, the signal is strongest over the North Pacific and the eastern North Atlantic, however, the SNRs do not exceed 2 poleward of about 30°. The correlations for the perfect and imperfect models are again quite similar. The main differences are the weaker imperfect model correlations over the southern Indian Ocean and the central/subtropical tropical Pacific.

The model does a credible job of reproducing the observed total variance in the North American surface temperature (Figure 7). As a note of caution, the NCEP surface temperature, which we take here as the observed data, may be strongly influenced by model deficiencies, and are considered less reliable than the upper air reanalysis fields. Both the signal and noise are largest over western Canada, though the latter makes the largest contribution to the total variance. The SNRs are less than 1 over most of North America. In particular, the SNRs are less than 1/2 throughout the United States, with the smallest values occurring over the southwest where they drop below 0.1. The small SNRs are reflected in the small

perfect and imperfect model correlations (Figure 8) that occur throughout North America. Large correlations are mostly confined to the tropics and subtropics. The main exception is a region of correlations greater than 0.6 over northern Canada. Note that for this quantity, the correlations are computed only over land, since the model's sea surface temperatures are specified from observations.

Figures 9-16, are the same as Figures 1-8, except confined to the ENSO years (1983, 85, 87, 89, 92, 98, 99). The model's spatial pattern of total 200mb height variance (Figure 9) is similar to that for all years, however, in the tropical/subtropical central and eastern Pacific the variance is more than a factor of two larger. The increased tropical variance is primarily comprised of signal and reflects the response to the ENSO SST anomalies. The extratropical variance is also somewhat larger, though the contribution from the noise is largely unchanged from that for all years. The observed variance is again very similar to the model variance, however, the tendency for the model to underestimate the total variance is even more pronounced during these years. The SNR is, not surprisingly, generally larger for the ENSO years, with values greater than 3 over the eastern North Pacific and the south central United States. The imperfect and perfect model correlations (Figure 10) are again similar to those for all years, though the correlations are generally larger. Compared with the correlations for all years, the ENSO years show increased correlations over eastern Europe (exceeding 0.6), and reduced (less than $\sim$0.6) imperfect model correlations over northeast Asia.

Precipitation variance (Figure 11) for the ENSO years shows a general increase compared with all years, though the basic patterns are quite similar. An increase (compared to all years) in the signal off the west coast of the United States and to the southeast of the United States results in SNRs that exceed one in these regions. Otherwise, the extratropical SNRs poleward of 30° are predominantly less than one. The precipitation correlations (Figure 12) show a picture similar to that of the SNRs, with the largest correlations occurring over the tropical Pacific Ocean. In particular, the increase in correlation off the west and southeast coasts of the United States occurs in both the the imperfect and perfect model correlations.

The total sea level pressure variance (Figure 13) in the North Pacific for the ENSO years is double that for all years, largely due to an increase in signal. The basic tendency for the model to over estimate the total sea level pressure variance is similar to that for all years. The noise is also very similar to that for all years. The SNRs are larger than those for all years, with values greater than 5 occurring throughout much of the tropics, except over land. In the extratropics, SNRs exceed 3 off the west coast of the United States, and off the southeast coast of the United States. The SLP correlations (Figure 14) are again generally larger for the ENSO years. Compared with the 200mb height correlations, the SLP correlation maxima in the North Pacific, tend to occur further east just off the west coast of North America. It is noteworthy that the equatorward extension of the central North Pacific minimum in the imperfect model correlations is not found in the perfect model correlations.

For the ENSO years, the model's total variance in surface temperature (Figure 15) over Canada does not change much from that for all years. In contrast, the observations indicate a reduction in variance. There is also surprisingly little change in the signal compared to all years, in fact, over parts of the United States there is less signal during ENSO years.

The correlations (Figure 16) are generally higher for the ENSO years compared with those for all years. High perfect and imperfect model correlations occur over much of Africa, and the Americas. It is interesting that, over North America, the perfect model correlations are largest over Canada and Mexico, while the imperfect model correlations are largest along the west coast and the northeast.

Figures 17-24, are the same as Figures 1-8, except for the non-ENSO years. Compared with all years, the total 200mb height variance for non-ENSO years (Figure 17) is considerably reduced in many regions. The reduction is almost entirely due to the reduction in signal, since the noise is virtually unchanged compared with the results for all years. In the tropical Pacific the signal is a factor of 4 smaller, while over the North Pacific it is a factor of 2 smaller than the signal for all years. The SNR is large in the tropics, though the large values extend less far into the extratropics compared with those for all years. SNR is less than one everywhere in the extratropics, except for a region near 30N in the central North Pacific. Both the imperfect and perfect model correlations (Figure 18) of the 200mb height are substantially reduced compared with those from all years. Correlations above 0.4 are rare in the extratropics. In the tropics, there are considerable differences in the magnitudes of the perfect and imperfect model correlations: the perfect model correlations exceed 0.8 over much of the tropics, while the imperfect model correlations never reach 0.8.

The precipitation variance during non-ENSO years (Figures 19) is substantially reduced compared with the variance from all years. This is also evident in the observations. The change is again almost entirely due to the reduction in signal, since the noise is unchanged. Large SNRs are now confined to the central tropical Pacific. This reduced SNR is reflected in the correlations (Figure 20). Imperfect model correlations are quite low everywhere except over a few regions of the tropics (e.g. the central tropical Pacific where over a relatively small region the values reach 0.8). The perfect model correlations are somewhat larger and extend over a larger region with values greater than 0.6 occurring over much of the deep tropics.

The results for non-ENSO years for sea level pressure (Figure 21) are somewhat different from the results for the 200mb height. While the signal is again reduced over the North Pacific, there is also some increase in noise. The net effect is a less dramatic non-ENSO year reduction in total variance compared with what was found for the 200mb height variance. There is a dramatic reduction in the SNR compared with all years. With few exceptions, the SNR is less than 2 including the tropics. The model tendency to overestimate the total sea level pressure variance noted before is also present for the non-ENSO years. Both the perfect and imperfect model sea level pressure correlations (Figure 22) are weak. Correlations exceeding 0.6 occur only over parts of the tropical/subtropical oceans.

The model reproduces the observed surface temperature total variance reasonably well for the non-ENSO years (Figure 23). Note, however, for the model, the variance is reduced compared to that for the ENSO years, while for the observations, the variance is increased compared with the ENSO years. There is a surprising increase in signal over the north central United States compared with the ENSO years, though the SNR is still well below 1 in that region. The imperfect model correlations (Figure 24) are substantially reduced from those for all years (Figure 8). The perfect model correlations are largely unchanged in the tropics compared with those for all years, though outside the tropics the correlations

are somewhat weaker.

4.2 Model comparison

In this section we compare the NSIPP 1 results with those from several other AGCMs. In particular, the comparison includes the Center for Ocean-Land-Atmosphere Studies (COLA), the National Center for Atmospheric Research (NCAR), the Geophysical Fluid Dynamics Laboratory (GFDL), and the Goddard Earth Observing System (GEOS-2) AGCMs. These models all participated in the Dynamnical Seasonal Prediction (DSP) project (Shukla et al., 2000), allowing us to take advantage of the results from the ensemble forecasts already carried out for that project. In fact, the NSIPP 1 model runs described here follow the experimental design of the DSP project. We note that, while the NSIPP 1 AGCM was not ready in time for the first phase of the DSP project, it will be included in future DSP project comparisons. The DSP project AGCMs, runs, and results for the Northern winter are summarized in Shukla et al.(2000).

Figure 25 shows the spatial anomaly correlations (see section 3.2) between the AGCM ensemble mean and observations for the 500mb height over the Pacific-North American region. The figure also includes the Niño-3 SST anomalies. Note that results are available from all the models for only a subset of the 20 years. The comparison shows that the NSIPP 1 model correlations are very consistent with those from the other AGCMs. Correlations tend to be high for all models during the ENSO years, while that is generally not true during other years. Exceptions are the 1985 cold event, for which the models show low correlations, and 1990, for which the models show consistently high correlations, yet the Niño-3 SST anomalies are weak.

Figures 26-33 compare the variance, signal and noise from the various AGCMs. The results are for the five ENSO years (1983, 85, 87, 89, 92) for which we have results available from all the other models. Figure 26 shows the total variance of the JFM 200mb height for the AGCMs and the observations over the Pacific/North American region. The models show a large range of variability in the North Pacific. The NSIPP and COLA models show variance similar to the observed, while the othe models (especially GEOS-2) have considerably less variance. The SNRs (Figure 27) also differ substantially among the models. Values in the North Pacific range from greater than 10 for the COLA model to less than 2 for the NCAR model. All models show large SNRs in the tropics though here too the values show a wide range. The differences in the total variance and SNRs over the North Pacific reflect primarily the differences in the AGCMs signals (Figure 28), with the COLA and NSIPP models having the strongest signals. The only exception is the NCAR model, which has noise over the North Pacific (Figure 29) that is a factor of two to three larger than that of the other models.

Figure 30 shows the total variance of the JFM precipitation for the AGCMs and observations. All the models show a maximum in variability over the central tropical Pacific, consistent with the observations. There are, however, considerable differences among the models and observations in both the magnitude and spatial patterns of the variance. The COLA, GFDL and NCAR models have comparable variance magnitudes in the central tropical, with peak values somewhat larger than observed. The GEOS-2 model has variance

weaker than observed. The NSIPP model precipitation variance has very realistic peak values, though the area of large tropical variability is smaller than observed. Most of the models show enhanced variability off the west coast of Central America that is not present in the observations. All the models show large SNRs (Figure 31) in the tropical precipitation, though there is little agreement in magnitude or spatial distribution. The differences in the signal (Figure 32) are largely the same as those already discussed for the total variance. The noise (Figure 33), while a small component of the total variance, is considerably different among the models. The COLA and GFDL models, for example, have substantially more noise in the tropical precipitation than the other models. The relatively small noise in the NCAR precipitation is rather surprising in view of the very strong noise in the extratropical 200mb height field for that model (cf., Figure 29).

4.3 Individual years

In this section we present the ensemble mean anomalies, the observed anomalies, and the intra-ensemble variance for each of the 20 years. Results are presented for the global 200mb height (Figures 34-40), global precipitation (Figures 41-47), and North American surface temperature (Figures 48-54). The local statistical significance of the ensemble mean height, precipitation, and surface temperature anomalies are determined using a t-test. We apply the test to the precipitation anomalies under the assumption that the seasonal averaging results in approximately normally distributed variables (see e.g. von Storch and Zwiers 1999, p54).

The basic wave-like ENSO response emanating from the central and eastern Pacific is clearly evident in the observed and model ensemble mean height anomalies during 1983, 1985, 1987, 1989, 1992, 1998 and 1999. The 1999 cold event is, however, not well simulated by the model over the Pacific/North American region. In addition to the ENSO years, the model also generates significant height anomalies over the Pacific/North American region that verify in the observations during 1982 and 1990 (see also Figure 25). There is considerably interannual variability in the JFM noise, though nine ensemble members are not enough to provide reliable estimates of the intra-ensemble variance (see next section). There is some tendency for the noise over the North Pacific to be less than normal (cf. Figure 17) during strong warm events (e.g. 1983, 1987 and 1998). The noise is greater than normal during cold events (1985, 1989, and 1999). The largest noise occurs during 1982 though, as we shall see in the next section, the results are sensitive to the sample size.

The main tropical precipitation anomalies reflect the ENSO events of 1983, 1985, 1987, 1989, 1992, 1998 and 1999. For 1983, 1989, 1998 and 1999 the area of significant precipitation anomalies extend well into the extratropics especially over the North Pacific and North American regions. The precipitation noise appears to be greater (less) than normal over the central and eastern tropical Pacific (subtropical North Pacific) during major warm events (1983, 1987 and 1998), while the reverse is true during cold events (1985, 1989, 1999)and during 1982.

The model's ensemble mean JFM surface temperature anomalies are shown in Figures 48-54. They are compared against the surface temperature anomalies from the NCEP reanalysis data. Again, we note that the NCEP surface data may be strongly influenced by model

deficiencies, and are considered less reliable than the upper air reanalysis fields. During El Niño events the model consistently generates large warm ensemble mean anomalies extending southeast from northwestern Canada into the Great Lakes region (e.g., 1983, 1987, 1992, 1998). The observations show a similar, though less consistent response. For example, during 1983 the observations show warm anomalies confined to the north United States/southern Canda, while during 1998 the warm anomalies were confined to the eastern United States and Canada. The cold events show a less consistent response. For example, during 1985 there are almost no significant surface temperature anomalies over North America. In contrast, during 1989, the model produces cold anomalies over much of Canada and warm anomalies over the United States that, to a large extent, verify against the observations. These discrepancies between the ensemble mean anomalies and the observations, are consistent with large noise values, and the small signal to noise ratio for this quantity (see Figure 7). The model shows surprising agreement with the observations during 1982 with both showing strong cold anomalies over western and central Canada extending into the Great Lakes region.

4.4 Large ensemble comparisons

For some years of special interest we have increased the ensemble size to either 18 or 36 ensemble members in order to provide improved estimates of the intra-ensemble statistics. Selected results from these larger ensemble runs are presented in this section.

We start by comparing the two most extreme warm events (1983 and 1998) to determine whether the model responses for these two years are significantly different from one another. Figure 55 shows the 200mb height ensemble mean anomalies and intra-ensemble variance for these two years based on 36 ensemble members. The ensemble mean anomalies are generally quite similar to those from the 9-member ensembles shown earlier (section 4.3), with a deeper low over the North Pacific for the 1983 event. The major differences between the 9 member and 36 member results are in the estimates of the intra-ensemble variance. In particular, the 36 member results show clearly that the variance over the North Atlantic is about a factor of two larger than that over the North Pacific for both years. This was not evident in the 9 member results. A maximum in variance also occurs for both years over Northern Asia.

Figure 56 shows the difference between the 1998 and 1983 200mb height anomalies for the 36 member ensemble mean model results and the observations. The ensemble mean differences are largest over the North Pacific, the middle east, and the Southern Hemisphere (SH) middle and high latitudes. In the SH the differences show a pronounced zonally-symmetric signal, with higher heights during 1998 in the middle latitudes, and lower heights at high latitudes. The differences are significant at the 5% level over most regions, except for parts of northern Asia, Canada, the northern and eastern United States, and the North Atlantic between 30N and 60N. The observed differences are generally consistent with the statistically significant model anomalies. Regions where the model does not agree with the observations (e.g. the large observed anomalies over the North Atlantic and Asia) also tend to be regions where the intra- ensemble variance is large. In the next section we will examine whether the observations are outliers from the model's intra- ensemble variability in these regions.

Figure 57 compares the 36-member ensemble mean precipitation anomalies and their intra-ensemble variance for 1983 and 1998. The 36 member results are generally similar to those for the 9 member results shown earlier. The precipitation difference fields (Figure 58) from both the model and observations highlight that the tropical precipitation anomalies were larger during 1998 over the western Indian Ocean, and far eastern Pacific. On the other hand, the tropical precipitation anomalies were larger during 1983 over most of the central and western Pacific and eastern Indian Ocean. The model also shows enhanced precipitation during 1998 just north of the equator over the central and western Pacific that is not found in the observations. In the northern extratropics, the model anomalies show enhanced precipitation over the west coast of Canada, and reduced precipitation off the southern and eastern coasts of the United States. These anomalies, while significant, are in many regions different from the observed anomalies. We shall examine in the next section whether the observations are outliers with respect to the model ensemble members in these regions.

Figures 59 and 60 compare the 36-member ensemble mean sea level pressure anomalies and their intra-ensemble variance for 1983 and 1998. Over the North Pacific, the anomalous low pressure center is more than 6mb deeper during 1983, consistent with the 200mb heights shown earlier (Figure 55). The intra-ensemble variance of the sea level pressure is very similar in the two years. In the Southern Hemisphere, the differences have a strong zonally symmetric component with lower pressures between 30°S and 60°S and higher pressures at high latitudes during 1983, again consistent with the height differences shown earlier. The North Pacific and Southern Hemisphere differences are not inconsistent with the observations. The major discrepancy with the observations occurs over the North Hemisphere high latitudes where the observations show large positive pressure differences that are not found in the model results.

Figures 61 and 62 compare the 36-member ensemble mean surface temperature anomalies and their intra-ensemble variance for 1983 and 1998. Both years show positive temperature anomalies over much of Canada and the northern United States, northern South America, north eastern Australia, and South Africa. Positive anomalies (greater than 1 degree) occur over northern Asia during 1983. These areas also tend to have high intra-ensemble variance. Both the model and observed difference fields show warmer temperature over Asia during 1983. In other regions, for example North America, the model and observations show considerable disagreement.

The ratio of the 1998 to 1983 intra-ensemble variances are shown in Figures 63 and 64. The results show that during 1998 there was less 200mb height and sea level pressure variance over parts of the subtropics and more variance over Northern Asia and North America. For the precipitation, the interannual differences in variance tend to follow the differences in the ensemble mean precipitation. For example, the larger variance over the western Indian Ocean and southern Asia during 1998 occurs in regions with greater mean precipitation (cf. Figure 58). Surface temperature variance was smaller during 1998 over much of the tropical land masses, and parts of Asia.

Two other years of interest are 1982 and 1989. For these years, the ensemble size was increased to 18. As noted earlier, 1982 showed unusually large intra-ensemble variance in the 200mb height field from the nine-member ensemble (see section 4.3). Figure 65 shows the ensemble mean precipitation and 200mb height anomalies, and their intra-ensemble

variances. The ensemble mean anomalies are generally quite similar to those computed with just 9 members. With 18 ensemble members there is some reduction in the North Pacific maximum in the 200mb height variance (cf. Figures 34 and 65), though the variance is still unusually large compared with the other years, and the overall pattern is unchanged.

Next, we compare the the 1989 cold event with the 1983 warm event. Here, we are particularly interested in the interannual differences in the 200mb noise, and whether the results from the NSIPP model are consistent with those obtained with the GEOS-2 model (Schubert et al., 2000). We show first the 1989 18-member ensemble mean anomalies and intra-ensemble variance of the 200mb height and precipitation (Figure 66). The results are generally similar to those based on 9 ensemble members shown earlier. The differences between the 1989 and 1983 200mb height anomalies are shown for the model and observations in Figure 67. The precipitation differences are shown in Figure 68. The model produces very realistic height and precipitation differences, both in the tropics and extratropics. In fact, even some of the regional differences, such as greater precipitation over California and the southeast United States during the warm event, are well captured by the model. The ratio of the 1989 to 1983 intra-ensemble variance for the 200mb height and precipitation are shown in Figure 69. The results show a significantly greater height variance over the North Pacific, and decreased variance over the North Atlantic, consistent with the results of Schubert et al.(2000). The variance is significantly reduced over parts of the subtropics, and a substantial fraction of the Atlantic Ocean. The tropical/subtropical Pacific precipitation variance is larger north of the equator, and smaller on, and south of, the equator during 1989.

4.5 Probabilistic Verification

One of the difficulties with validating the model statistics (e.g. total variance) against observations is that the observations represent only a single realization from nature that is subject to sampling variability. For example, even though the NSIPP model generates total variance in the 200mb height field that is very similar to the observed variance (e.g. top panels of Figure 1), it may be that a hypothetical second 20 year realization of nature would produce a substantially different observed variance estimate, so that the good correspondence we obtained is simply a chance occurrence. While we clearly cannot produce further realizations of nature, we can carry out a more detailed analysis of the sampling variability of the model variance estimates, and ask whether the one realization from nature is an outlier with respect to the model's distribution of outcomes.

As an example, Figure 70 shows the probability density functions (pdfs) of the variance of the PNA index for the NSIPP, COLA and GFDL models (see section 3.3 for a description of the pdf calculation and the PNA index), computed for the years 1982 -1996. In short, the pdf describes the probability of the possible outcomes of the variance of the PNA index over the 15 years. The vertical line indicates the variance corresponding to the one outcome from the observations. Since the observed outcome is not an outlier for any of the model pdfs (does not fall in the tails of the distributions) we conclude that, for this quantity, the observations are not inconsistent with any of the models. It is interesting that there are substantial differences in the pdfs. In particular, the NSIPP model pdf shows a broader

distribution (larger variance), and it is shift to the right with respect to the other two models so that it has a mean closer to the observed value.

The analogous pdf plots for the variance of the area mean tropical precipitation are shown in Figure 71. Since the models show rather different latitudinal distributions of the tropical precipitation, we present results for two different regions, one north (0–8° N), and the other south (0–8° S) of the equator.[2] Both regions span the same longitude band (160° – 200°). For this quantity, there are clear discrepencies with the observations. North of the equator, all three models appear to be inconsistent with the observations: the NSIPP and GFDL model precipitation variability is too weak, and the precipitation variability of COLA model is too strong. South of the equator the precipitation variability is too weak for the NSIPP and COLA models, while that of the GFDL model appears to be consistent with the observations.

Figure 72 shows scatterplots comparing the variance of the PNA index with the variance of the area mean tropical precipitation for the two regions discussed in Figure 71. The observation is indicated by a dark square. These plots provide information on the joint distribution of the PNA index and precipitation variance. Surprisingly, none of the models show a clear linear relationship between the variance of the PNA index and the precipitation variance in either of the two regions, suggesting that the extratropical noise in these models is not strongly tied to noise in the tropical forcing. This is not inconsistent with our earlier results from the NCAR model (section 4.2) which showed unusually large noise in the extratropics, but only modest noise in the tropical precipitation.

We next examine the ability of the AGCM to reliably forecast the probability of occurrence of particular events. For dichotomus events, the reliability diagram (see Wilks 1995 and section 3.4) provides a useful display of the frequency of occurrence of events as a function of the forecast probability. By examining the results as a function of forecast probability, we can determine the ability of the model to distinguish between highly probable and less probable occurrences of the events. Figures 73-78 show reliability diagrams for various quantities and regions for the following two events: 1) the value of the quantity in question exceeds +1 standard deviation (std, panels on the right), and 2) the value falls below -1 std (panels on the left). The model std is computed at each grid point from the 20 years of hindcasts using (3). The std of the observations is computed from (4). The results in each plot are obtained by averaging the probabilities over all gridpoints included in the region of interest, and over either all, ENSO, or non-ENSO years. Figure 73 shows the results for the 200mb height averaged over the Southern and Northern Hemispheres. The predictions are biased toward higher probabilities than the observed frequency would suggest for both the +1 and -1 std events. For example, in the NH for all years and for ENSO years, when the model predictions show an 80% probability of occurrence of the event, the actual observed frequency is only about 55%. The results are even more biased for non-ENSO years, in which case the observed frequency is only about 40%. Overall, the model is more reliable for ENSO than non-ENSO years when it is predicting the event with high probability.

Figure 74 is similar to Figure 73 except that the results are for North America and Europe. These regions show interesting asymmetries between the ±1 std events. For the PNA region, the model is more reliable for +1 std events for all years and ENSO years. For the -1 std

[2]In practice, we chose for each model the gridpoint closest to 8° latitude.

event, the model shows a bias similar to that seen for the entire NH in Figure 61. A similar result is obtained for Europe. In this case, the model is more reliable for ENSO years for the -1 std event than for the +1 std event. Over Europe results must, however, be treated with more caution since there are few cases for which the model actually predicted the events with high probability.

Figures 75 and 76 are the reliability diagrams for precipitation. The curves in Figure 75 for both hemispheres are rather flat for all but the highest probability forecasts and do not distinguish very much between ENSO and non-ENSO years. This indicates that the probability forecasts for precipitation are generally not very reliable. That is not true for small regions. For example, over North America (Figure 76) , the -1 std event forecasts are very reliable for the ENSO years. Also, over the tropical Pacific area (10S to 10N, 160W to 160E), the +1 std events forecasts are reliable for ENSO years.

Figures 77 and 78 are the same as Figures 73 and 74, respectively, except for sea level pressure. The results are very similar to those for the 200mb heights. The forecasts over North America are, however, about equally reliable for both + 1 and -1 std events.

We next assess the reliability of the differences in the 1998 and 1983 anomalies generated by the model. In section 4.4 we established the statistical significance of the differences, based on the intra-ensemble noise produced by the model. Our purpose here is to establish whether the 1998/1983 differences in model response are consistent with the differences actually observed, or whether they reflect deficiencies in the model or boundary forcing. Our approach to this is described in section 3.4. The basic idea is that we determine, at each grid point, the 5^{th} the 95^{th} percentiles of the 1998/1983 differences from various permutations of the ensemble members. We then determine whether the observations fall outside these percentiles. If the total area of the regions where the observations fall outside the percentiles is field significant (see section 3.4), we conclude that the observed and model discrepancies must be due to model and/or boundary forcing deficiencies.

Figure 79 shows the 5^{th} the 95^{th} percentiles for the 1998-1983 200mb height differences. This shows, for example, that over the region of maximum differences in North Pacific, the values can locally be larger than 160 meters, and smaller than 0. The observations lie outside the 95^{th} percentile over much of the tropical/subtropical Indian and western Pacific oceans. Altogether observations are above 95^{th} percentile over 14.7% of the globe. This value is marginally field significant (values above 15.4% are field-significant at the one-sided 5% level.) On the other hand, observations are below the 5^{th} percentile over only 2.4% of the globe. This value is not field-significant at the one-sided 5% level.

Figure 80 is the same as Figure 79, except for the precipitation differences. The observed differences lie outside the 95^{th} percentile primarily over regions scattered about the tropics and subtropics (total area covers 3.6% of the globe). Similarly, the regions where the observations lie outside the 5^{th} percentile are scattered throughout the globe (total area also covers 3.9% of the globe). Neither area is field-significant at the one-sided 5% level.

Figure 81 shows the 5^{th} and 95^{th} percentiles for the 1998-1983 surface temperature differences. Over the winter continents the values reflect the predominantly positive differences over North America and the predominantly negative differences over Asia. The observations lie outside the 5^{th} (95^{th}) percentiles over an insignificant 2.2% (2.3%) of the land area

(percentages must be less than 0.65% to be field significant at the one-sided 5% level).

The 5^{th} the 95^{th} percentiles for the 1998-1983 sea level pressure differences are shown in Figure 82. For this quantity, the observed differences lie outside the 5^{th} and 95^{th} over less than 1% of the globe. These percentages are field significant at the one-sided 5% level, indicating that the model's range in sea level pressure differences is larger than that of the observations.

5 Conclusions

This report presents the results of 20 ensembles of nine January-February-March (JFM) hindcasts made with the NSIPP 1 AGCM for the period 1980-1999. Sea surface temperature (SST) and sea ice are specified from observations. The quantities examined are the global 200mb height, precipitation, sea level pressure, and North American surface temperature. The results show that the NSIPP AGCM produces very realistic January-February-March (JFM) mean interannual variability. Comparisons with several other AGCMs demonstrate that the NSIPP model is state-of-the-art. The comparisons also show that there is a wide range among the models in the signal-to-noise ratios (SNRs), with the NSIPP model producing SNRs in the extratropics that are on the high side compared with most of the other models.

The results further show that the potentially predictable signal in the extratropics is predominantly associated with ENSO. The wave-like ENSO response emanates from the central and eastern tropical Pacific and, in the Northern Hemisphere, extends across much of the North Pacific Ocean and North America. The latter two regions are characterized by some of the largest extratropical signal-to-noise ratios with, for example, maximum values exceeding 5 for the 200mb height field during the ENSO years.

A comparison of the 1983 and 1998 El Niño events using larger ensembles suggests that there are potentially predictable differences in the extratropical response for these two extreme events.

Various probablistic verification measures are presented that provide further indications of the reliability of the AGCM hindcasts. These include reliability diagrams, and various model probablity density functions (pdfs) that help to determine whether the observations fall within the range of model solutions.

6 References

Bacmeister J., P.J. Pegion, S.D. Schubert, and M.J. Suarez , 2000: Atlas of seasonal means simulated by the NSIPP 1 atmospheric GCM, *NASA Tech. Memo.* 104606, Vol. 17.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, J. Derber, L. Gandin, S. Sara, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A Leetma, R. Renolds, R Jenne, 1995: The NMC/NCAR Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.

Koster, R., and M. Suarez, 1992: Modeling the land surface boundary in climate models as a composite of independent vegetatio n stands. *J. Geophys. Res.*, **97**, 2697-2715.

Koster, R. and M. Suarez, 1996: Energy and Water Balance Calculations in the Mosaic LSM, *NASA Tech. Memo.* 104606, Vol. 9.

Livezey, R.E. and W.Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46-59.

Lorenz, E.N. 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289-307.

Moorthi, S. and M. Suarez, 1992: Relaxed Arakawa-Schubert: a parameterization of moist convection for general circulation models. *Mon. Weather Rev.*, **120**, 978-1002.

Reynolds, W. R. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929-948.

Rowell,D.P., C.Folland, K.Maskell, and N.Ward, 1995: Variability of summer rainfall over tropical north Africa (1906-92): Observations and modelling. *Q. J. R. Meteor. Soc.*, **121**, 669-704.

Schubert,S. D., M. J. Suarez, Y. Chang and G. Branstator, 2000: The impact of ENSO on extratropical low frequency noise in seasonal forecasts. *Submitted to J. Climate.*

Shukla, J., J. Anderson, D. Baumhefner, Y. Chang, E. Kalnay, L. Marx, D. Paolino, S. Schubert, D. Straus, M. Suarez, J. Tribbia, 2000: Dynamical Seasonal Prediction, To appear in *Bull. Amer. Met. Soc.*

Storch, H.V. and F. Zwiers, 1999: Statistical Analysis in Climate Research, Cambridge University Press, 484pp.

Stuart, A. and J.K. Ord, 1994: Kendall's Advanced Theory of Statistics. Halsted Press, New York, 676pp.

Wallace, J.M. and D.S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 784-812.

Wilks, D. S., 1995: Statistical methods in the Atmospheric Sciences. International Geophysics Series, volume 59, Academic Press, 467pp.

Xie, P. and P. Arkin, 1997: Global precipitation, a 17-Year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. *Bull. Am. Met. Soc.*, **78**, 2539-2558.

SUMMARY STATISTICS

All Years, ENSO Years, Non-ENSO Years

Geopotential Height Variance at 200 mb

Precipitation Variance

Sea-Level Pressure Variance

Geopotential Height Correlation at 200 mb

Precipitation Correlation

Sea-Level Pressure Correlation

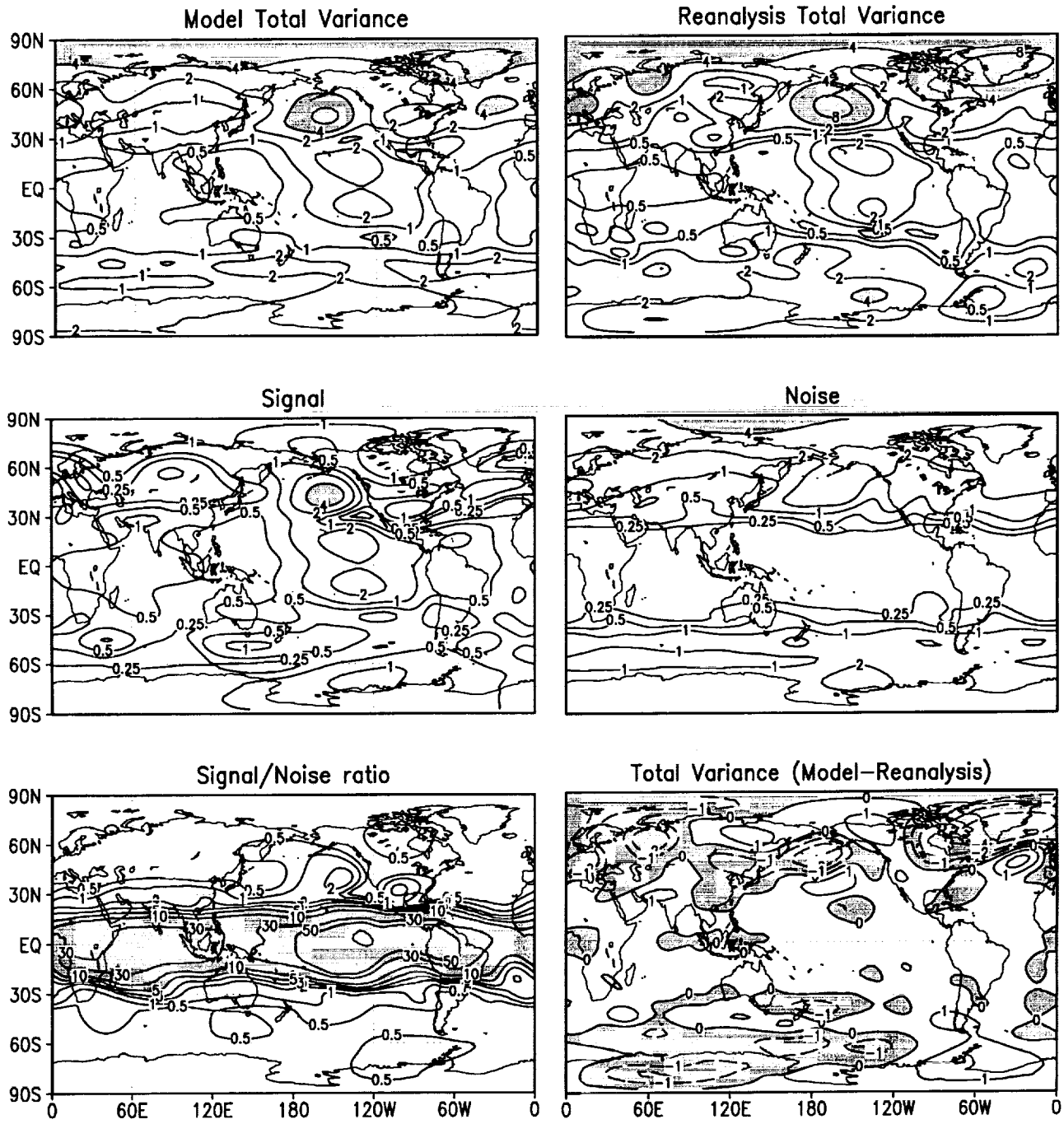

Figure 1: Variance of JFM 200mb geopotential height for all years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are .25 .5 1 2 4 8 16 32 $\times$ 10^3 m^2, with values above 4000 m^2 shaded. For the difference (bottom-right) the contours are $\pm$ 0 1 2 4 $\times$ 10^3 m^2, with negative values shaded. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100 200, with values above 5 shaded.

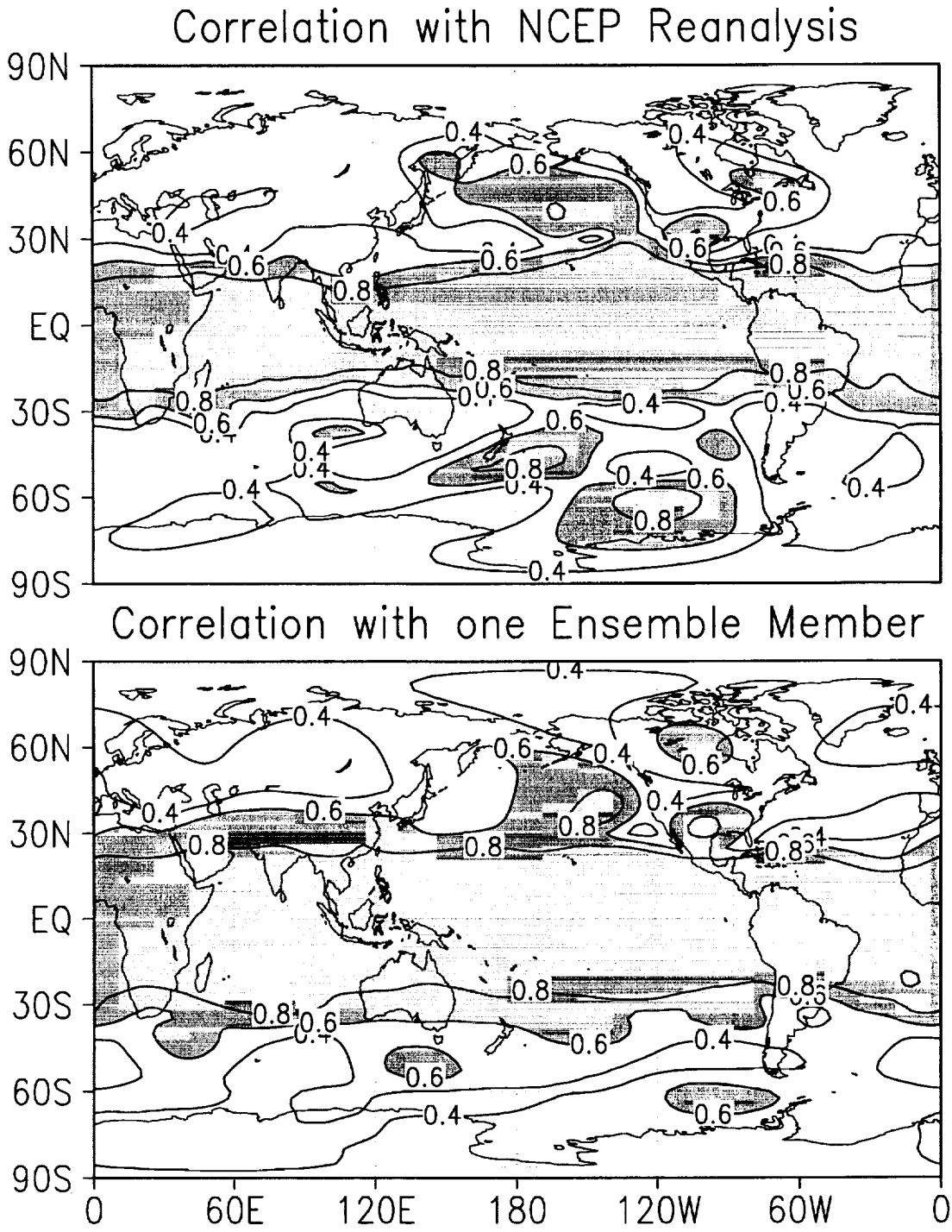

Figure 2: JFM 200mb geopotential height correlations for all years. Top: The mean correlation over all years of 8-member ensemble means with the observations. Bottom: The mean correlation over all years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ± 0.4 are not contoured. Negative contours are dashed.

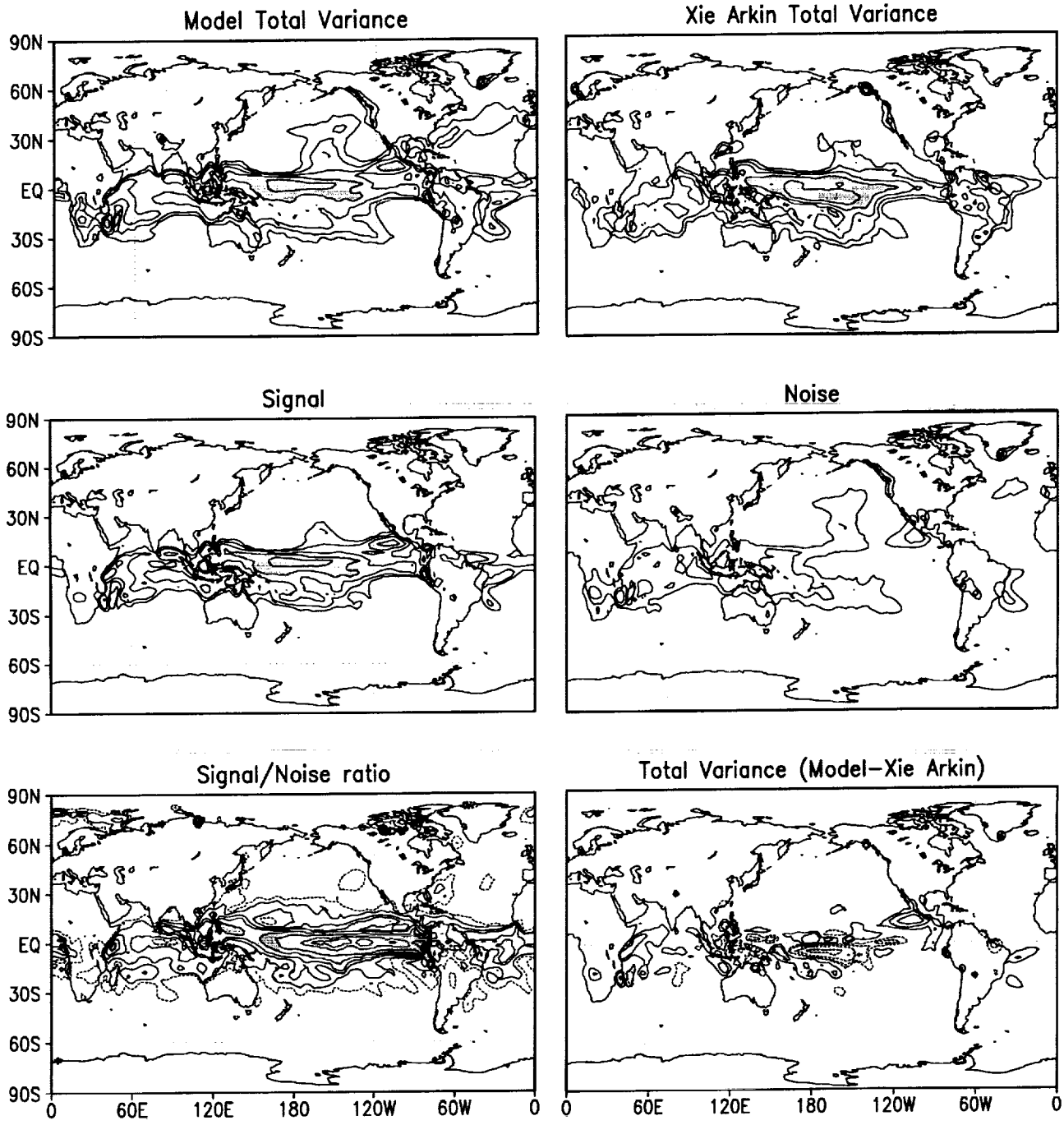

Figure 3: Variance of JFM precipitation for all years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are 1 2 4 8 16 32 64 $(\text{mm/day})^2$; values above 8 are shaded. For the difference (bottom-right) the contours are $\pm$ 2 4 8 16 32 64 $(\text{mm/day})^2$; negative contours dashed. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100; values above 5 are shaded and the 0.5 contour is dashed.

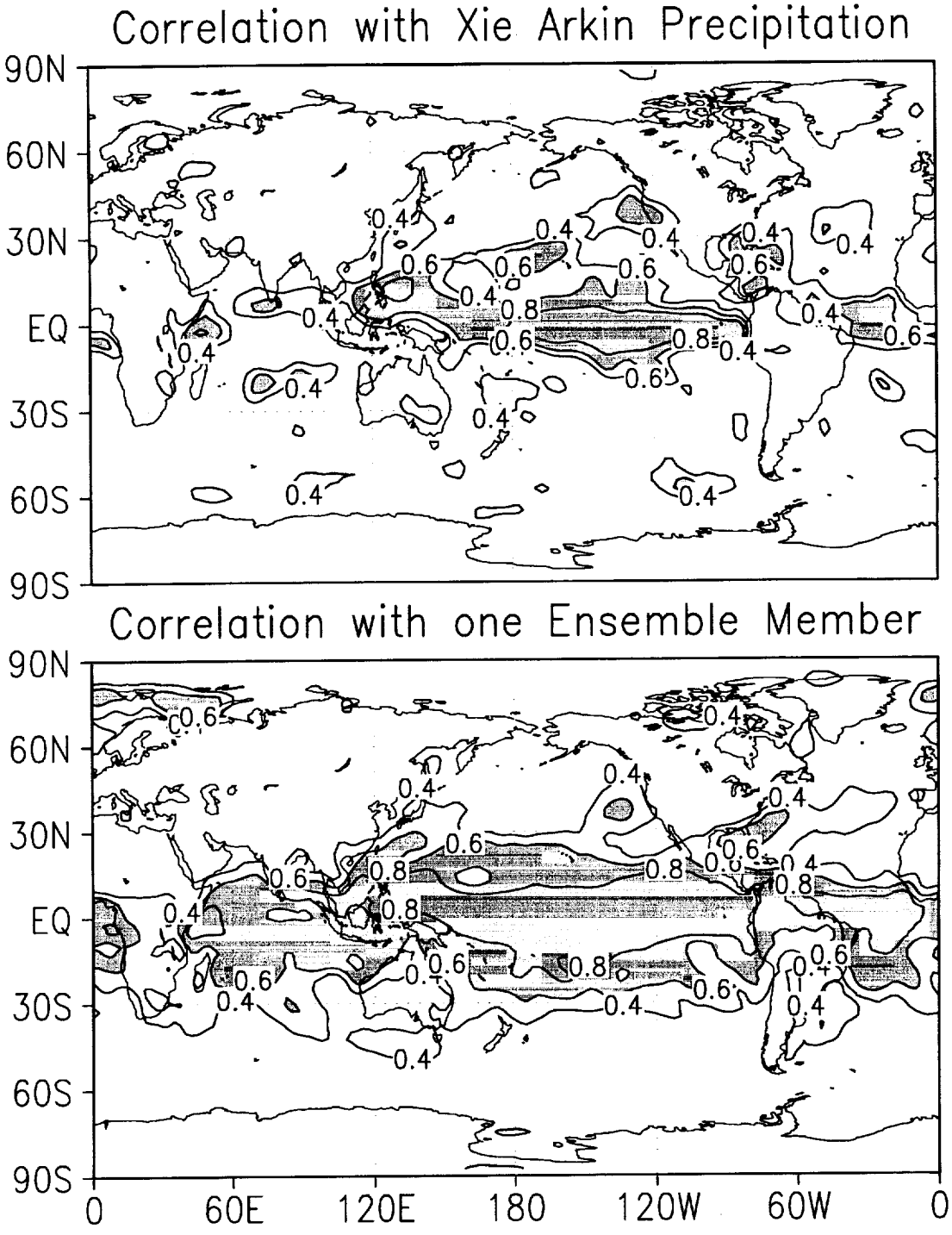

Figure 4: JFM precipitation correlations for all years. Top: The mean correlation over all years of 8-member ensemble means with the observations. Bottom: The mean correlation over all years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ±0.4 are not contoured. Negative contours are dashed.

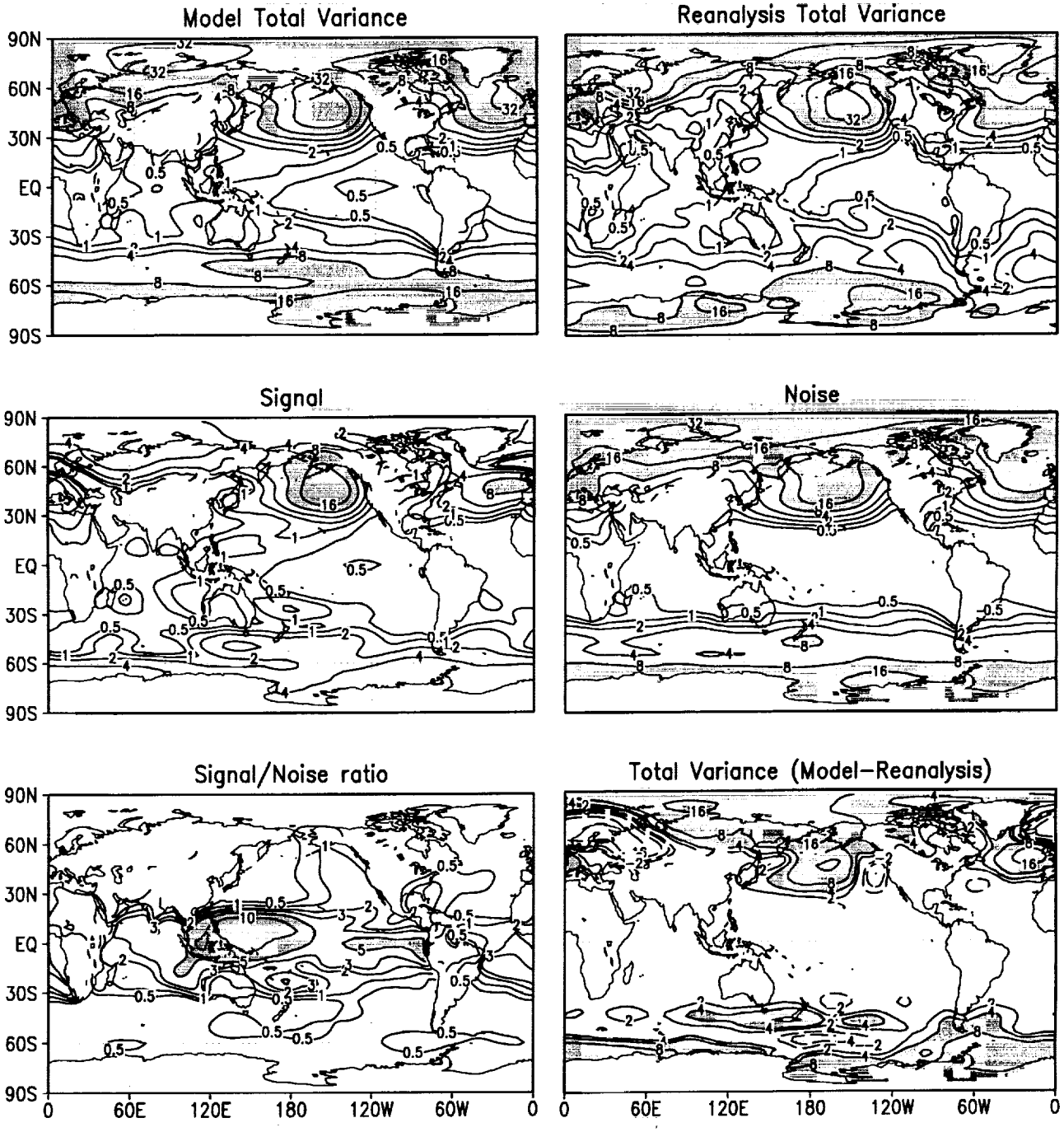

Figure 5: Variance of JFM sea-level pressure for all years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. The variances are in units of mb^2. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100 200, with values above 5 shaded. The contours for the difference are $\pm$ 2 4 8 16 mb^2, with values above +4 mb^2 shaded; negative contours are dashed.

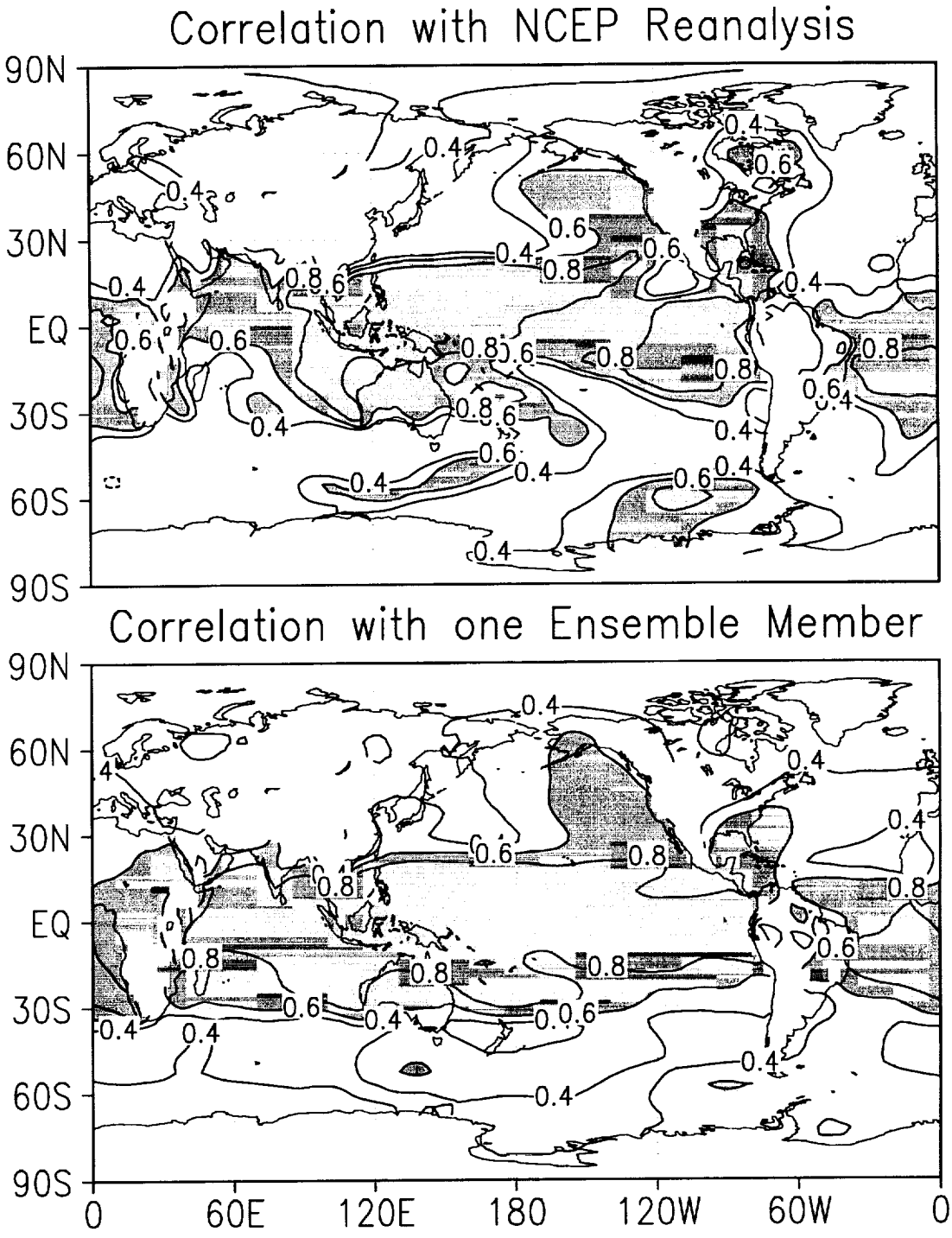

Figure 6: JFM sea-level pressure correlations for all years. Top: The mean correlation over all years of 8-member ensemble means with the observations. Bottom: The mean correlation over all years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ±0.4 are not contoured. Negative contours are dashed.

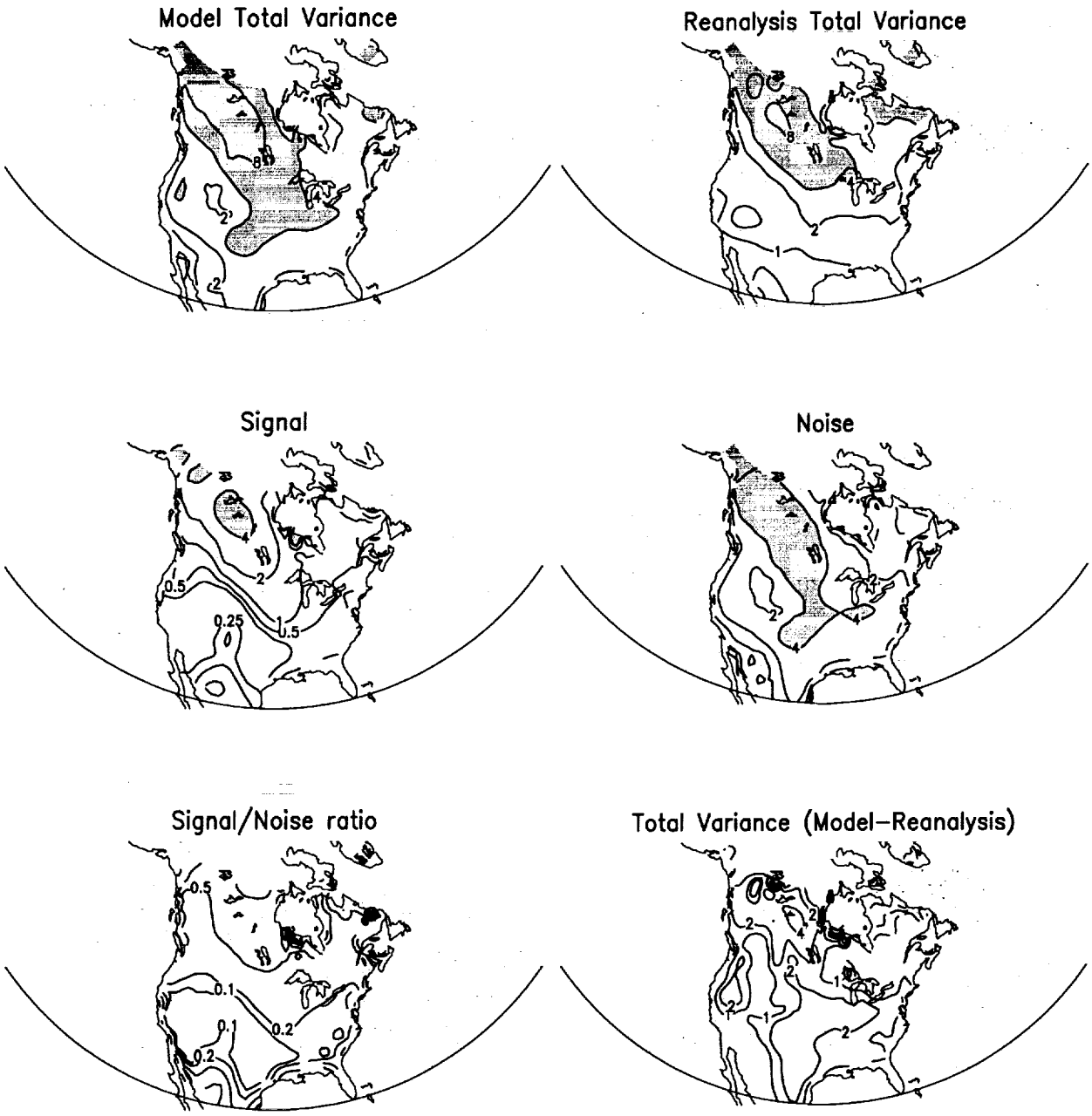

Figure 7: Variance of JFM surface temperature for all years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are 0.25 0.5 1 2 4 8 K^2, with values above 4 shaded. The contours for the difference are $\pm$ 1 2 4 8 K^2. The contours for the ratio are 0.1 0.2 0.5 1 2.

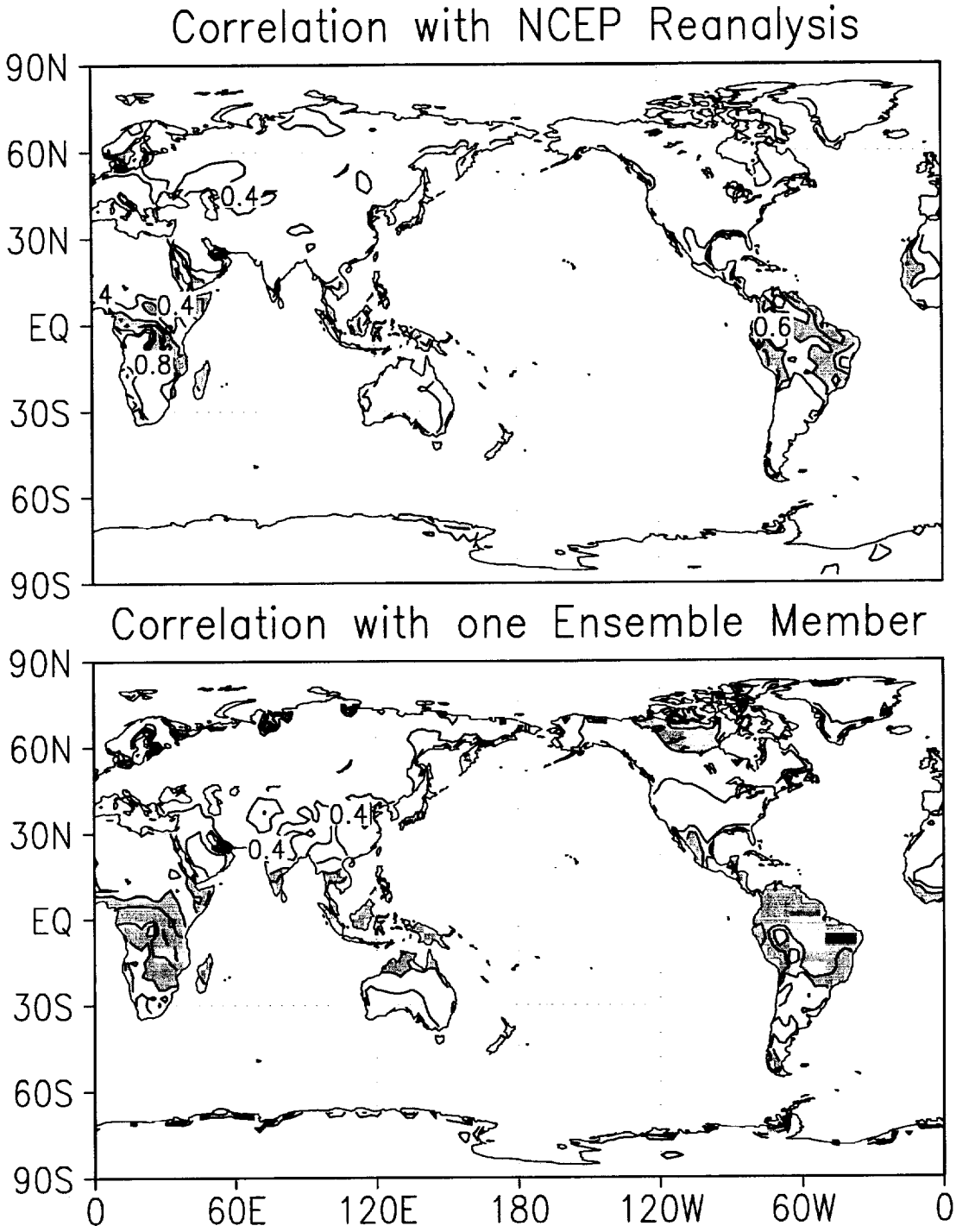

Figure 8: JFM surface temperature correlations for all years. Top: The mean correlation over all years of 8-member ensemble means with the observations. Bottom: The mean correlation over all years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ± 0.4 are not contoured. Negative contours are dashed.

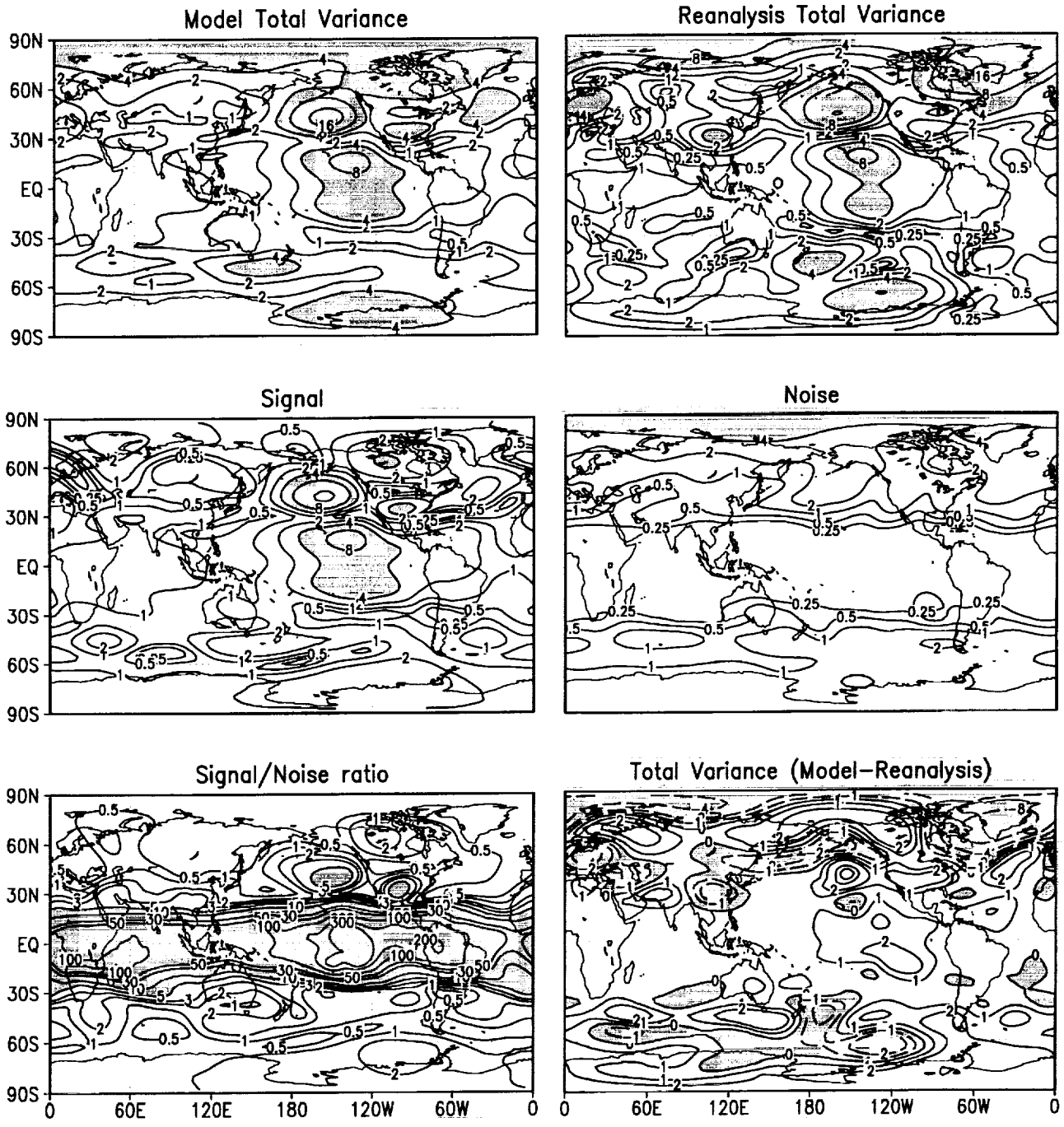

Figure 9: Variance of JFM 200mb geopotential height for ENSO years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are .25 .5 1 2 4 8 16 32 $\times 10^3$ m^2, with values above 4000 m^2 shaded. For the difference (bottom-right) the contours are $\pm$ 0 1 2 4 $\times 10^3$ m^2, with negative values shaded. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100 200, with values above 5 shaded.

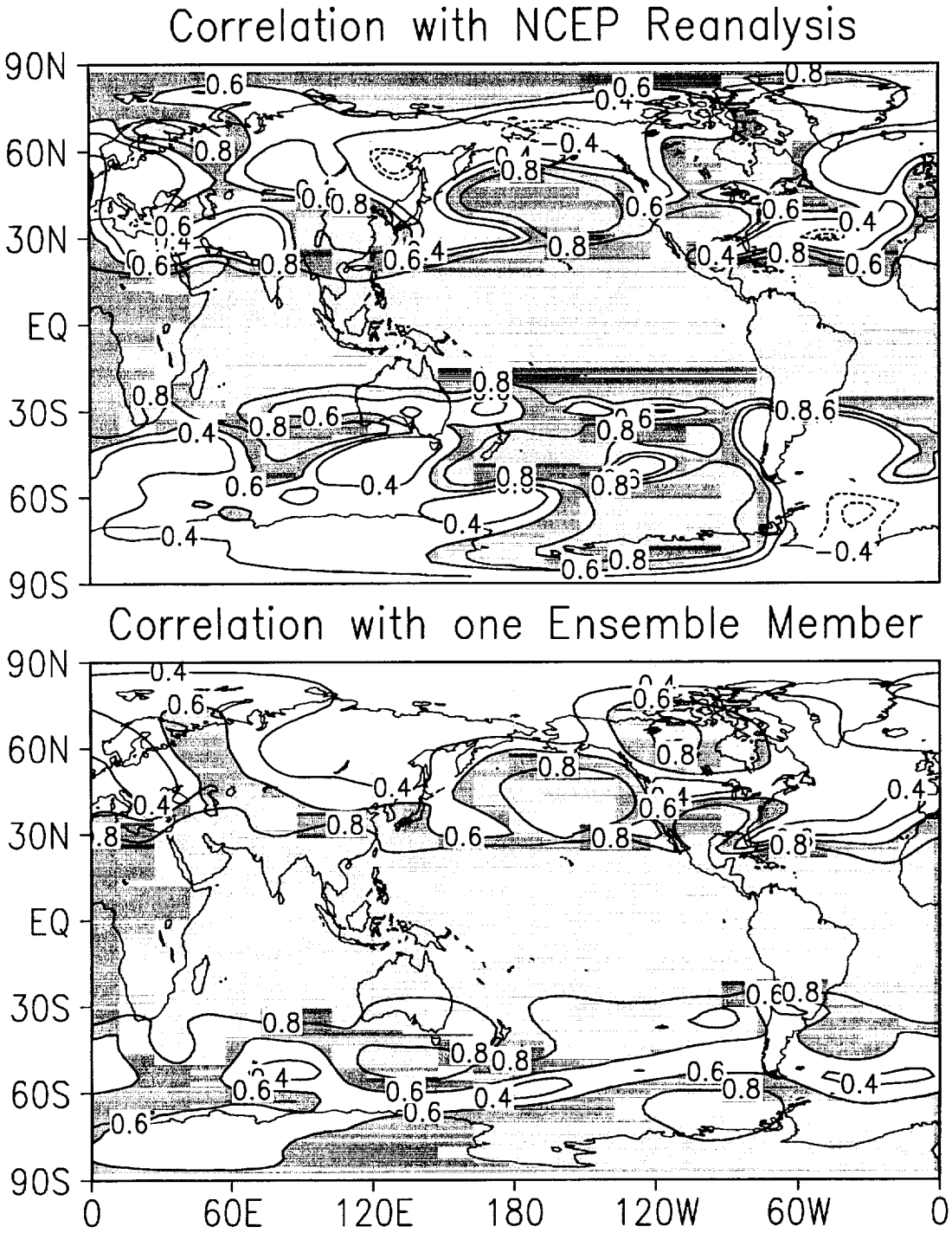

Figure 10: JFM 200mb geopotential height correlations for ENSO years. Top: The mean correlation over ENSO years of 8-member ensemble means with the observations. Bottom: The mean correlation over ENSO years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ±0.4 are not contoured. Negative contours are dashed.

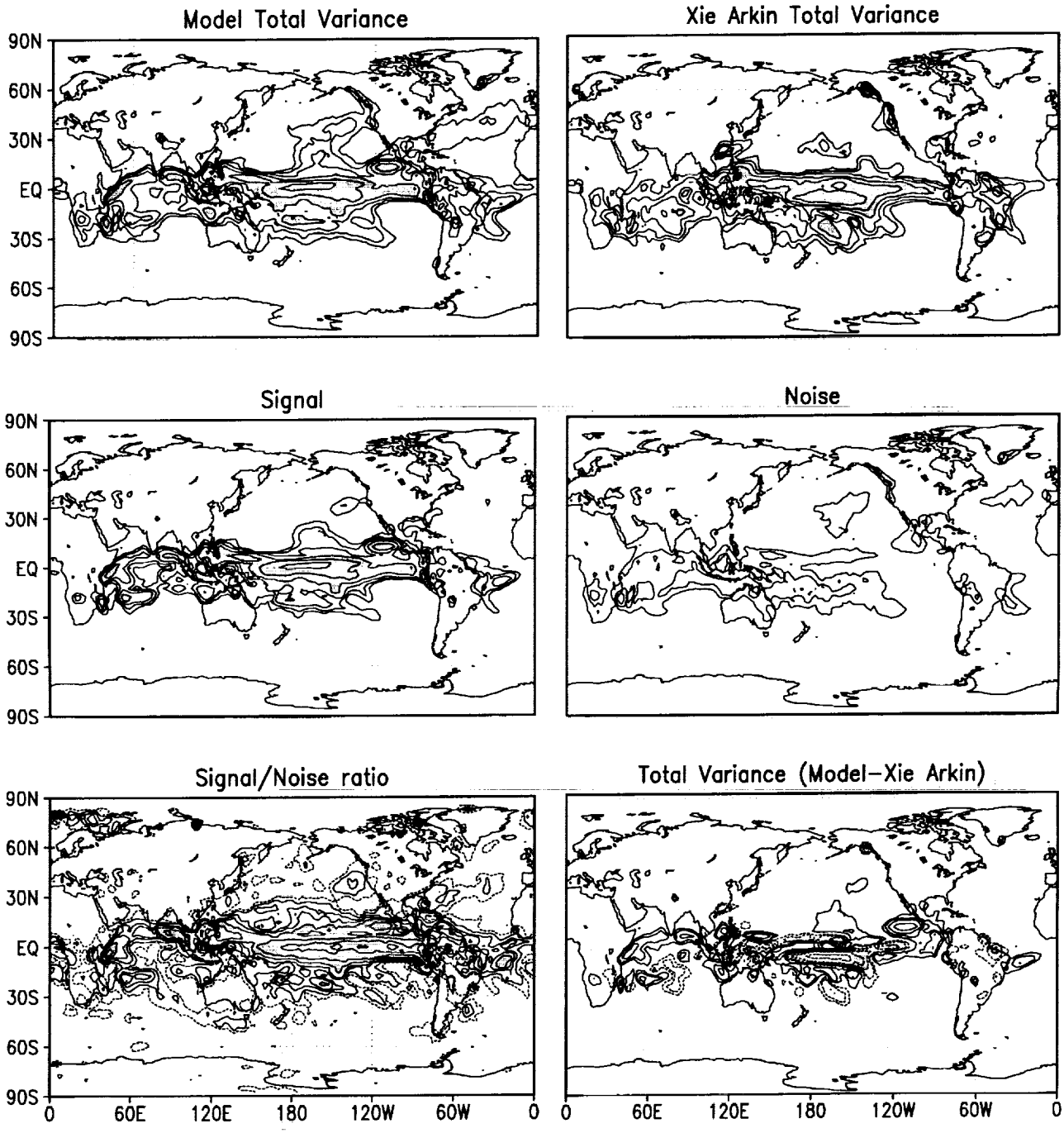

Figure 11: Variance of JFM precipitation for ENSO years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are 1 2 4 8 16 32 64 $(\text{mm/day})^2$; values above 8 are shaded. For the difference (bottom-right) the contours are $\pm$ 2 4 8 16 32 64 $(\text{mm/day})^2$; negative contours dashed. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100; values above 5 are shaded and the 0.5 contour is dashed.

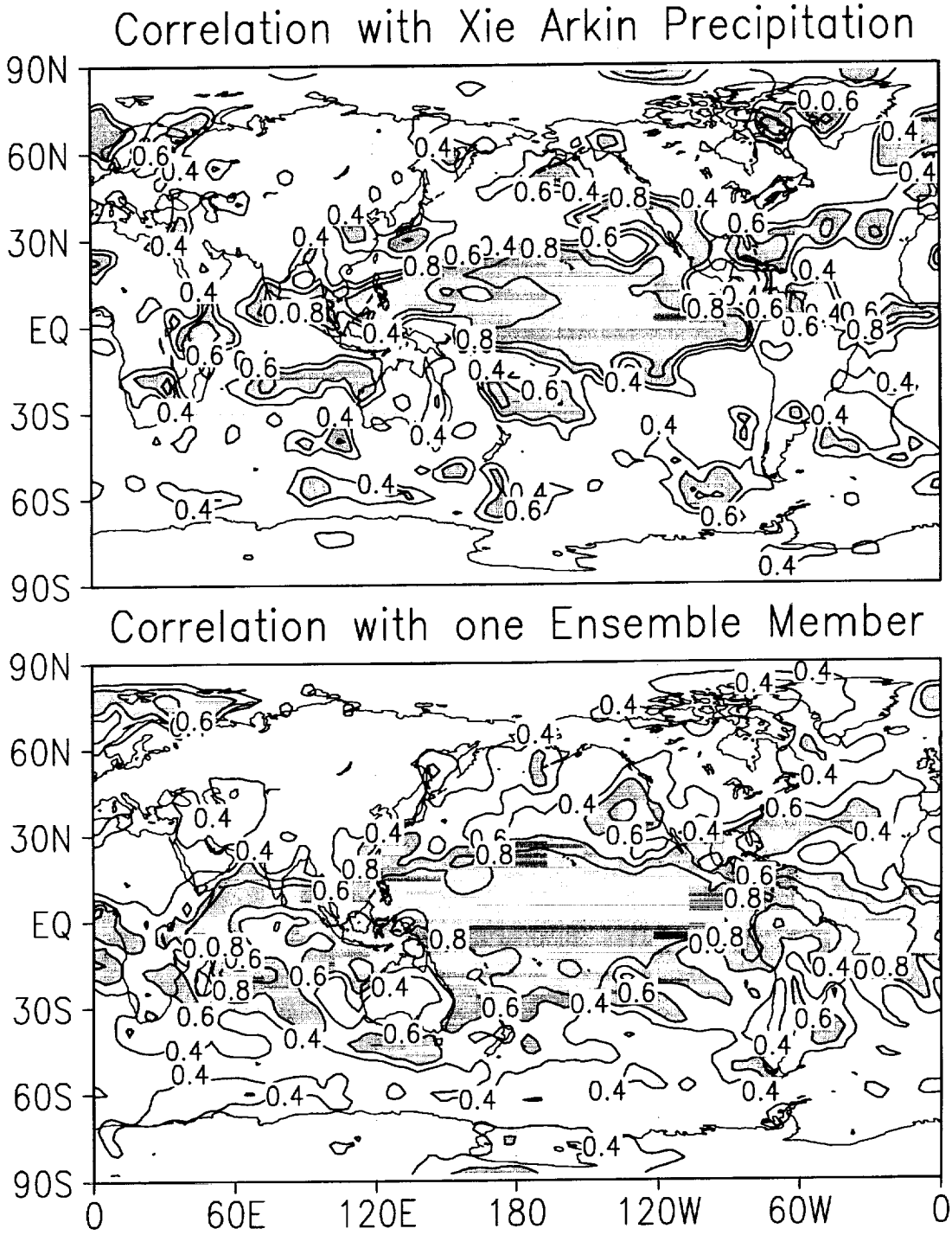

Figure 12: JFM precipitation correlations for ENSO years. Top: The mean correlation over ENSO years of 8-member ensemble means with the observations. Bottom: The mean correlation over ENSO years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ± 0.4 are not contoured. Negative contours are dashed.

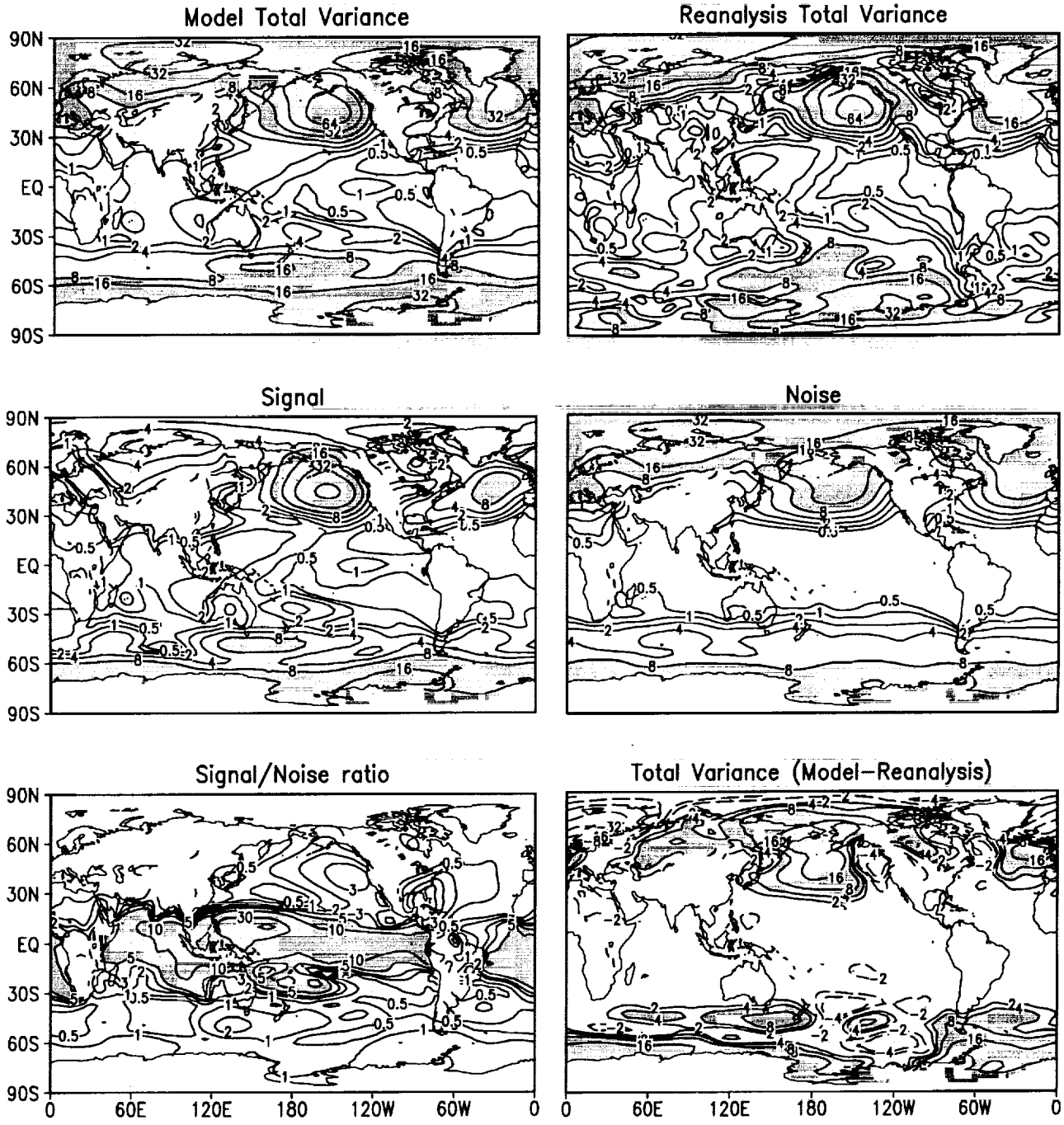

Figure 13: Variance of JFM sea-level pressure for ENSO years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. The variances are in units of mb^2. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100 200, with values above 5 shaded. The contours for the difference are $\pm$ 2 4 8 16 mb^2, with values above +4 mb^2 shaded; negative contours are dashed.

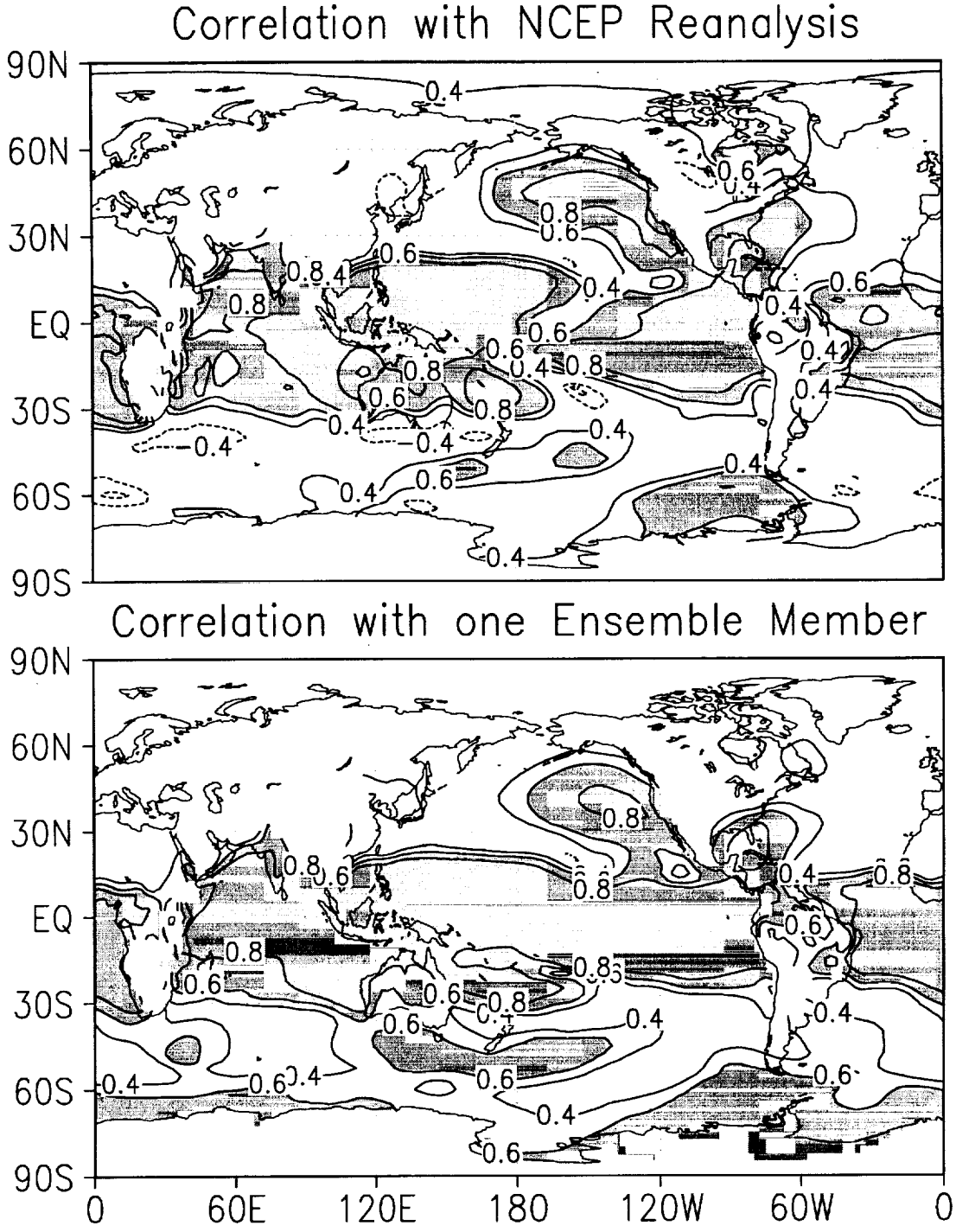

Figure 14: JFM sea-level pressure correlations for ENSO years. Top: The mean correlation over ENSO years of 8-member ensemble means with the observations. Bottom: The mean correlation over ENSO years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ±0.4 are not contoured. Negative contours are dashed.

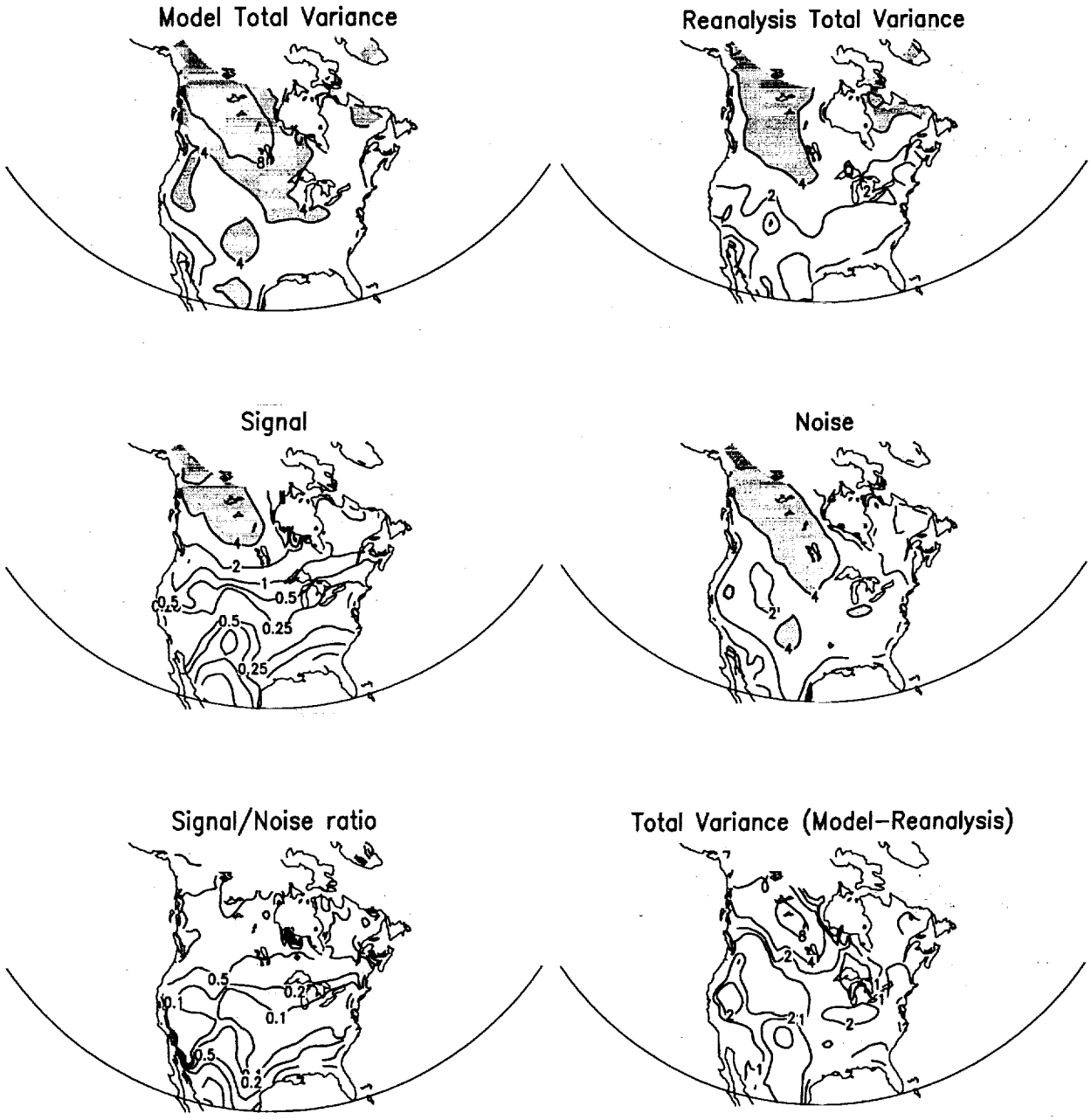

Figure 15: Variance of JFM surface temperature for ENSO years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are 0.25 0.5 1 2 4 8 K^2, with values above 4 shaded. The contours for the difference are $\pm$ 1 2 4 8 K^2. The contours for the ratio are 0.1 0.2 0.5 1 2.

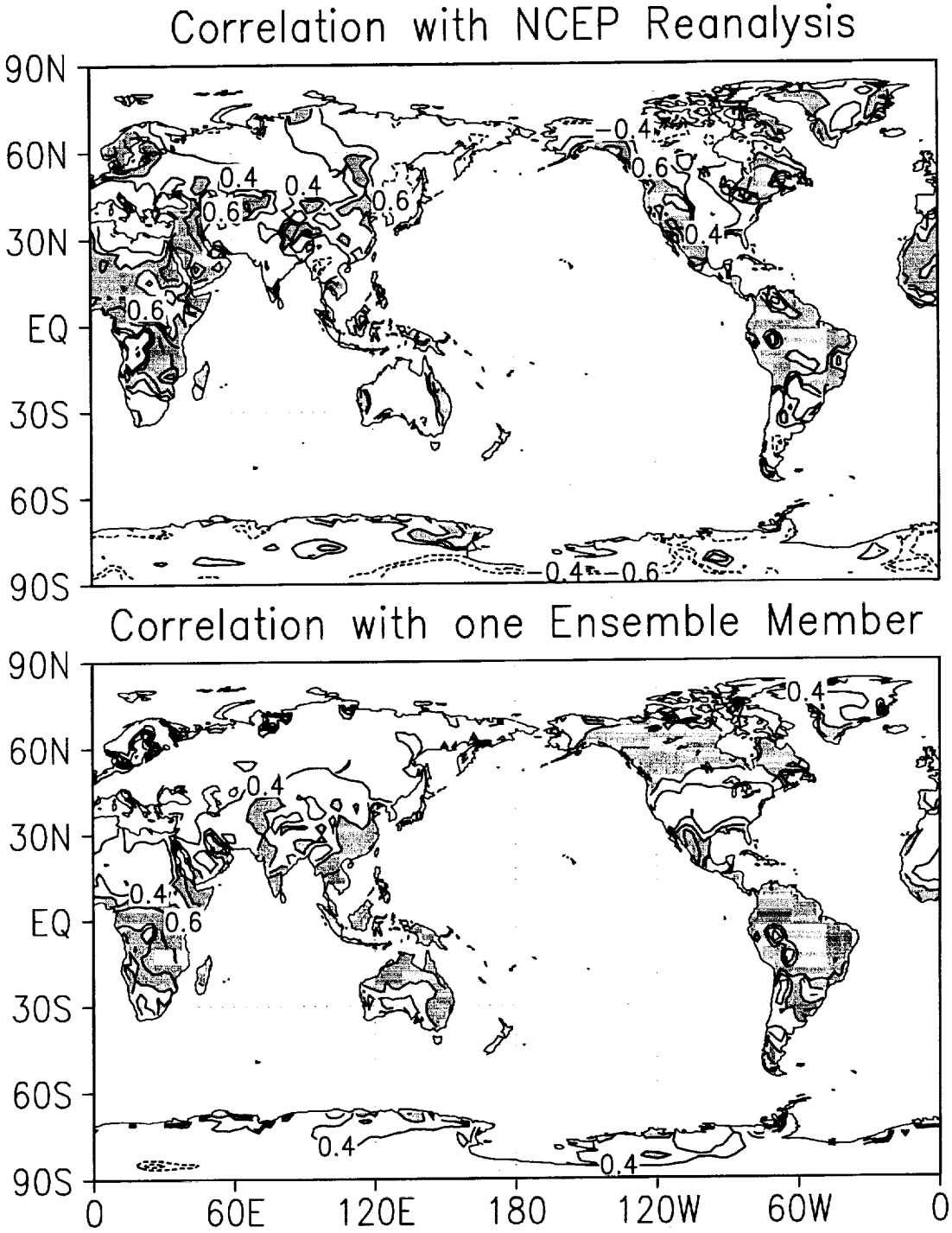

Figure 16: JFM surface temperature correlations for ENSO years. Top: The mean correlation over ENSO years of 8-member ensemble means with the observations. Bottom: The mean correlation over ENSO years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ±0.4 are not contoured. Negative contours are dashed.

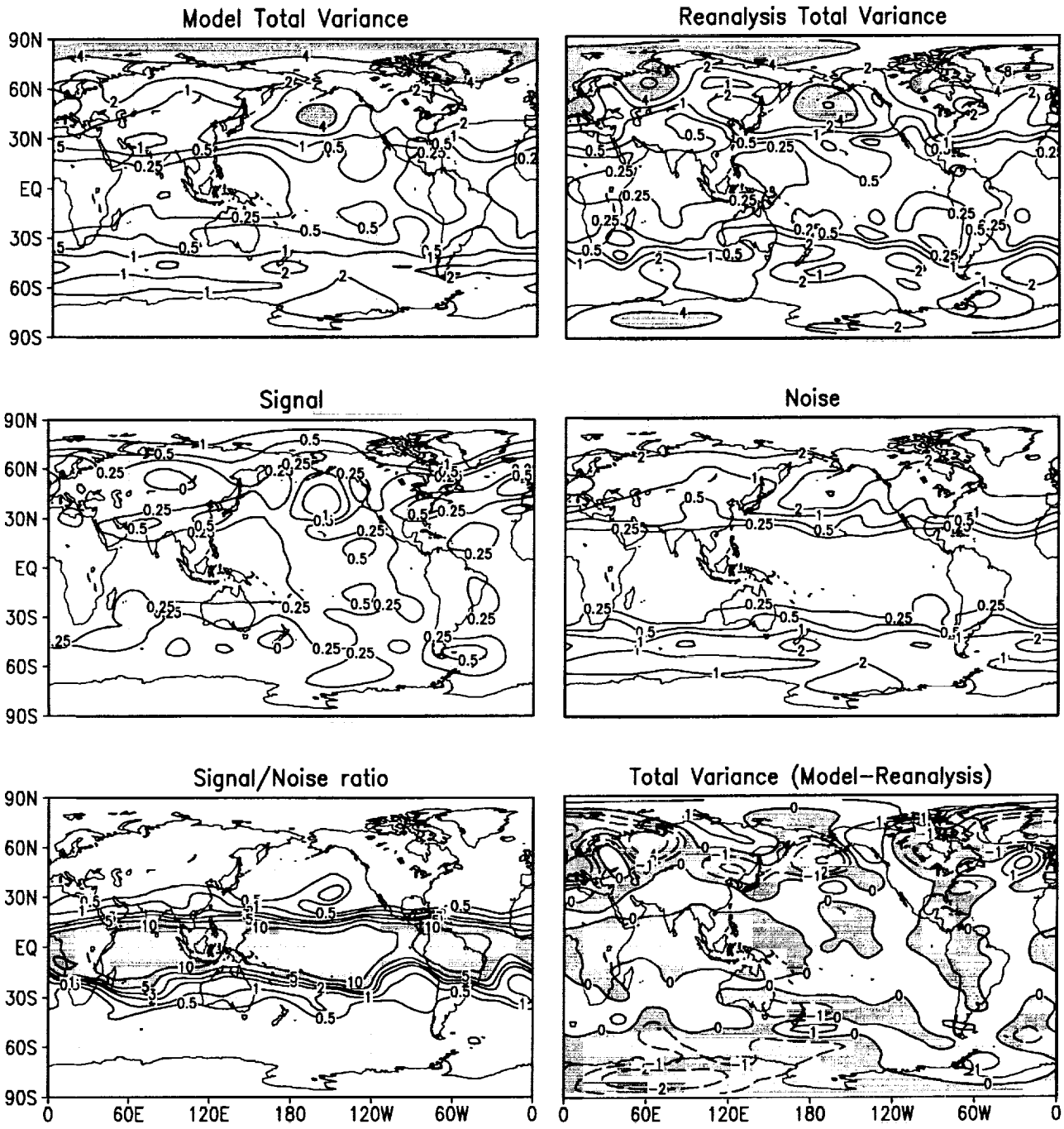

Figure 17: Variance of JFM 200mb geopotential height for non-ENSO years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are .25 .5 1 2 4 8 16 32 $\times 10^3$ m^2, with values above 4000 m^2 shaded. For the difference (bottom-right) the contours are $\pm$ 0 1 2 4 $\times 10^3$ m^2, with negative values shaded. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100 200, with values above 5 shaded.

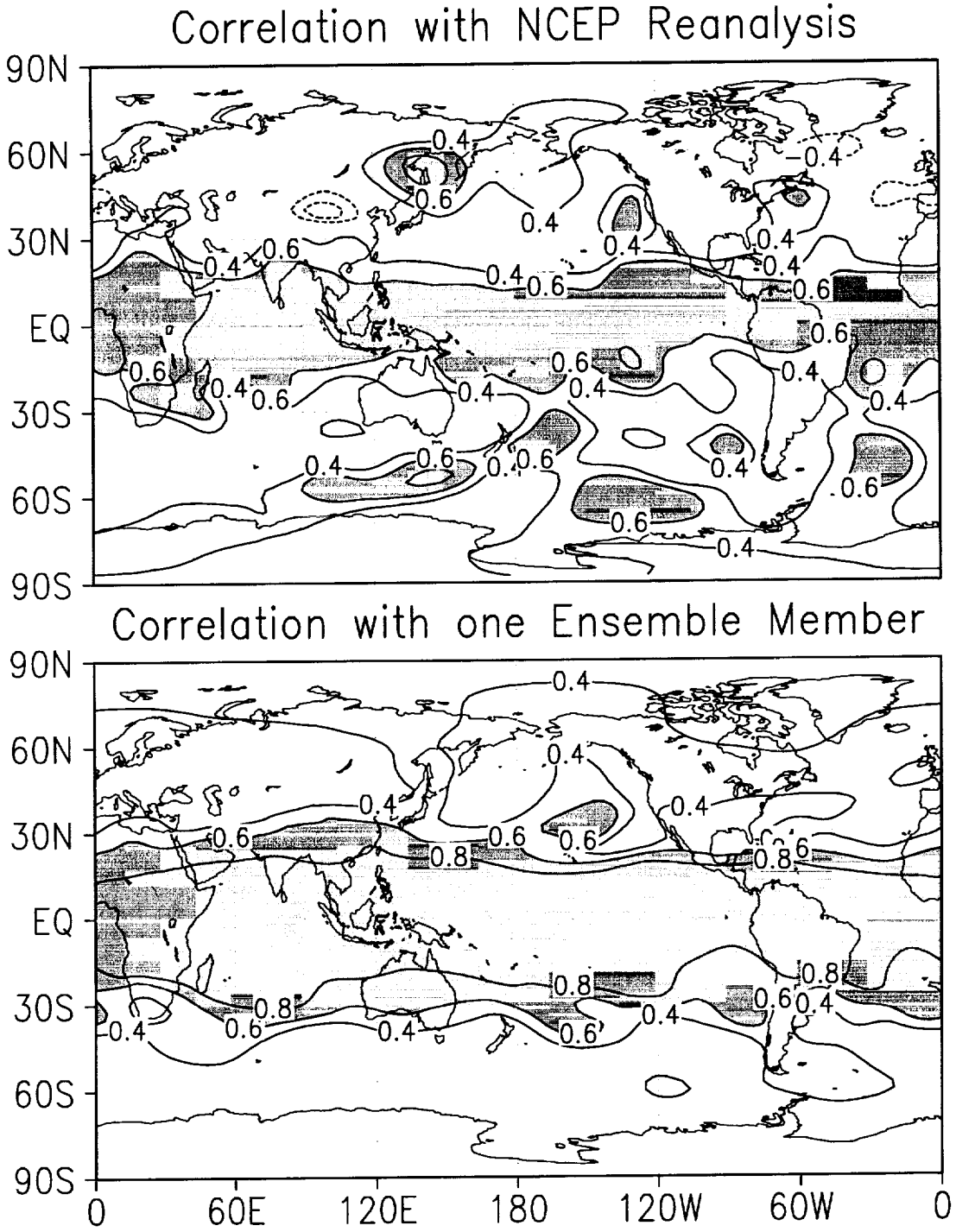

Figure 18: JFM 200mb geopotential height correlations for non-ENSO years. Top: The mean correlation over non-ENSO years of 8-member ensemble means with the observations. Bottom: The mean correlation over non-ENSO years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ± 0.4 are not contoured. Negative contours are dashed.

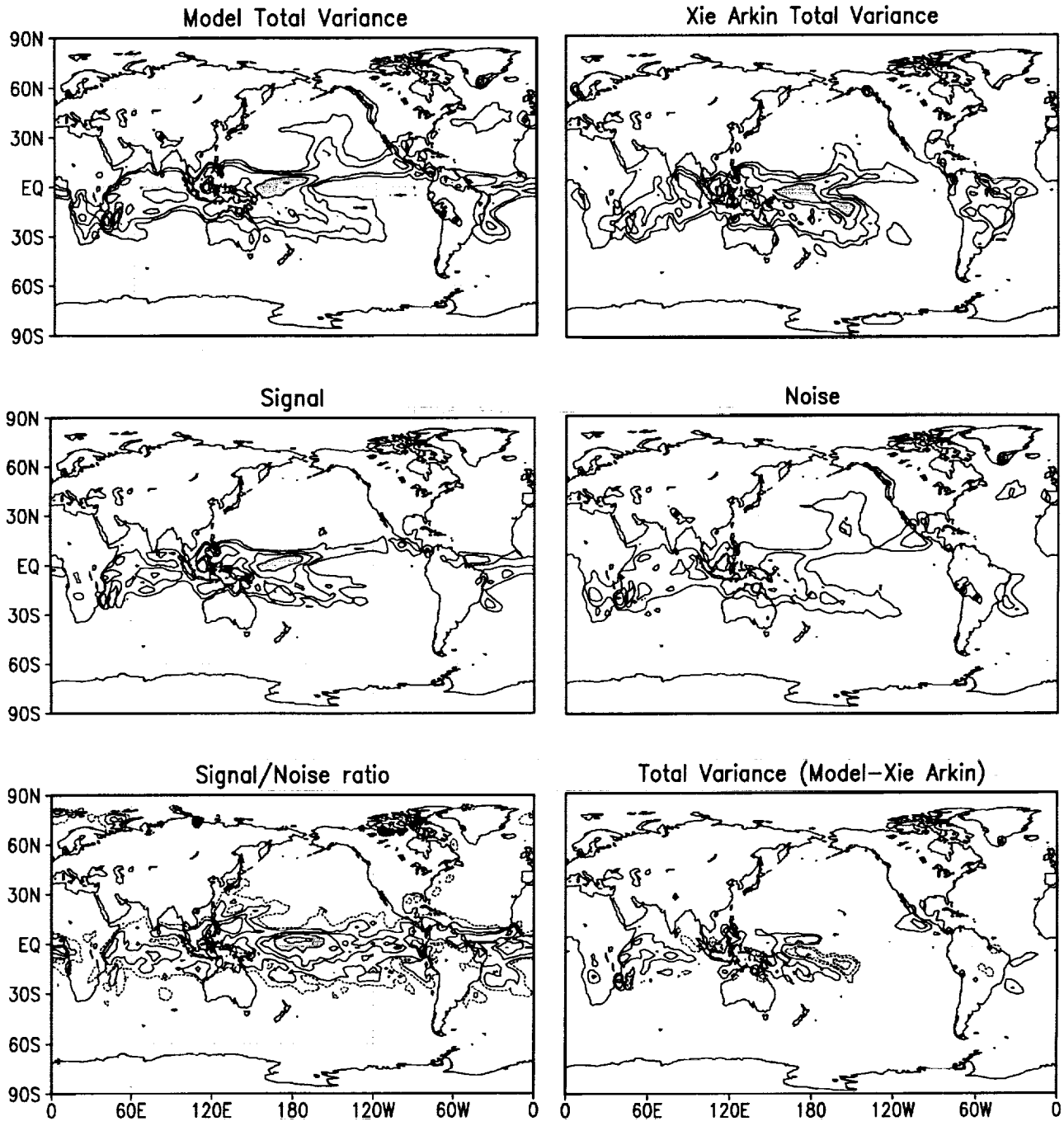

Figure 19: Variance of JFM precipitation for non-ENSO years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are 1 2 4 8 16 32 64 (mm/day)^2; values above 8 are shaded. For the difference (bottom-right) the contours are $\pm$ 2 4 8 16 32 64 (mm/day)^2; negative contours dashed. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100; values above 5 are shaded and the 0.5 contour is dashed.

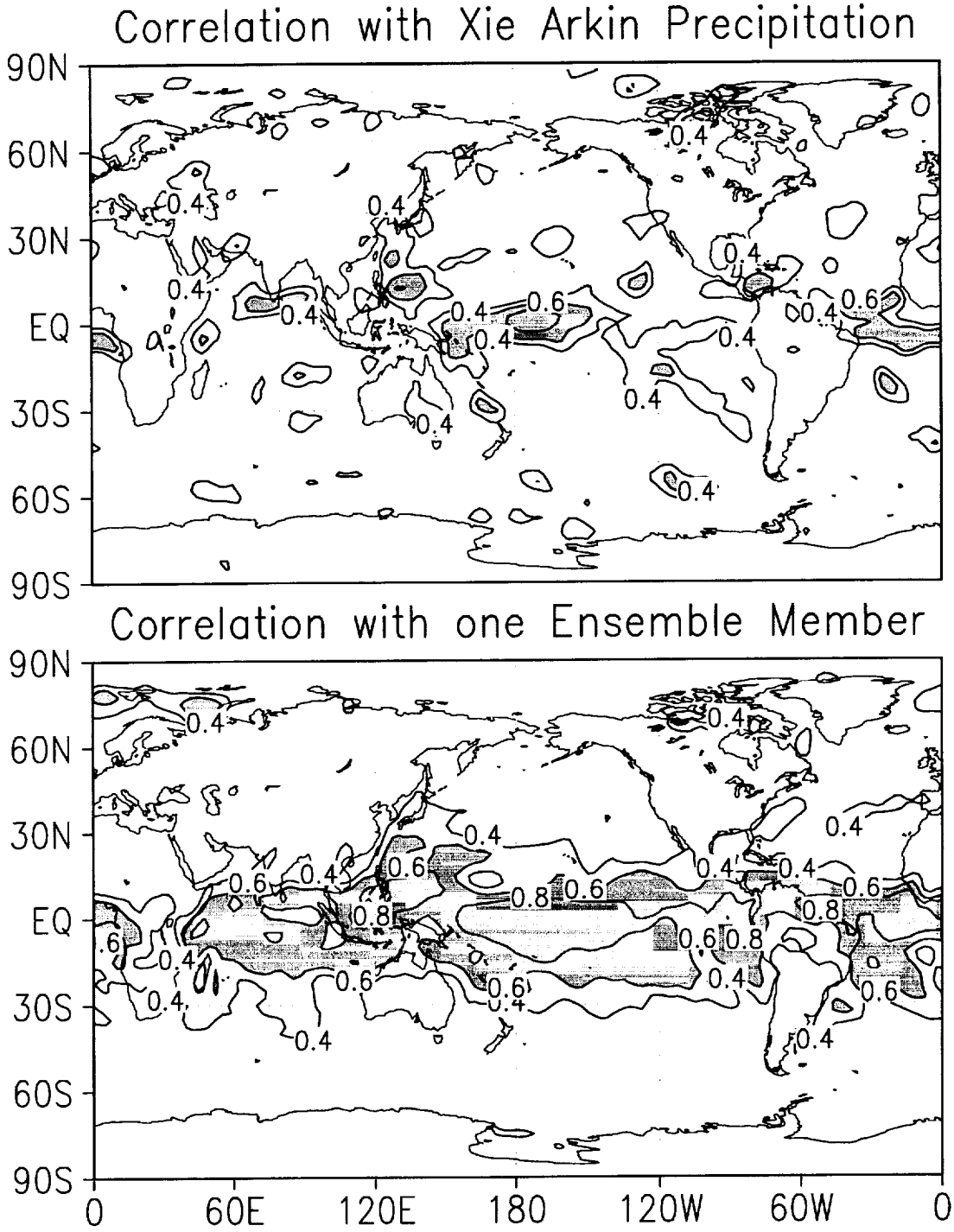

Figure 20: JFM precipitation correlations for non-ENSO years. Top: The mean correlation over non-ENSO years of 8-member ensemble means with the observations. Bottom: The mean correlation over non-ENSO years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ±0.4 are not contoured. Negative contours are dashed.

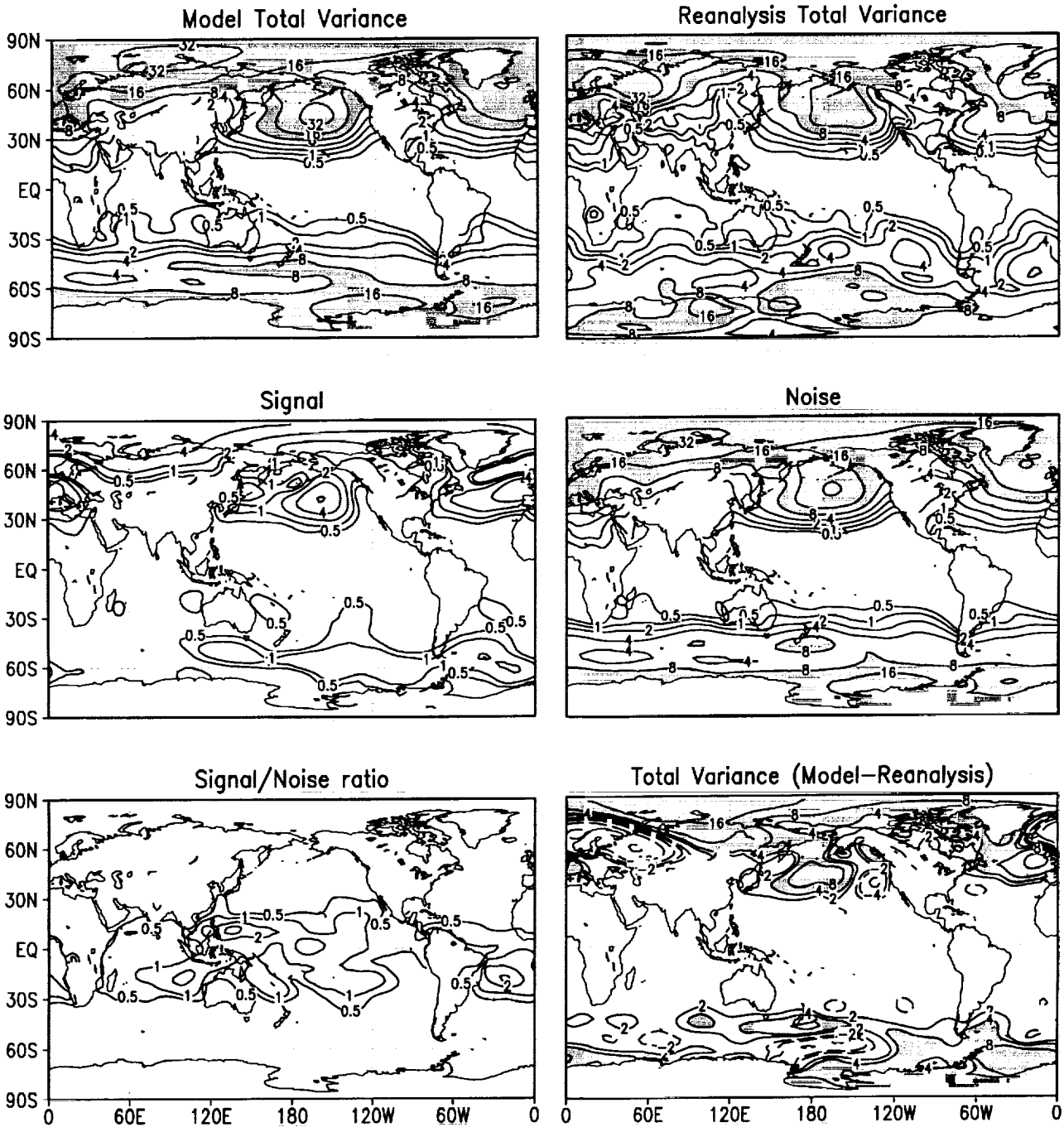

Figure 21: Variance of JFM sea-level pressure for non-ENSO years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. The variances are in units of mb^2. The contours for the ratio are 0.5 1 2 3 5 10 30 50 100 200, with values above 5 shaded. The contours for the difference are $\pm$ 2 4 8 16 mb^2, with values above +4 mb^2 shaded; negative contours are dashed.

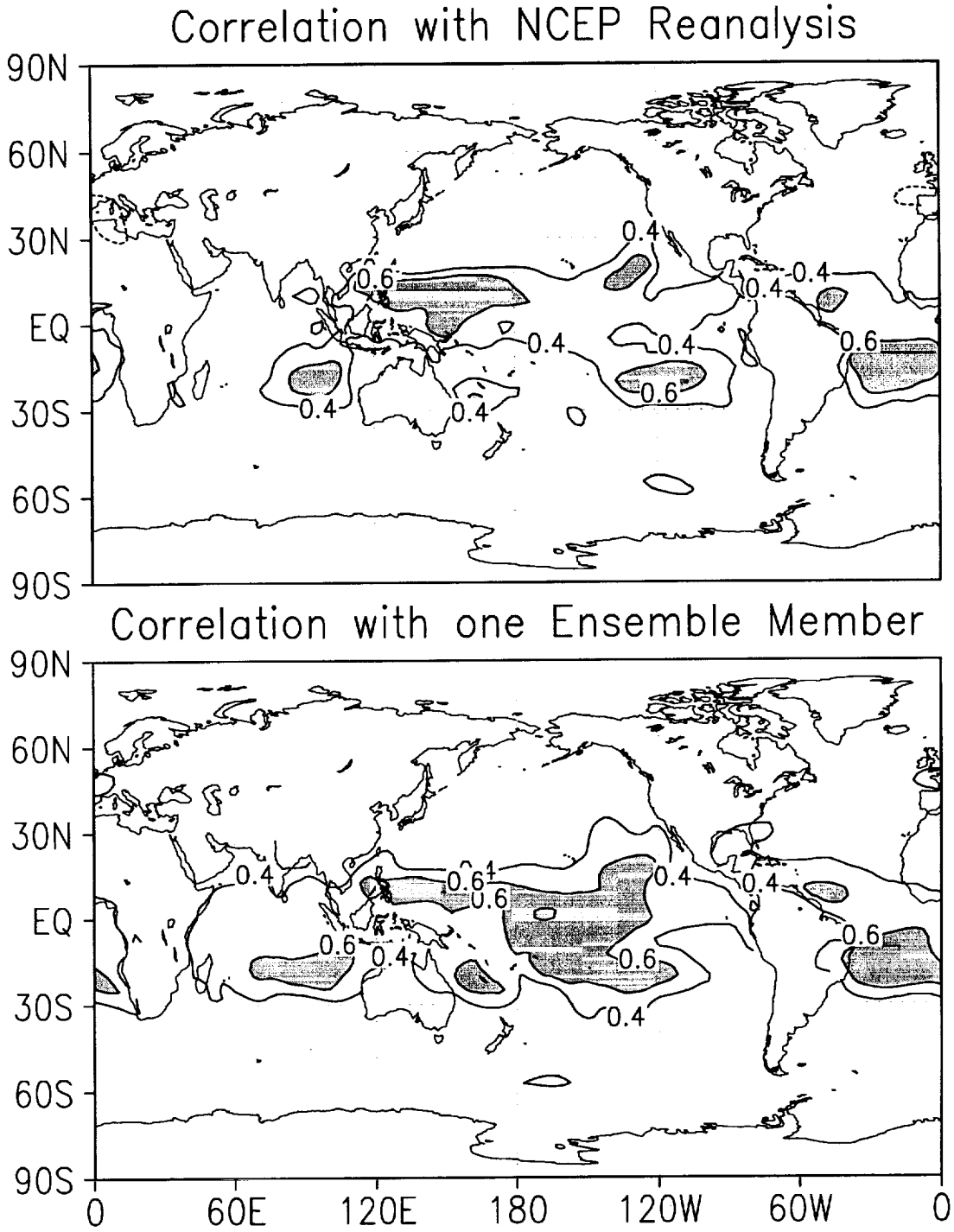

Figure 22: JFM sea-level pressure correlations for non-ENSO years. Top: The mean correlation over non-ENSO years of 8-member ensemble means with the observations. Bottom: The mean correlation over non-ENSO years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ± 0.4 are not contoured. Negative contours are dashed.

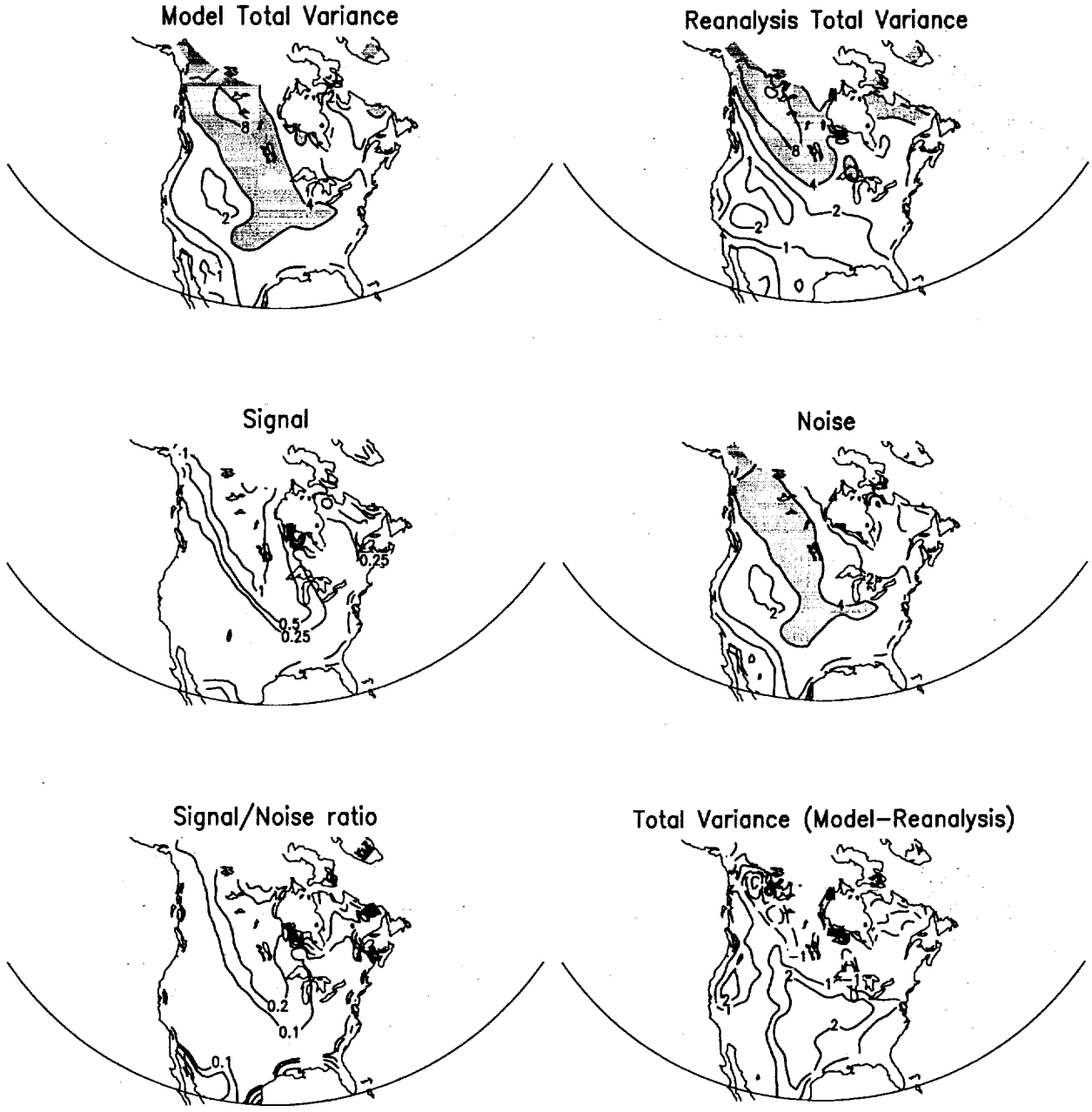

Figure 23: Variance of JFM surface temperature for non-ENSO years. The "signal" is the variance of the ensemble mean. The "noise" is the intra-ensemble variance. For the variances the contours are 0.25 0.5 1 2 4 8 K^2, with values above 4 shaded. The contours for the difference are $\pm$ 1 2 4 8 K^2. The contours for the ratio are 0.1 0.2 0.5 1 2.

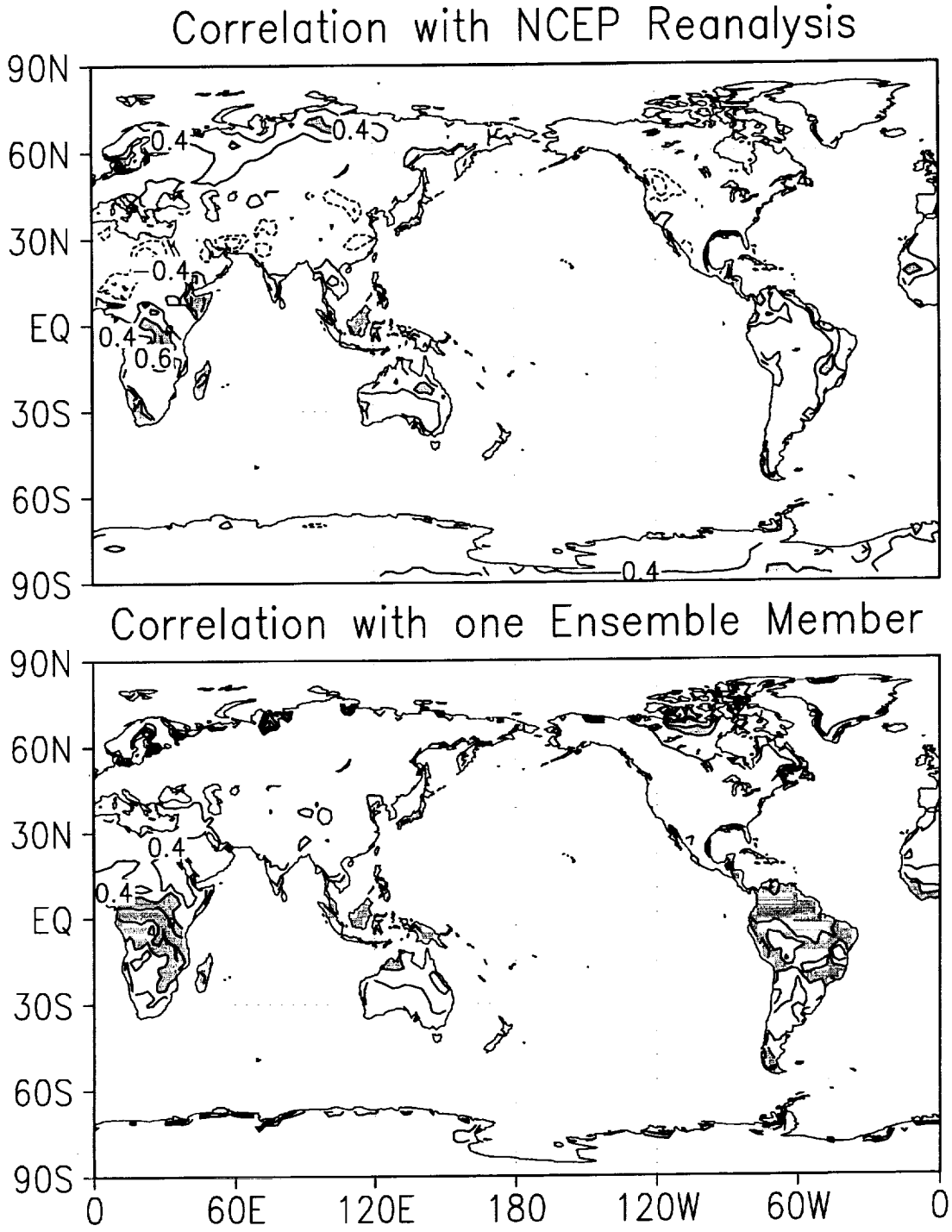

Figure 24: JFM surface temperature correlations for non-ENSO years. Top: The mean correlation over non-ENSO years of 8-member ensemble means with the observations. Bottom: The mean correlation over non-ENSO years between 8-member ensemble means and the ninth member. Values above 0.6 are shaded, values within ± 0.4 are not contoured. Negative contours are dashed.

MODEL COMPARISON

Anomaly correlations of Z_{500} in then Pacific–North American Sector

Total Variance of 200 mb Height in then Pacific–North American Sector

"Signal-to-Noise" Ratio of 200 mb Height in Pacific–North American Sector

"Signal" of 200 mb Height in the Pacific–North American Sector

"Noise" of 200 mb Height in the Pacific–North American Sector

Total Variance of Precipitation in the Tropical Pacific

"Signal-to-Noise" Ratio of Precipitation in the tropical Pacific

"Signal" Precipitation in the tropical Pacific

"Noise" Precipitation in the tropical Pacific

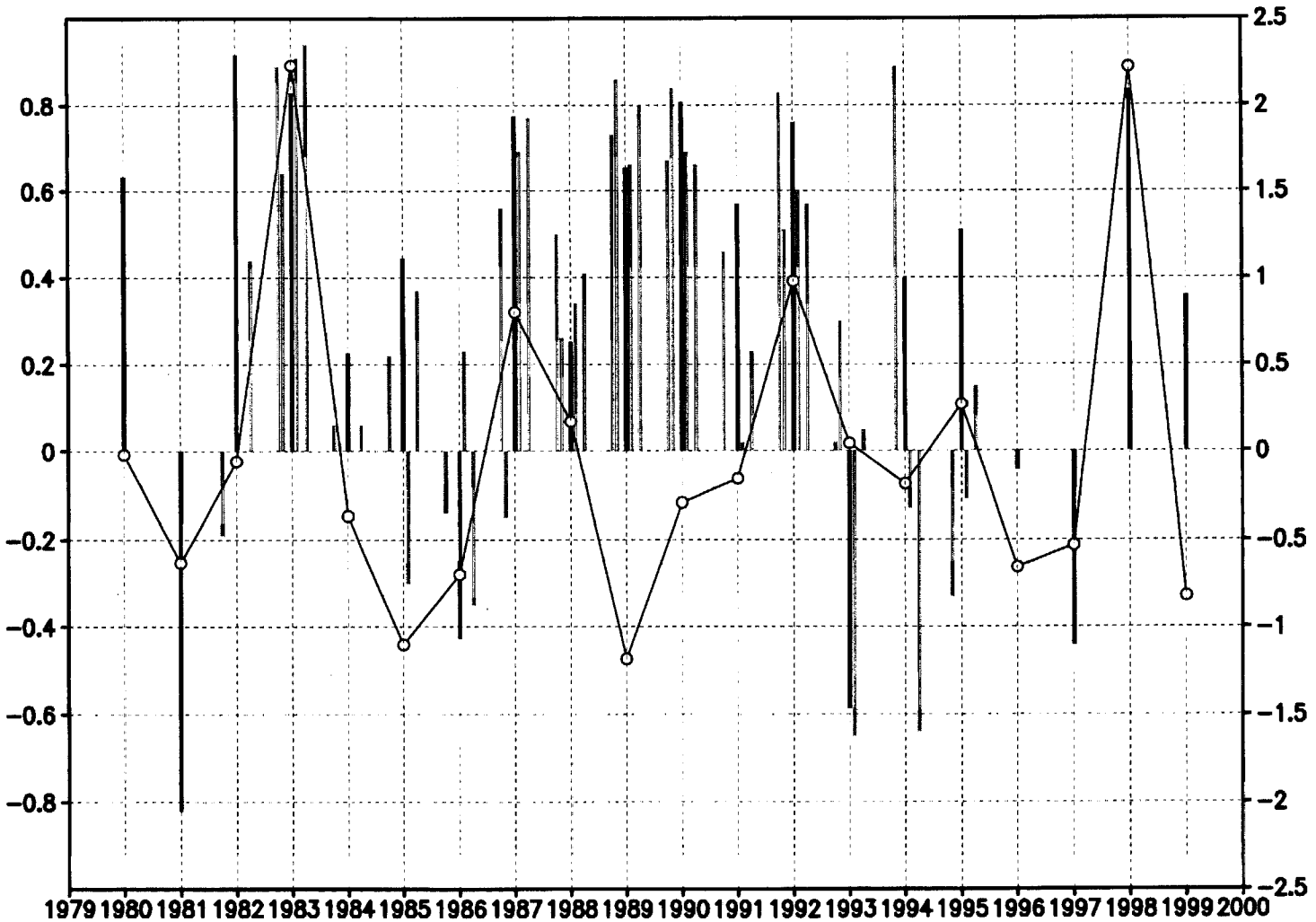

Figure 25: Spatial anomaly correlation of 500 mb height over the Pacific–North American sector (25N–70N and 60W–150W). Black bars are for the NSIPP-1 AGCM. Grey bars are for other models that participated in the DSP project. Note that for some models results are not available for all years. The open circles show the observed JFM Niño-3 SST anomalies (right ordinate, °C).

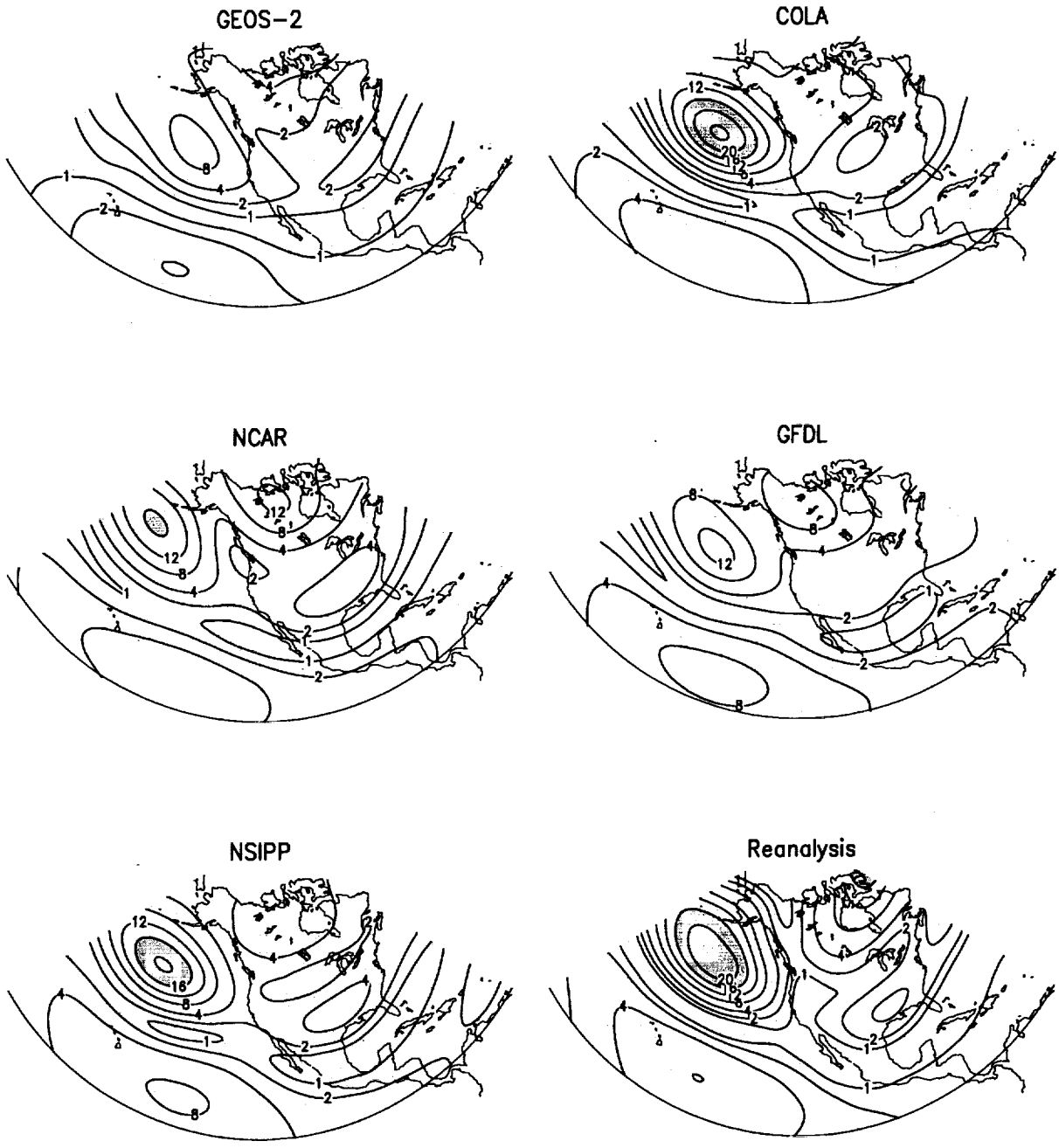

Figure 26: Total Variance of JFM 200 mb geopotential height in the Pacific–North American sector for the four models that participated in the DSP comparison, for the NSIPP-1 GCM, and for the reanalysis. The variance is computed only over the ENSO years (1983, 1985, 1987, 1989, 1992) for which results were available from all models. The contour levels are 1 2 4 8 12 16 20 24 $\times 10^3$ m^2. Values greater than 1600 m^2 are shaded.

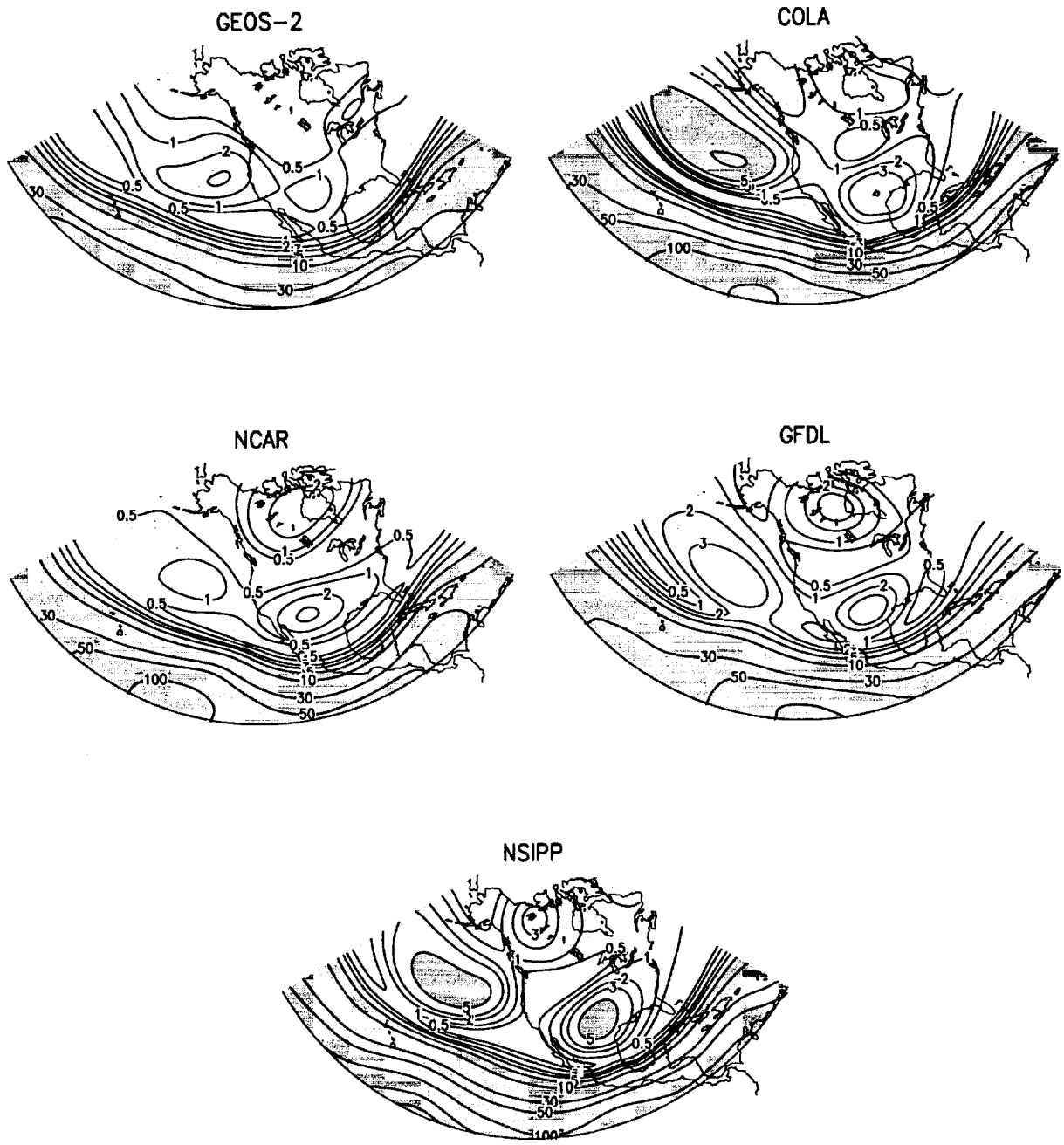

Figure 27: "Signal-to-Noise" ratio of the JFM 200 mb geopotential height in the Pacific–North American sector as in figure 26. Contour levels are 0.5 1 2 3 5 10 30 50 100 200. Values greater than 5 are shaded

Figure 28: "Signal" of the JFM 200 mb geopotential height in the Pacific–North American sector as in figure 26. The contour levels are 1 2 4 8 12 16 20 24 $\times 10^3$ m^2. Values greater than 1600 m^2 are shaded

Figure 29: "Noise"of the JFM 200 mb geopotential height in the Pacific–North American sector as in figure 26. The contour levels are 1 2 4 8 12 16 20 24 $\times 10^3$ m^2. Values greater than 1600 m^2 are shaded.

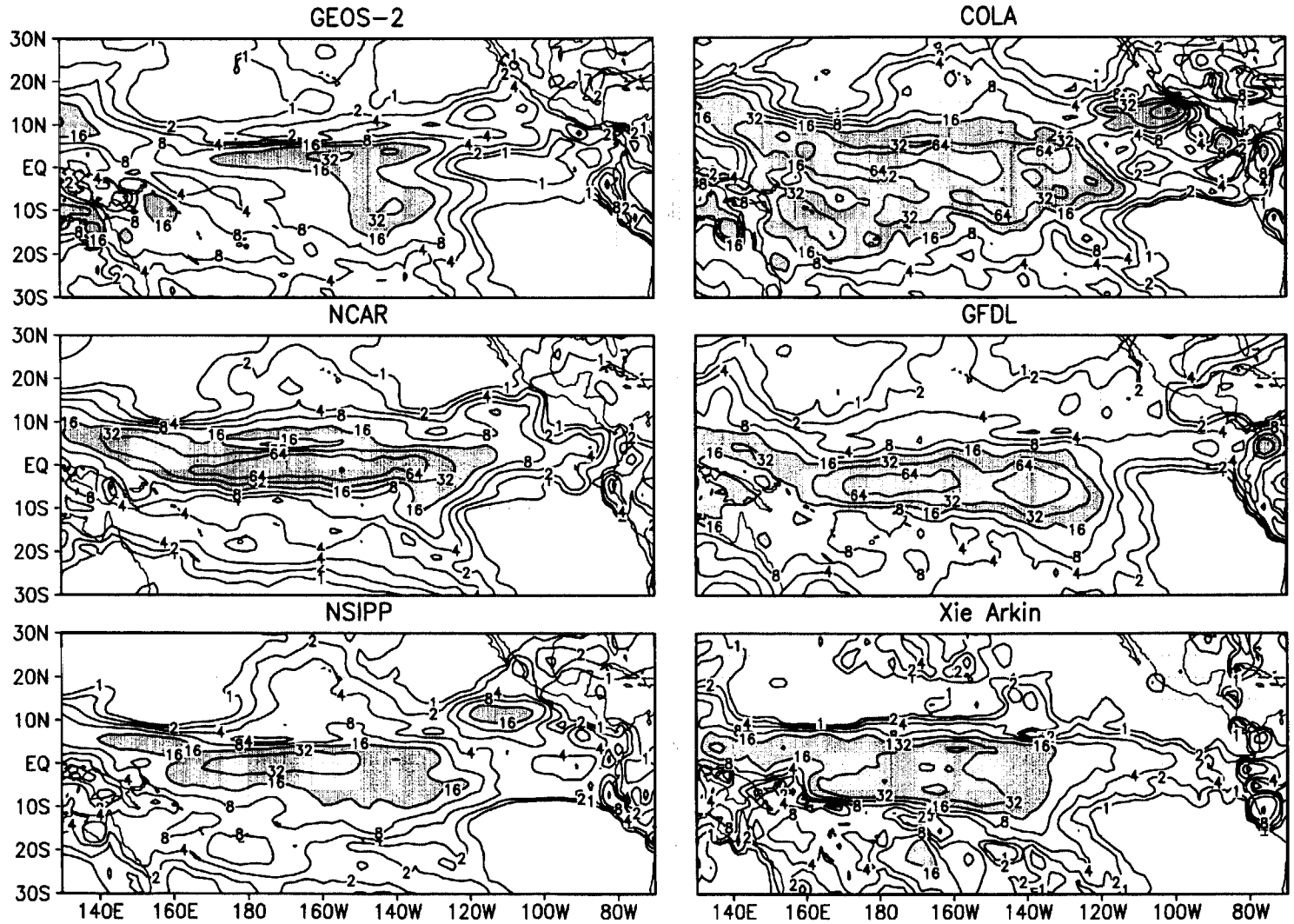

Figure 30: Total variance of precipitation in the tropical Pacific as in figure 26. Contour levels are 1 2 4 8 16 32 64 128 $(\text{mm/day})^2$. Values greater than 16 $(\text{mm/day})^2$ are shaded.

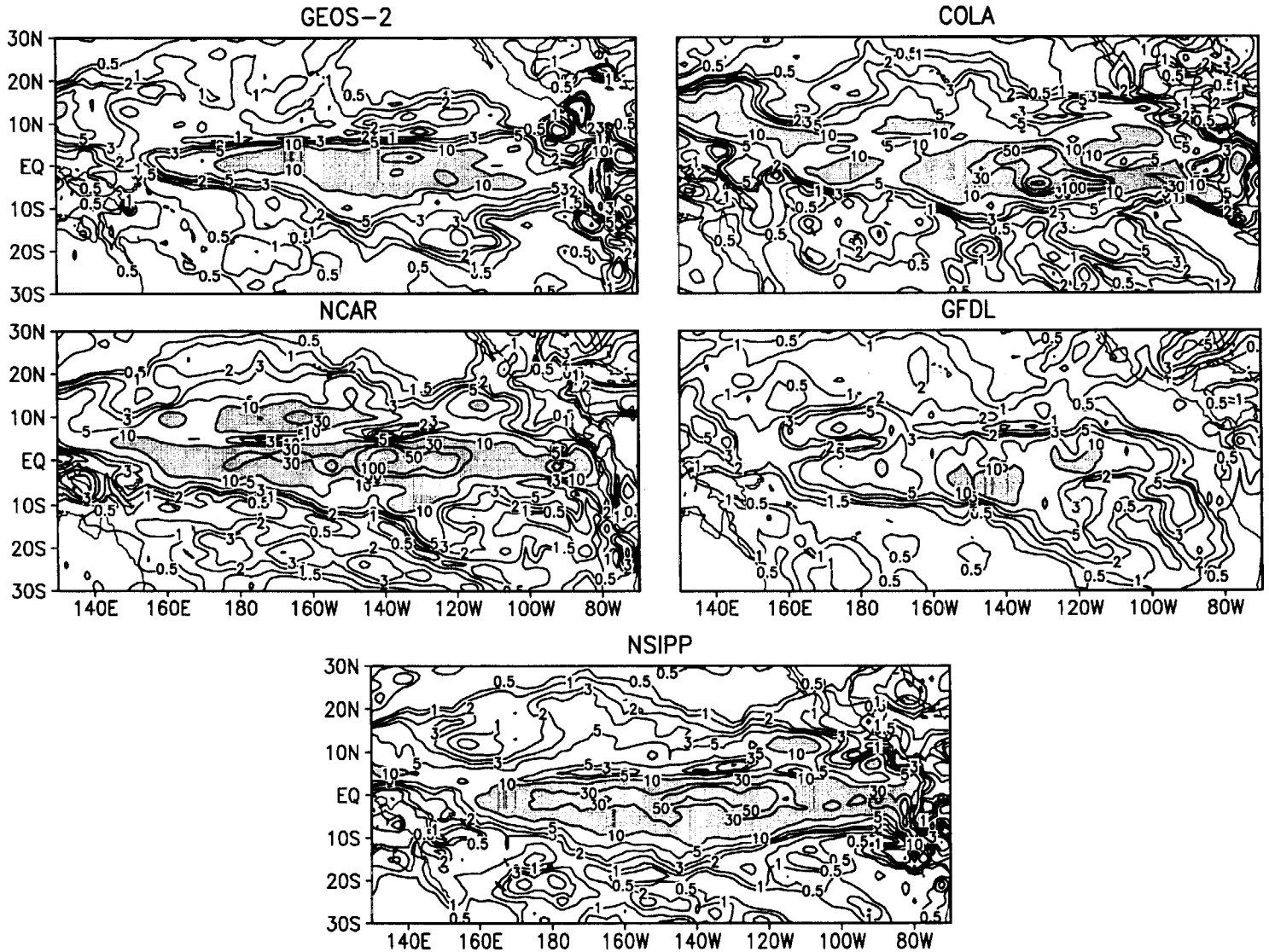

Figure 31: "Signal-to-Noise" ratio of precipitation in the tropical Pacific as in figure 26. Contour levels are 0.5 1 2 3 5 10 30 50 100 150. Values greater than 10 are shaded.

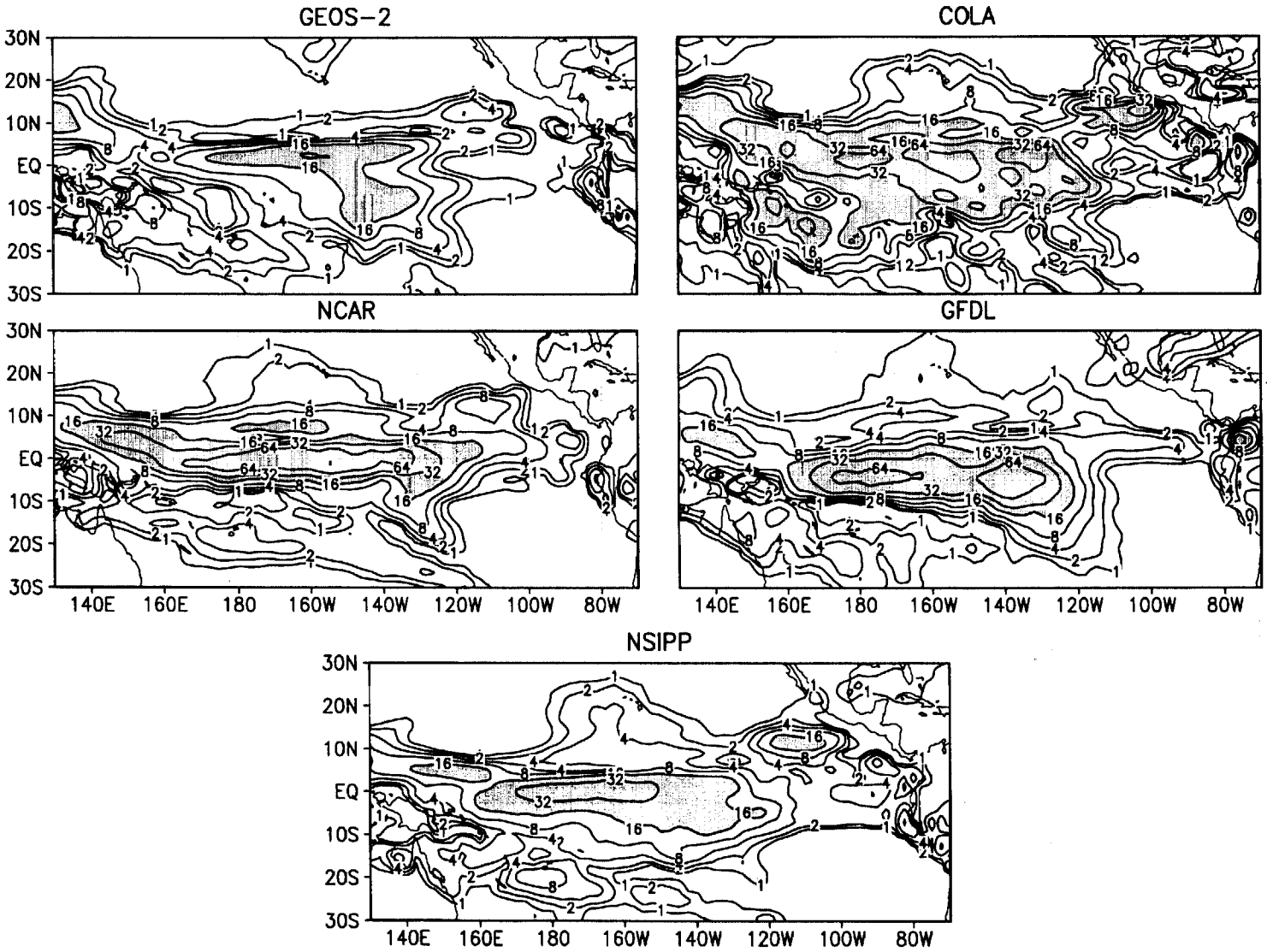

Figure 32: "Signal" of precipitation in the tropical Pacific as in figure 26. Contour levels are 1 2 4 8 16 32 64 128 (mm/day)^2. Values greater than 16 (mm/day)^2 are shaded.

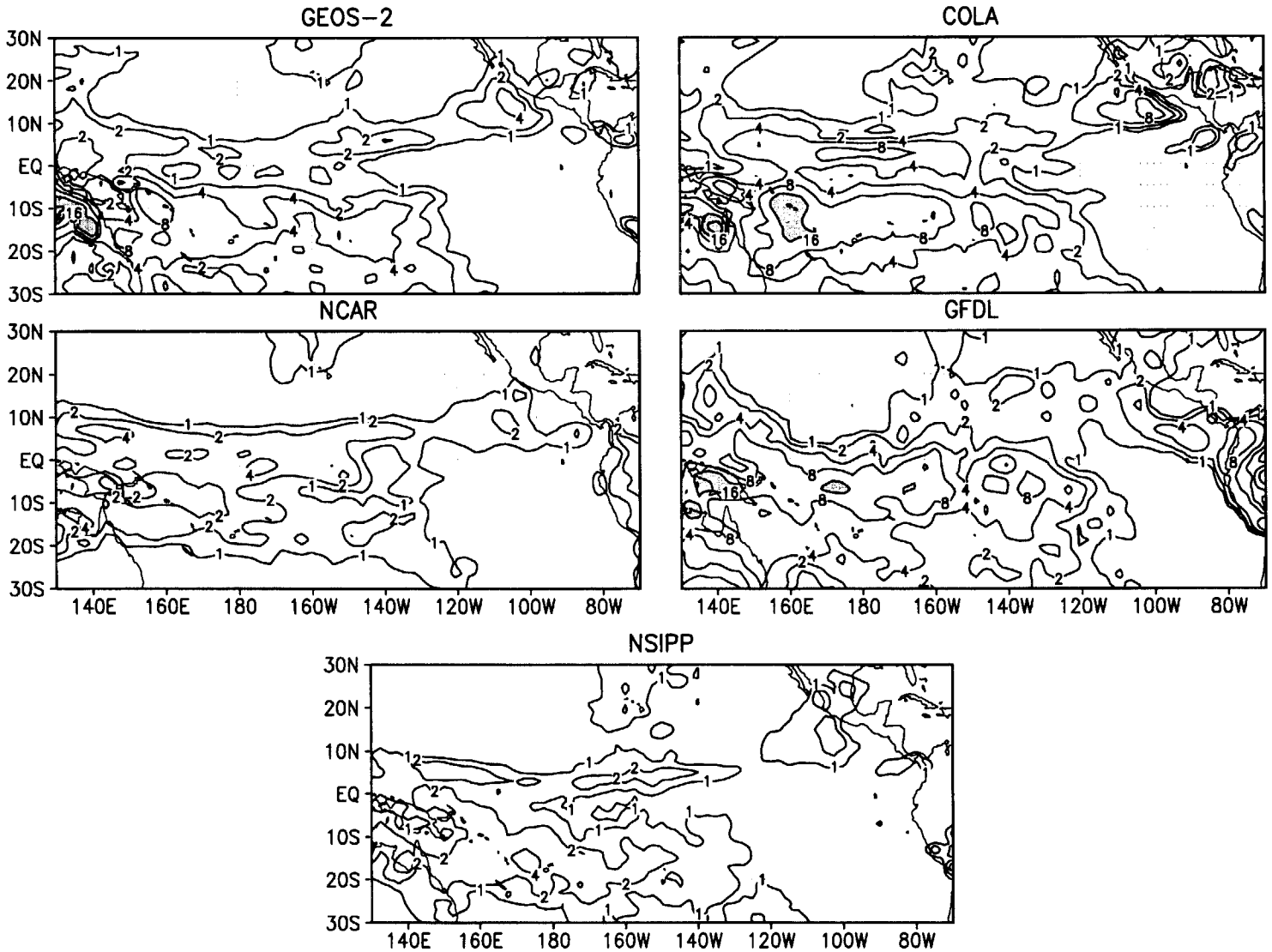

Figure 33: "Noise" of precipitation in the tropical Pacific as in figure 26. Contour levels are 1 2 4 8 16 32 64 128 (mm/day)^2. Values greater than 16 (mm/day)^2 are shaded.

INDIVIDUAL YEARS

(1980–1999)

200 mb Height

Precipitation

Surface Temperature over North America

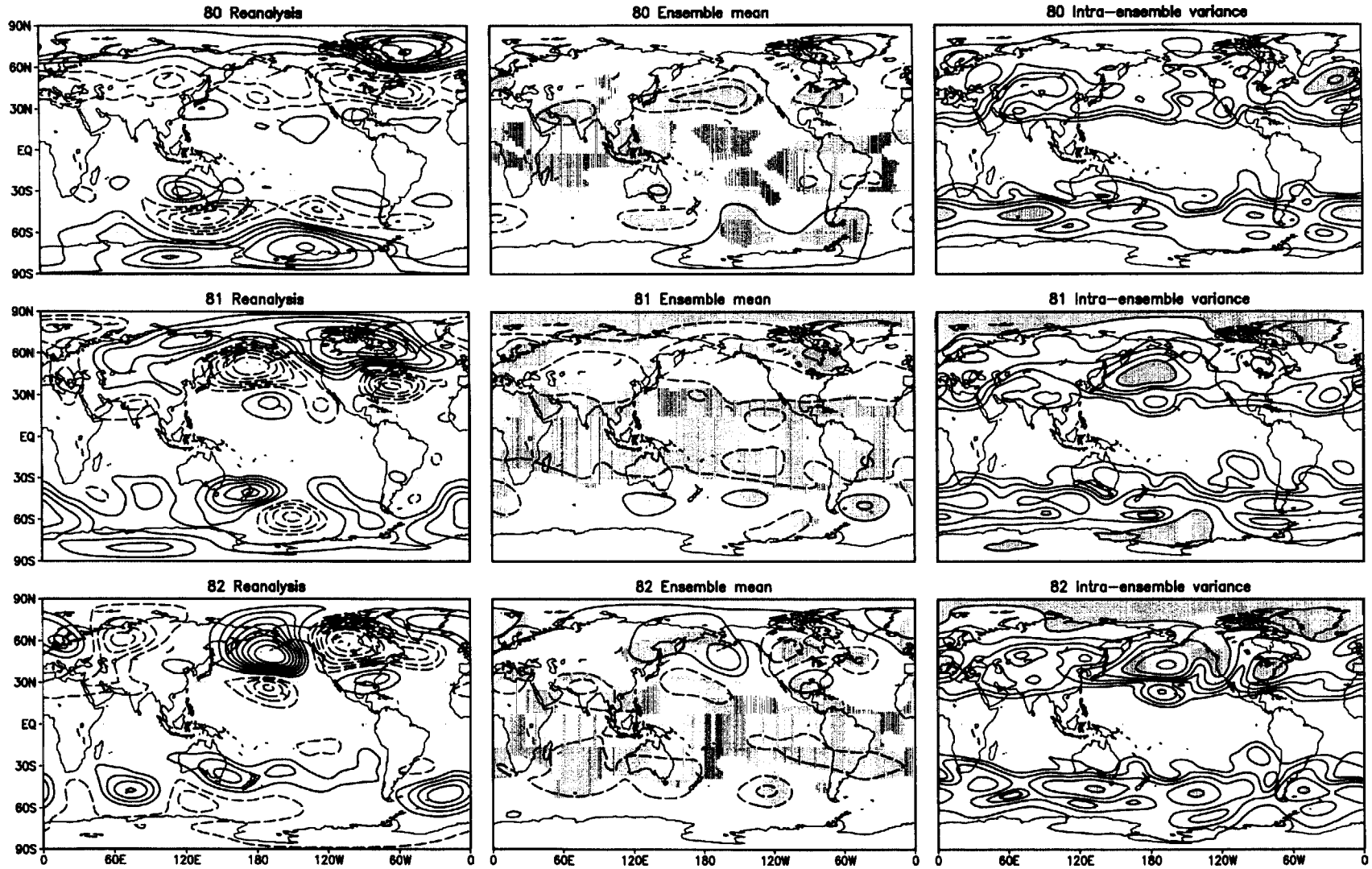

Figure 34: JFM geopotential height at 200 mb for 1980, 1981, and 1982. Left panels are the anomalies from the reanalysis (contour interval: 20 m); center panels are the ensemble-mean anomalies for the model (contour interval: 20 m); and right panels are the model's intra-ensemble variance (contour levels: 2 4 8 16 32 $\times 10^2$ m^2). For the ensemble mean, shading indicates values significant at the two-sided 5% level, the zero contour is omitted, and negative values are dashed. Variance values exceeding 3200 m^2 are shaded.

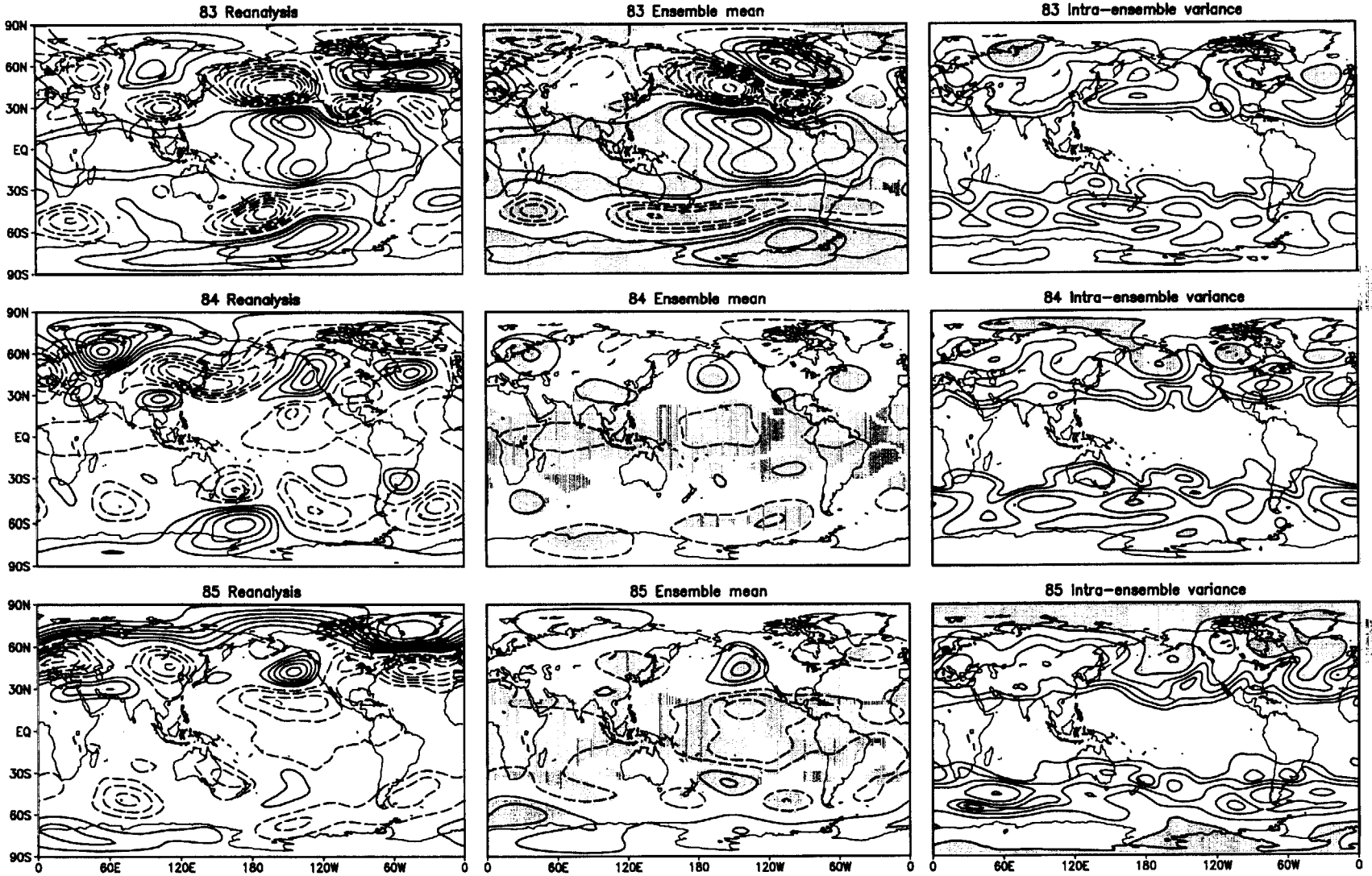

Figure 35: JFM geopotential height at 200 mb for 1983, 1984, and 1985. Left panels are the anomalies from the reanalysis (contour interval: 20 m); center panels are the ensemble-mean anomalies for the model (contour interval: 20 m); and right panels are the model's intra-ensemble variance (contour levels: 2 4 8 16 32 $\times 10^2$ m^2). For the ensemble mean, shading indicates values significant at the two-sided 5% level, the zero contour is omitted, and negative values are dashed. Variance values exceeding 3200 m^2 are shaded.

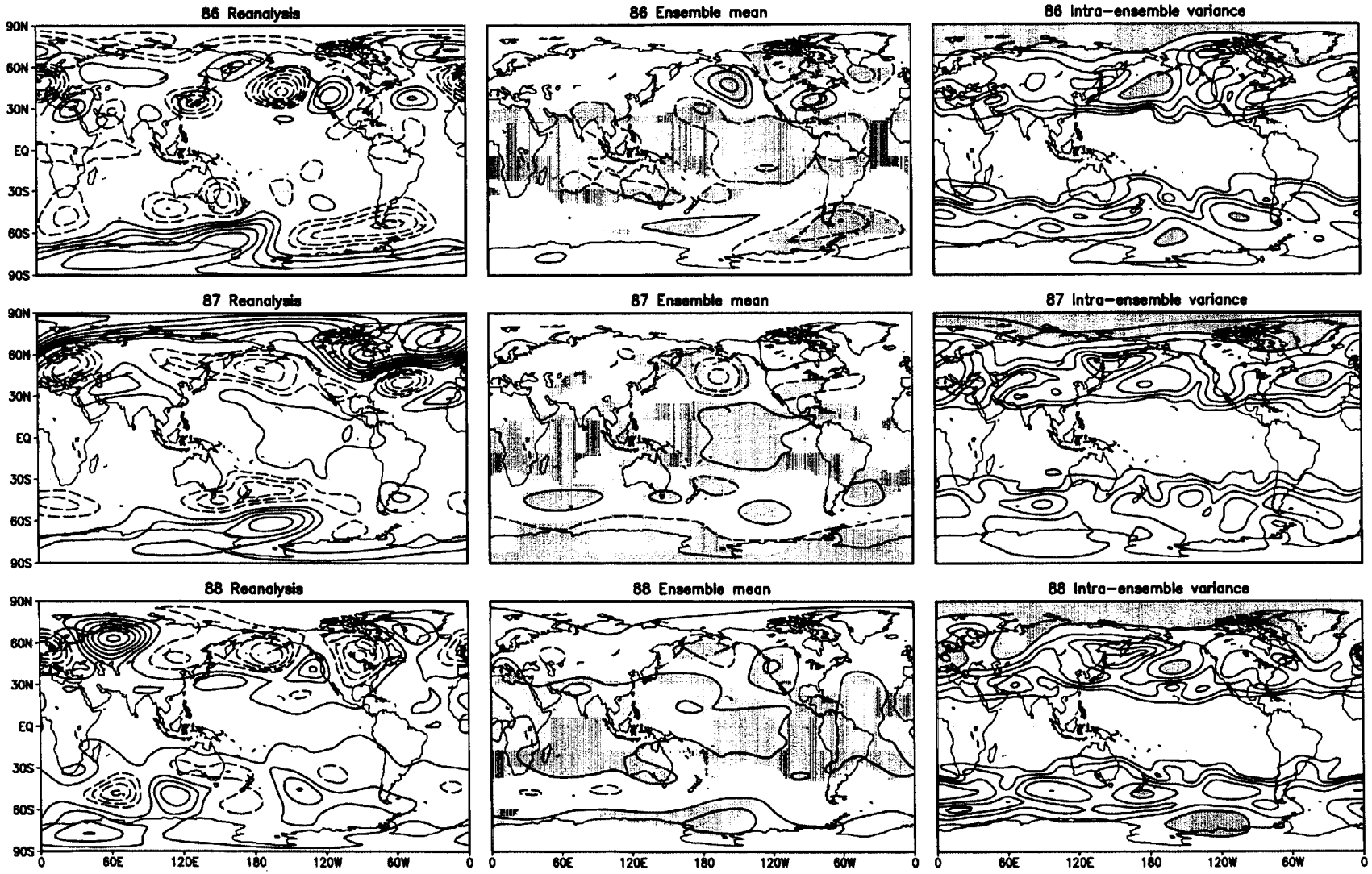

Figure 36: JFM geopotential height at 200 mb for 1986, 1987, and 1988. Left panels are the anomalies from the reanalysis (contour interval: 20 m); center panels are the ensemble-mean anomalies for the model (contour interval: 20 m); and right panels are the model's intra-ensemble variance (contour levels: 2 4 8 16 32 $\times 10^2$ m^2). For the ensemble mean, shading indicates values significant at the two-sided 5% level, the zero contour is omitted, and negative values are dashed. Variance values exceeding 3200 m^2 are shaded.

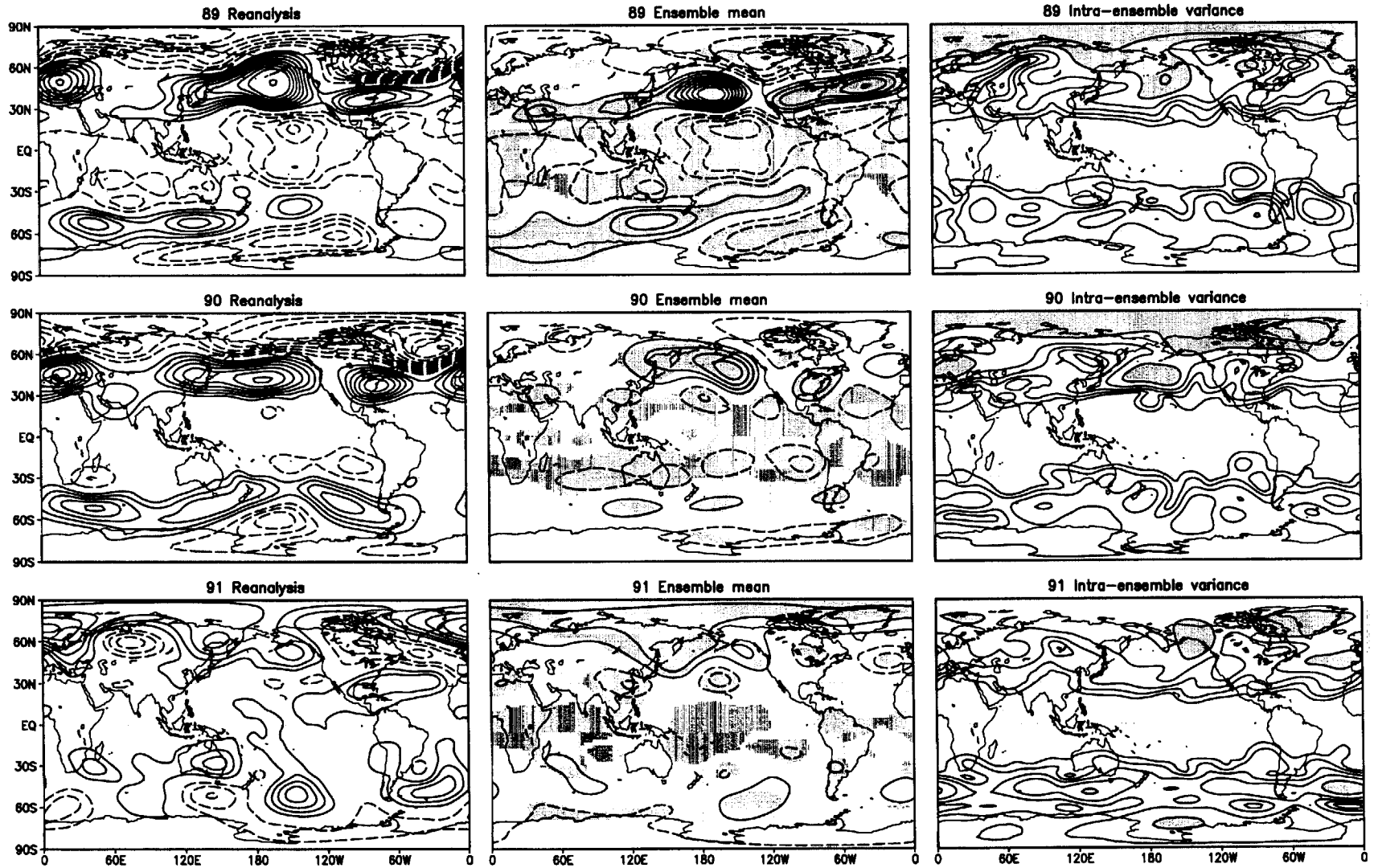

Figure 37: JFM geopotential height at 200 mb for 1989, 1990, and 1991. Left panels are the anomalies from the reanalysis (contour interval: 20 m); center panels are the ensemble-mean anomalies for the model (contour interval: 20 m); and right panels are the model's intra-ensemble variance (contour levels: 2 4 8 16 32 $\times 10^2$ m^2). For the ensemble mean, shading indicates values significant at the two-sided 5% level, the zero contour is omitted, and negative values are dashed. Variance values exceeding 3200 m^2 are shaded.

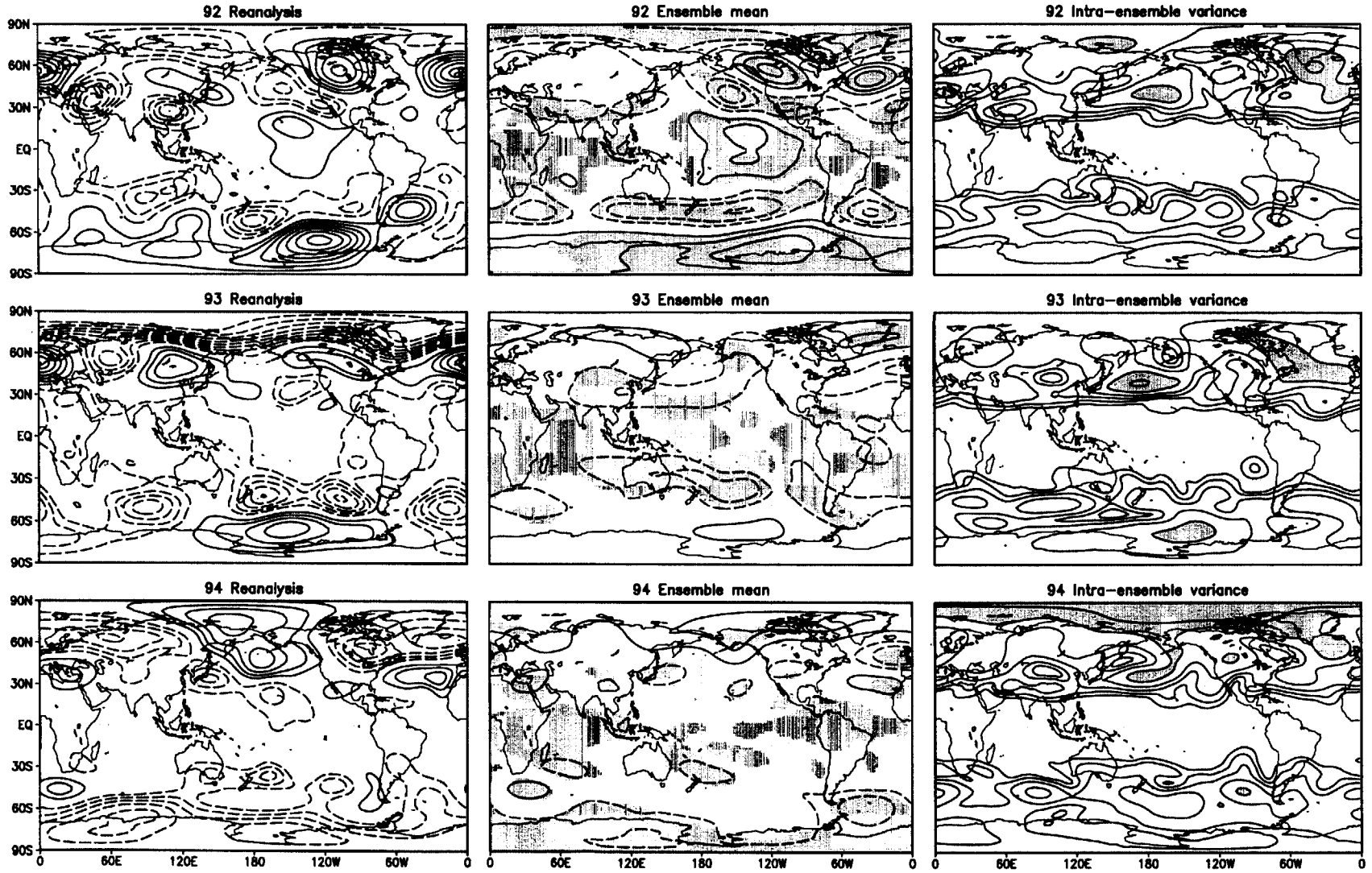

Figure 38: JFM geopotential height at 200 mb for 1992, 1993, and 1994. Left panels are the anomalies from the reanalysis (contour interval: 20 m); center panels are the ensemble-mean anomalies for the model (contour interval: 20 m); and right panels are the model's intra-ensemble variance (contour levels: 2 4 8 16 32 $\times 10^2$ m^2). For the ensemble mean, shading indicates values significant at the two-sided 5% level, the zero contour is omitted, and negative values are dashed. Variance values exceeding 3200 m^2 are shaded.

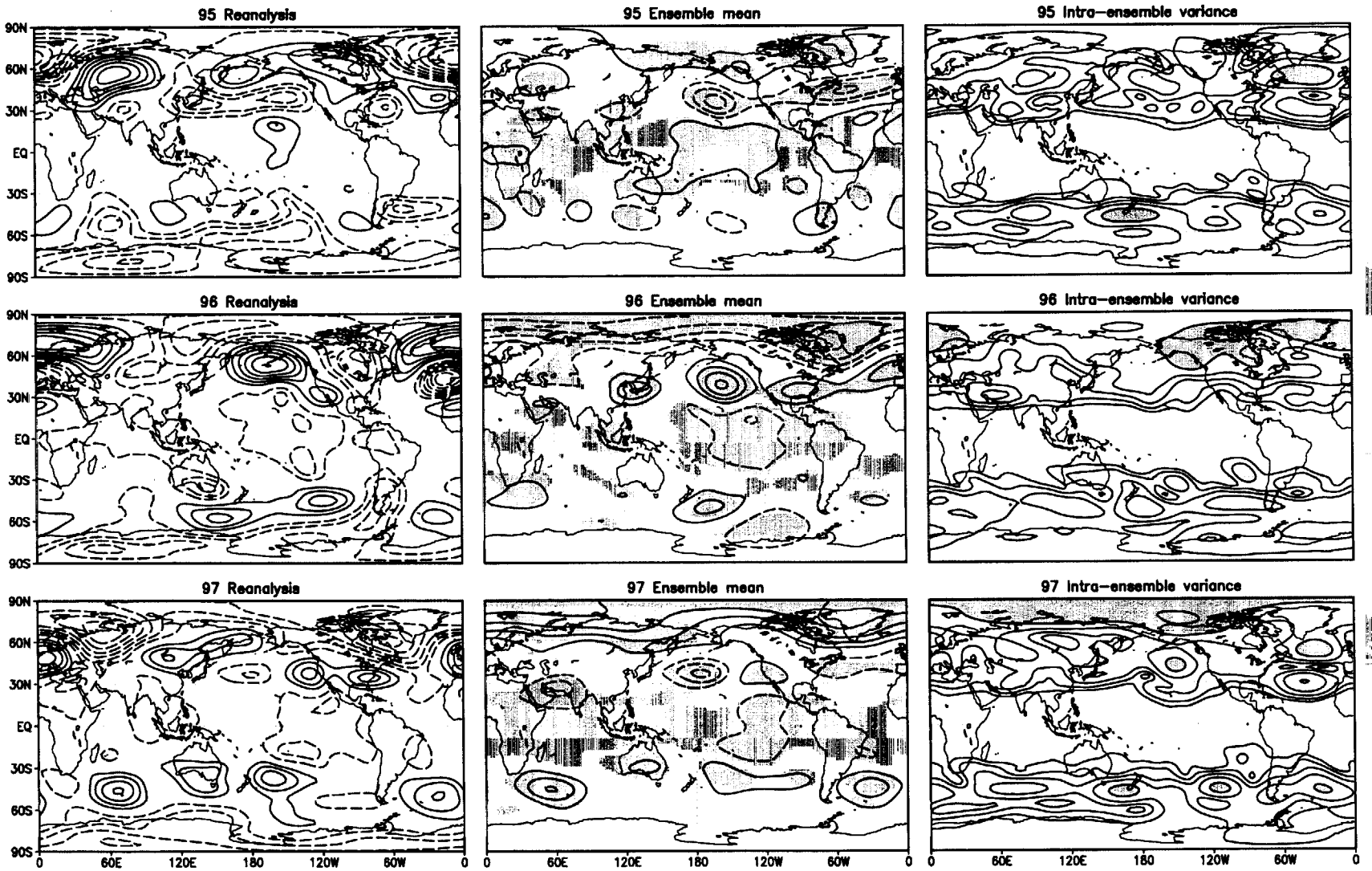

Figure 39: JFM geopotential height at 200 mb for 1995, 1996, and 1997. Left panels are the anomalies from the reanalysis (contour interval: 20 m); center panels are the ensemble-mean anomalies for the model (contour interval: 20 m); and right panels are the model's intra-ensemble variance (contour levels: 2 4 8 16 32 $\times 10^2$ m^2). For the ensemble mean, shading indicates values significant at the two-sided 5% level, the zero contour is omitted, and negative values are dashed. Variance values exceeding 3200 m^2 are shaded.

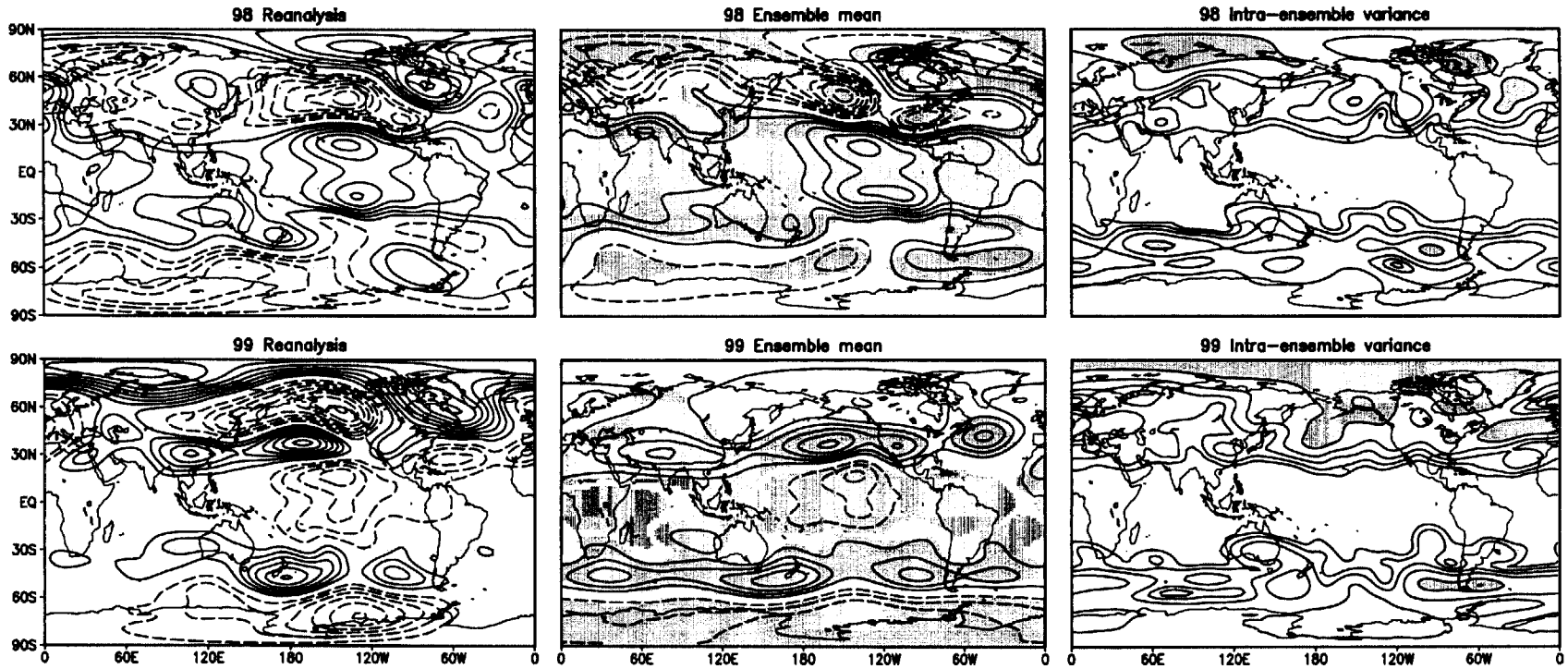

Figure 40: JFM geopotential height at 200 mb for 1998 and 1999. Left panels are the anomalies from the reanalysis (contour interval: 20 m); center panels are the ensemble-mean anomalies for the model (contour interval: 20 m); and right panels are the model's intra-ensemble variance (contour levels: 2 4 8 16 32 $\times 10^2$ m^2). For the ensemble mean, shading indicates values significant at the two-sided 5% level, the zero contour is omitted, and negative values are dashed. Variance values exceeding 3200 m^2 are shaded.

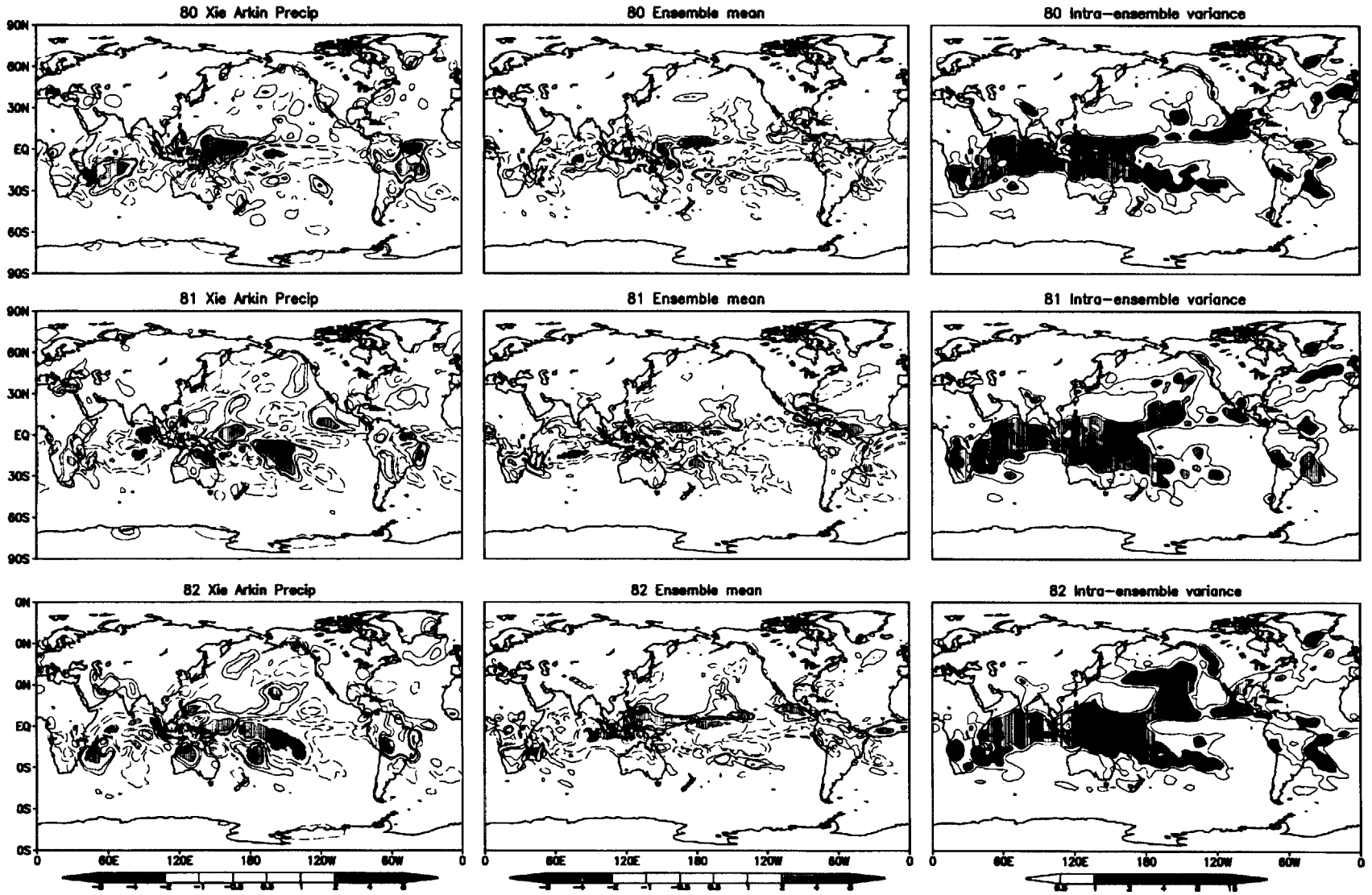

Figure 41: JFM precipitation for 1980, 1981, and 1982. Left panels are the anomalies from the Xie-Arkin data; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. The units of precipitation are mm/day.

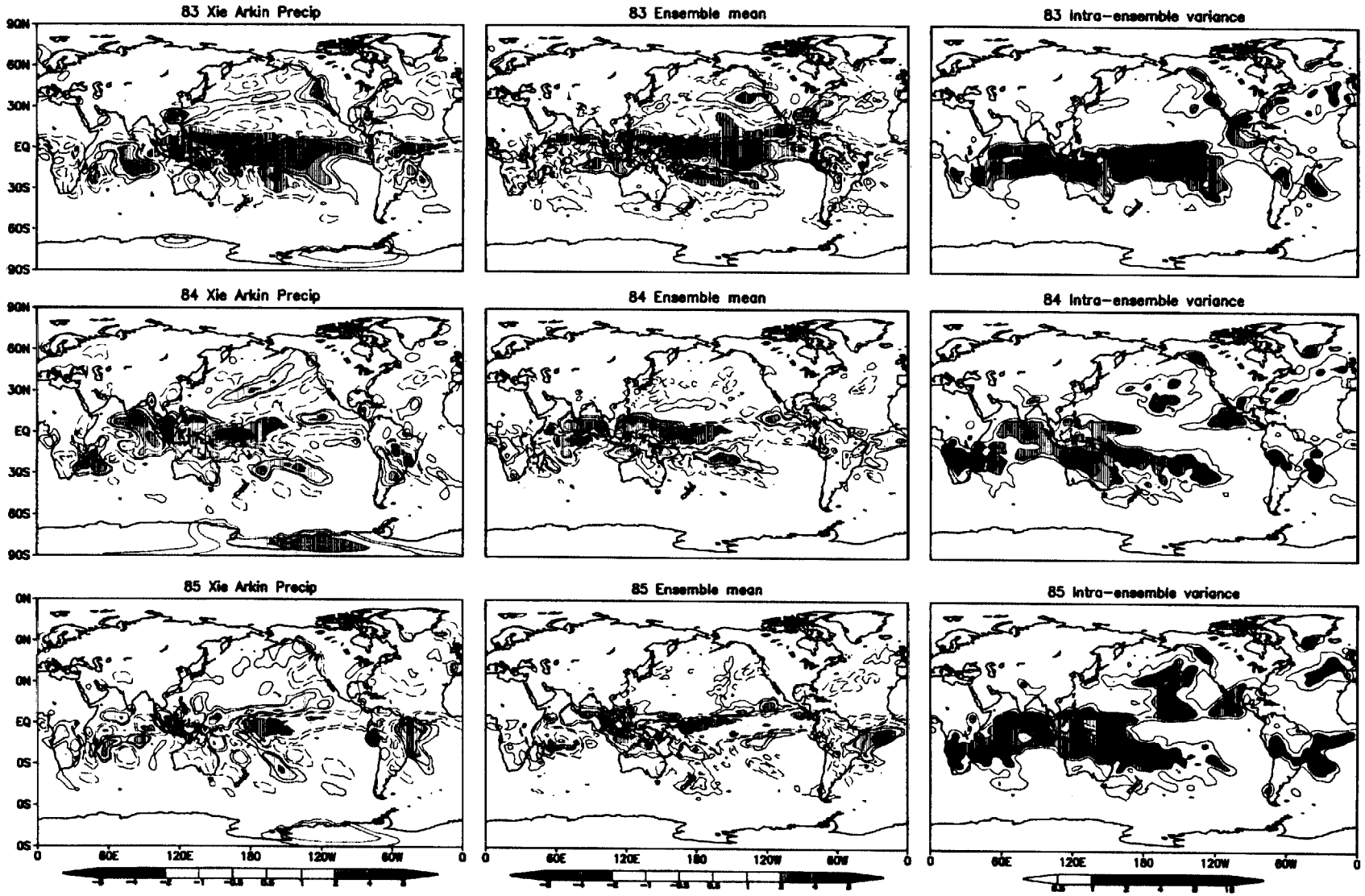

Figure 42: JFM precipitation for 1983, 1984, and 1985. Left panels are the anomalies from the Xie-Arkin data; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. The units of precipitation are mm/day.

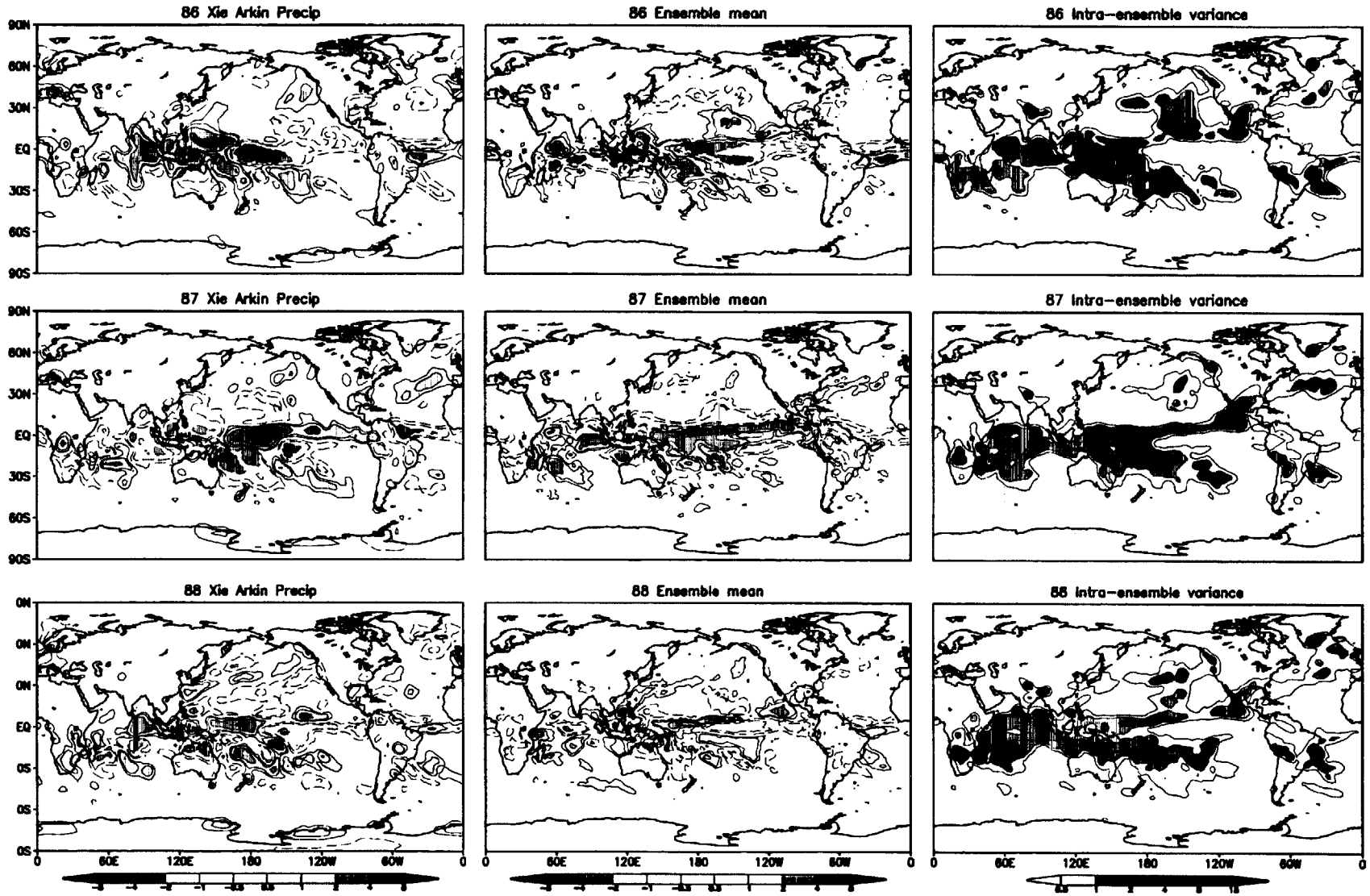

Figure 43: JFM precipitation for 1986, 1987, and 1988. Left panels are the anomalies from the Xie-Arkin data; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. The units of precipitation are mm/day.

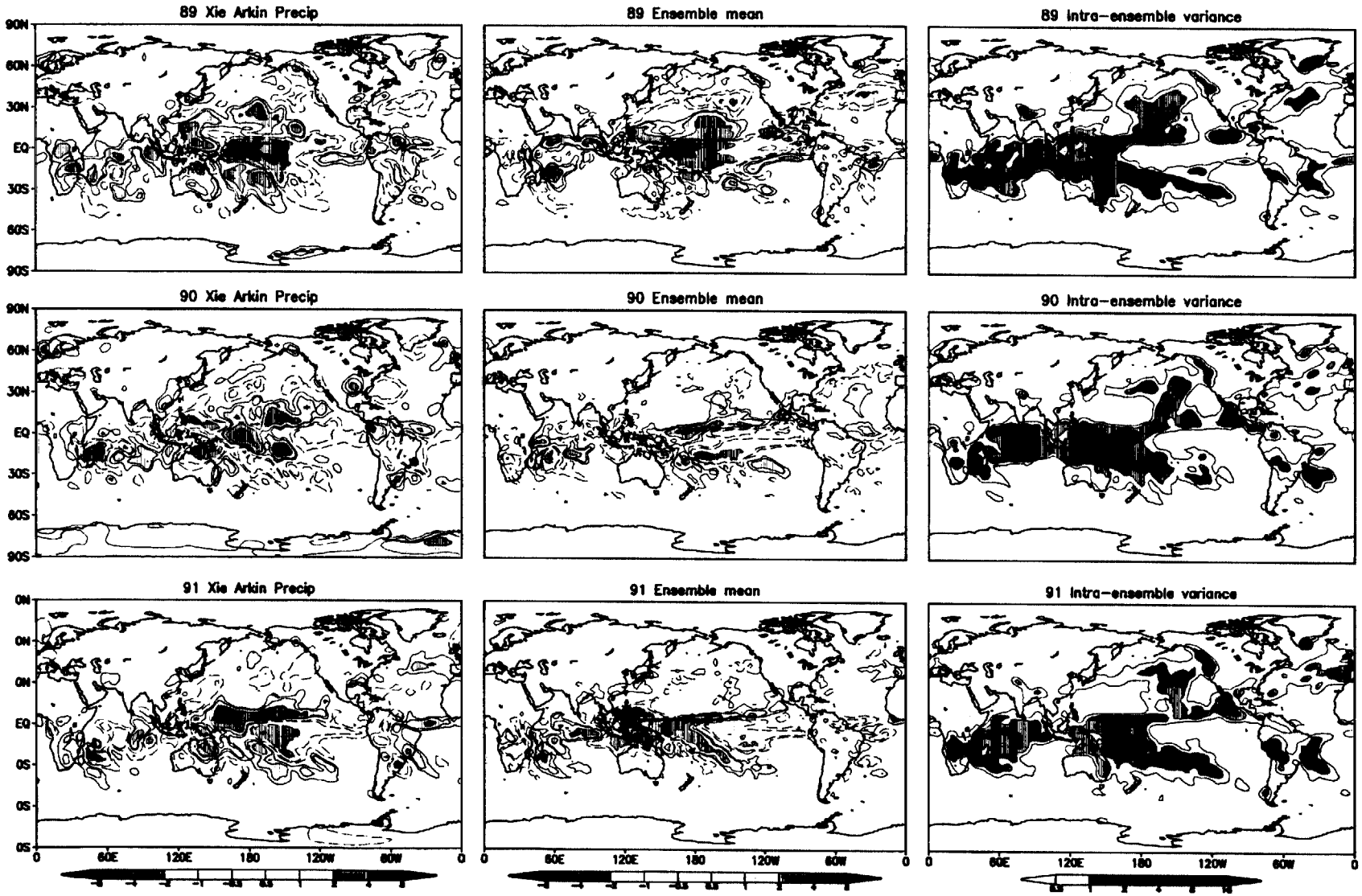

Figure 44: JFM precipitation for 1989, 1990, and 1991. Left panels are the anomalies from the Xie-Arkin data; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. The units of precipitation are mm/day.

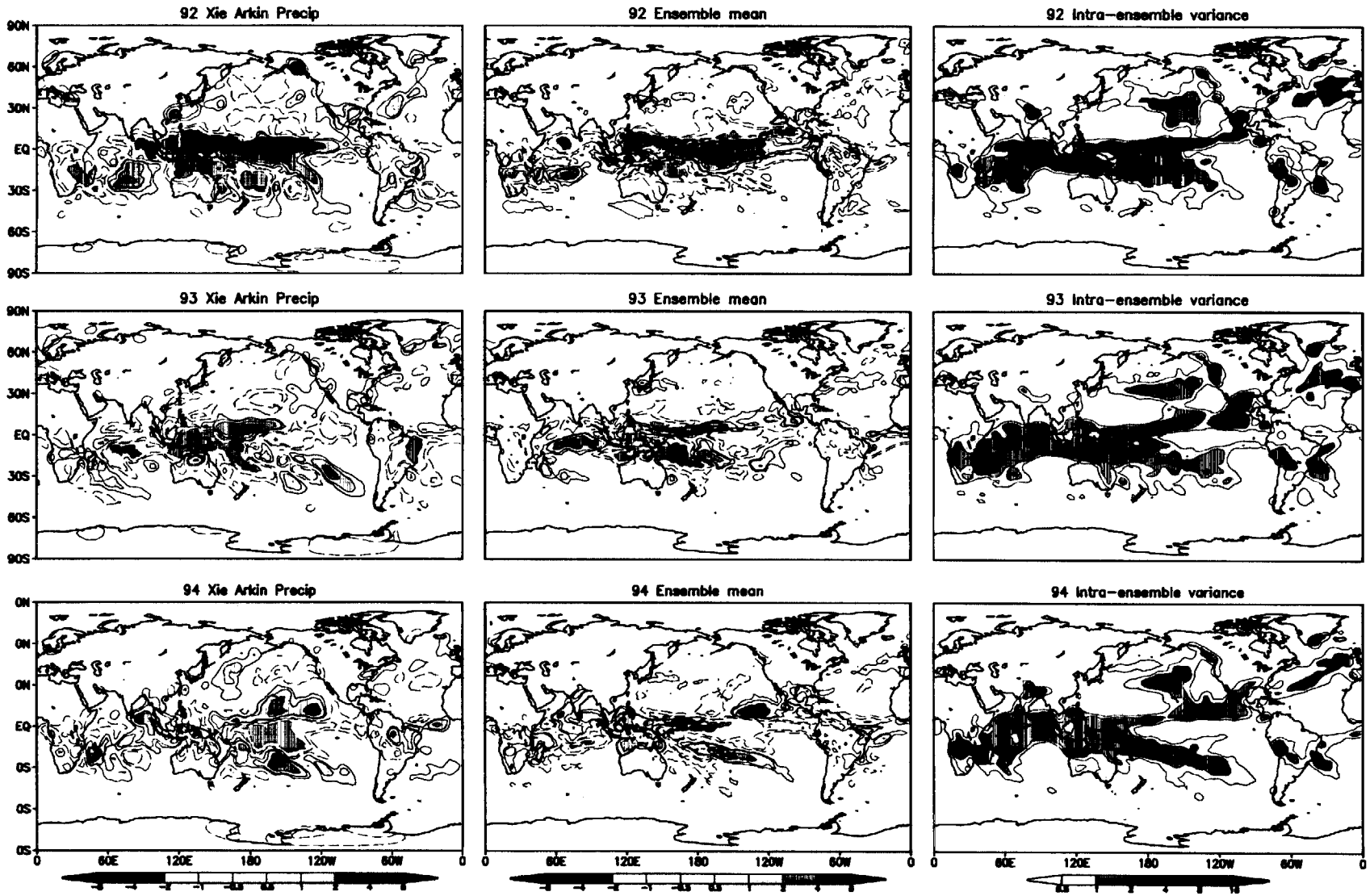

Figure 45: JFM precipitation for 1992, 1993, and 1994. Left panels are the anomalies from the Xie-Arkin data; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. The units of precipitation are mm/day.

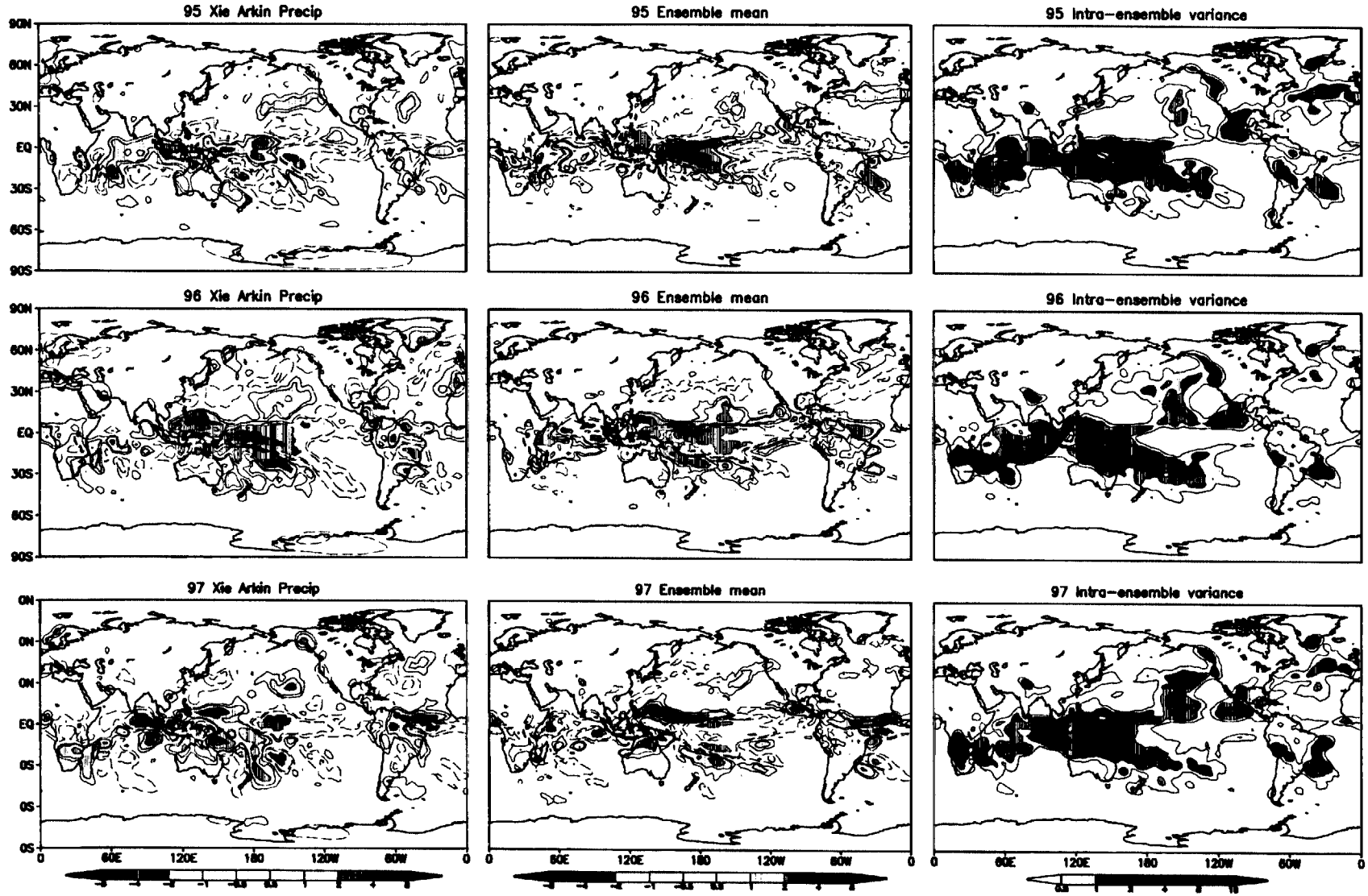

Figure 46: JFM precipitation for 1995, 1996, and 1997. Left panels are the anomalies from the Xie-Arkin data; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. The units of precipitation are mm/day.

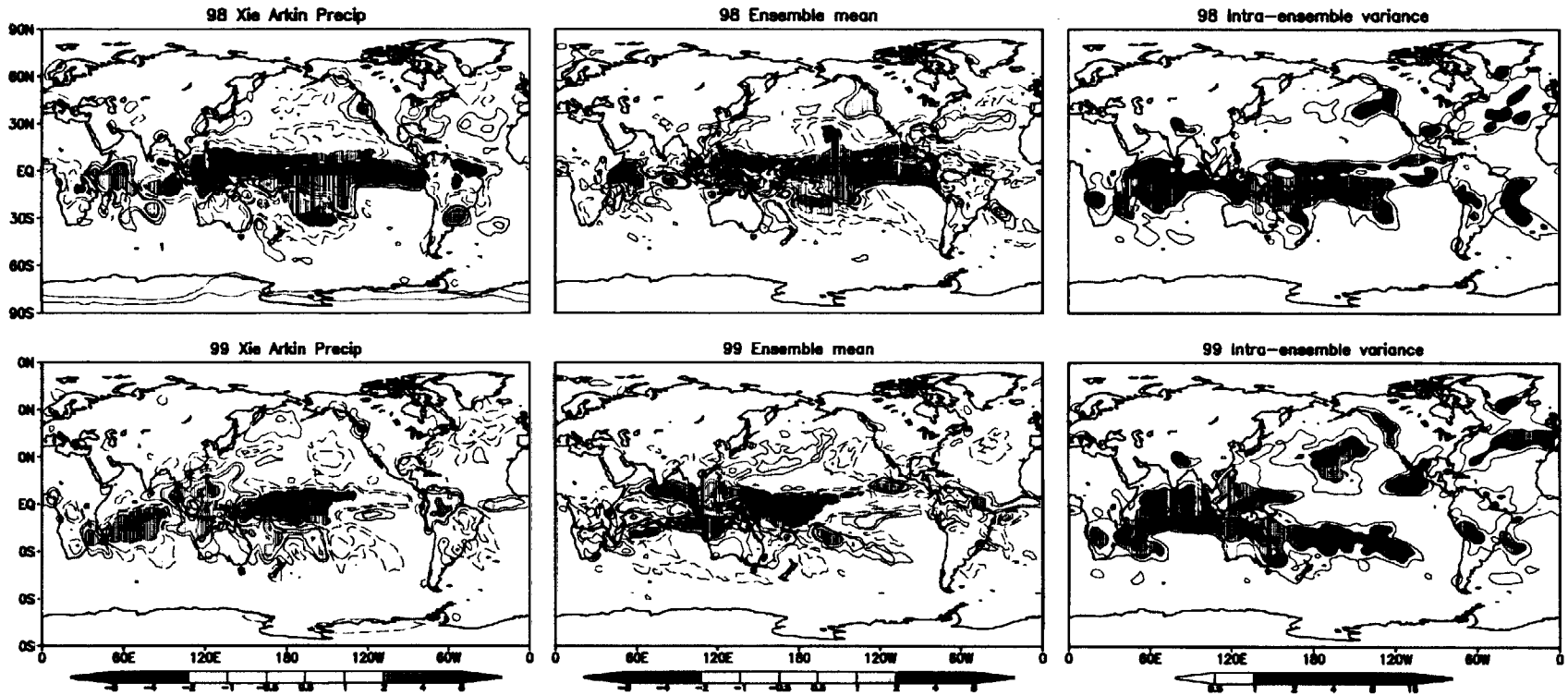

Figure 47: JFM precipitation for 1998 and 1999. Left panels are the anomalies from the Xie-Arkin data; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. The units of precipitation are mm/day.

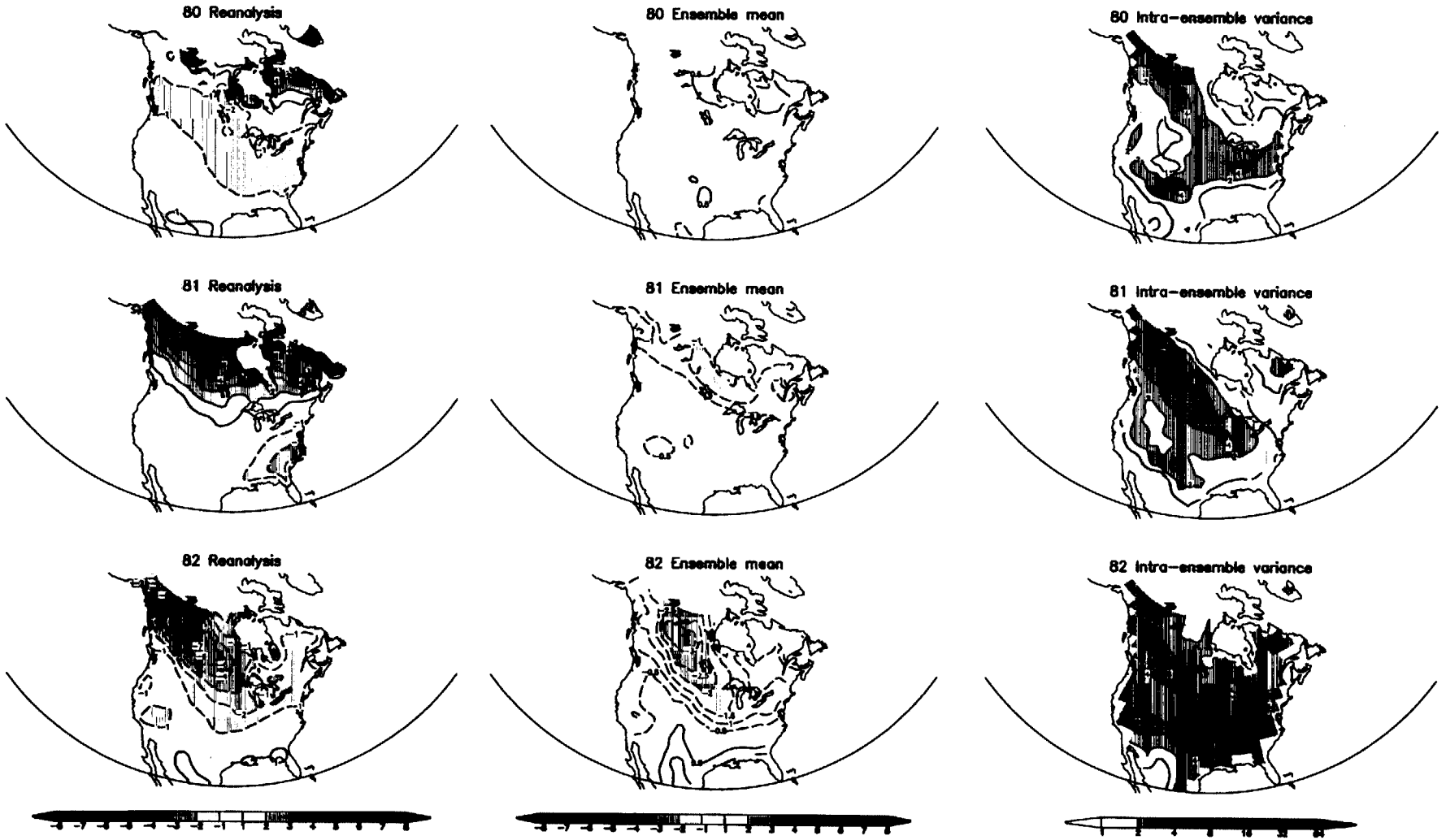

Figure 48: JFM surface temperature for 1980, 1981, and 1982. Left panels are the anomalies from the reanalysis; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. Units of temperature are degrees Kelvin.

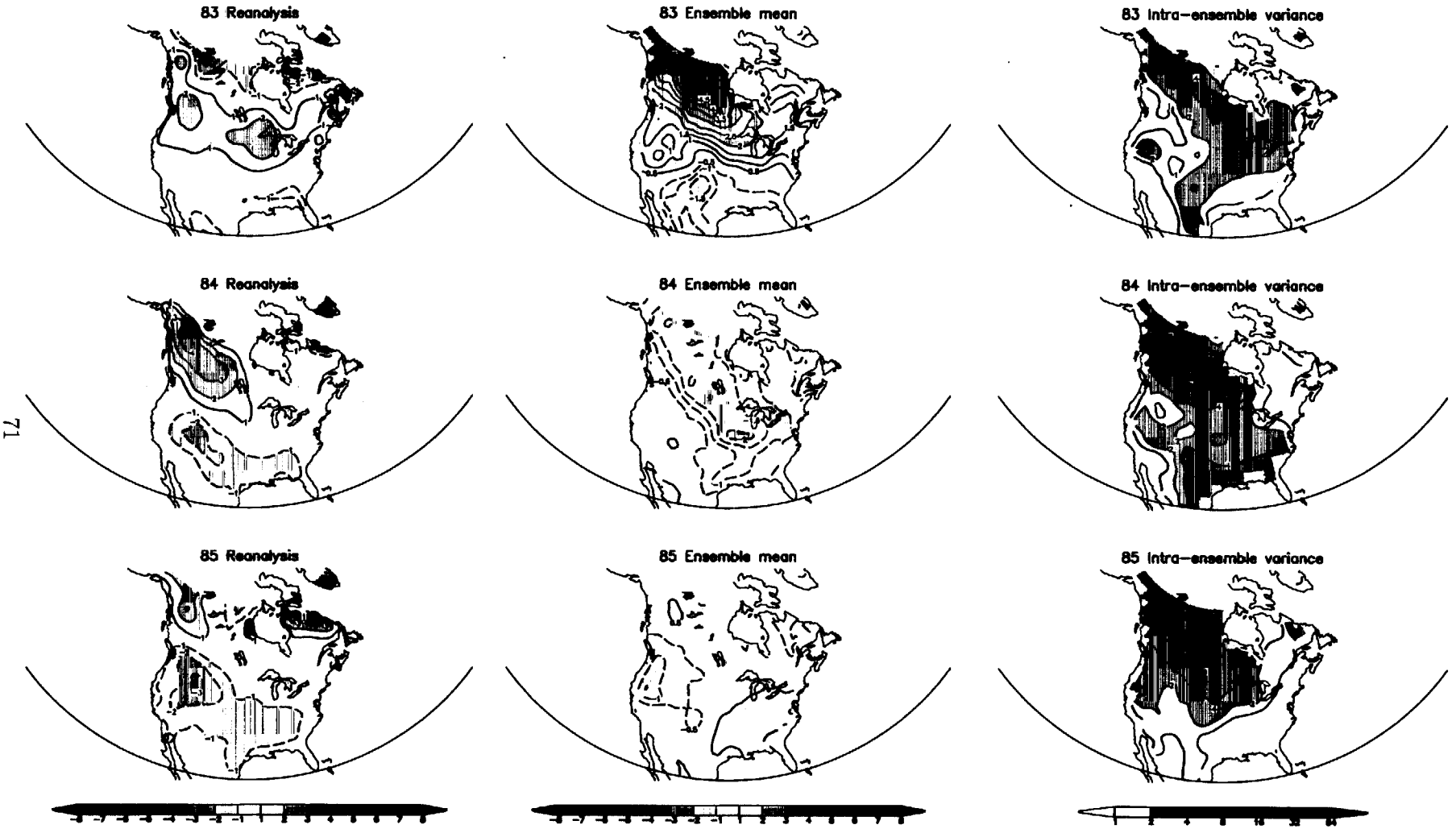

Figure 49: JFM surface temperature for 1983, 1984, and 1985. Left panels are the anomalies from the reanalysis; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. Units of temperature are degrees Kelvin.

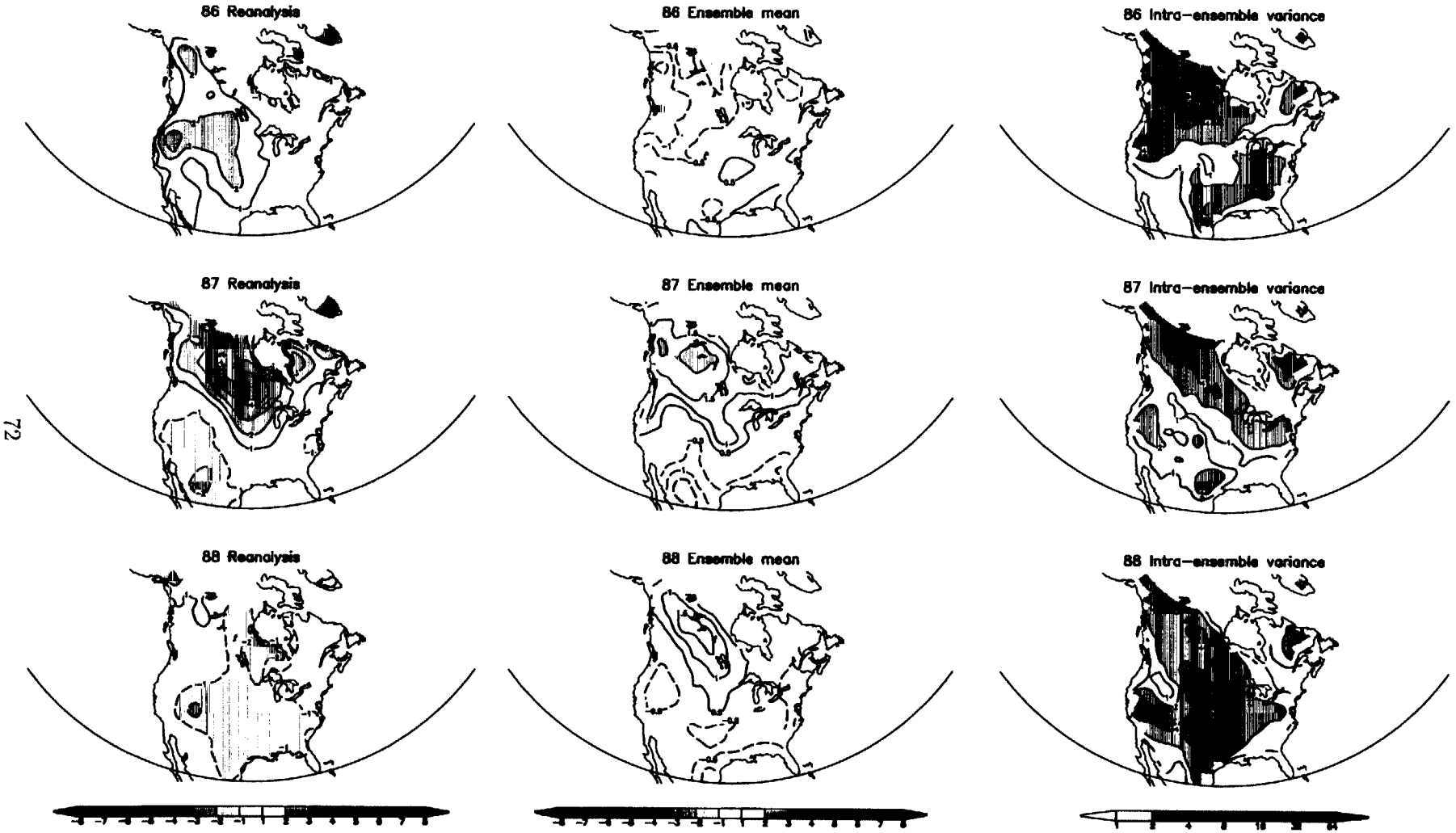

Figure 50: JFM surface temperature for 1986, 1987, and 1988. Left panels are the anomalies from the reanalysis; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. Units of temperature are degrees Kelvin.

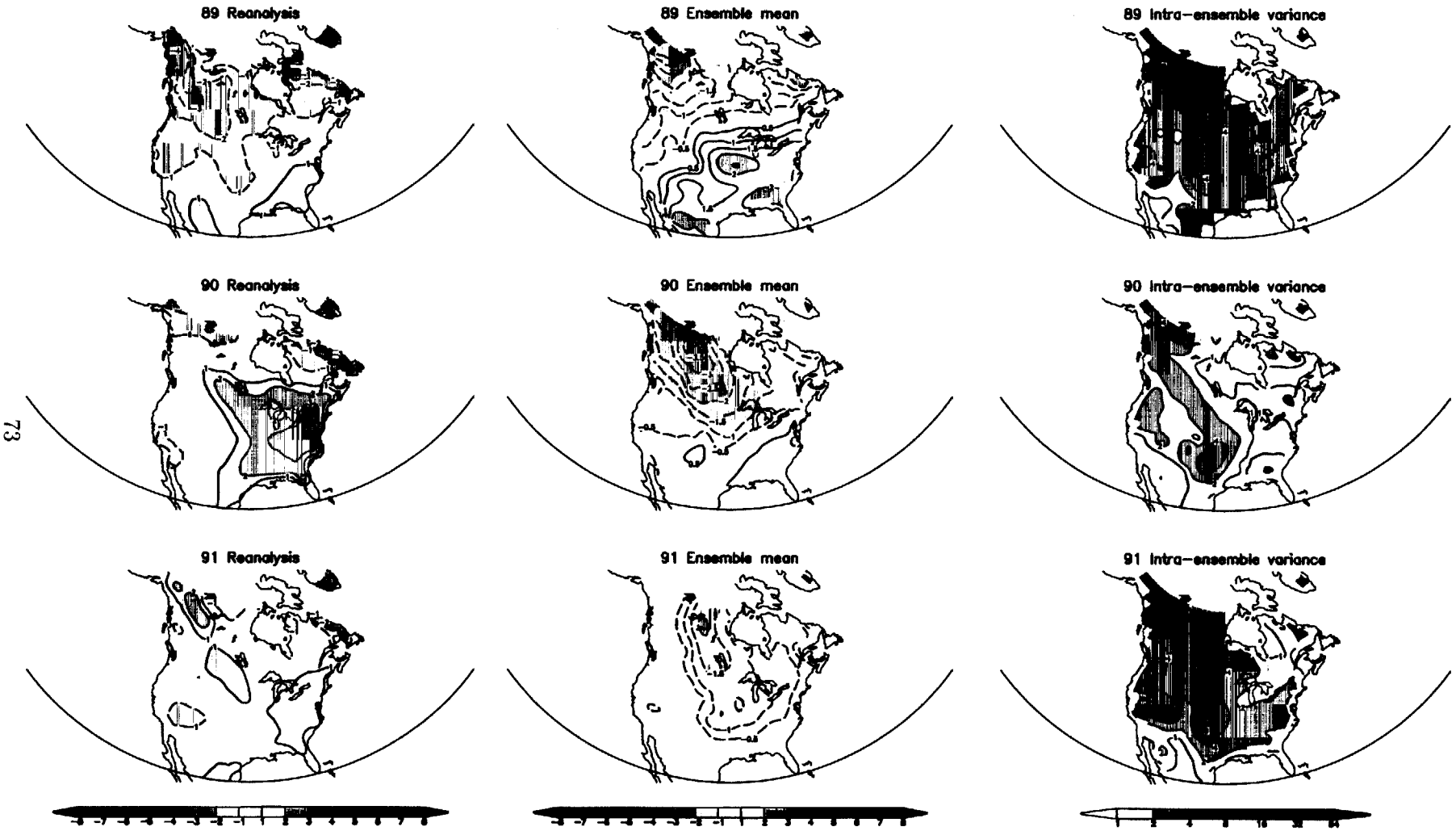

Figure 51: JFM surface temperature for 1989, 1990, and 1991. Left panels are the anomalies from the reanalysis; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. Units of temperature are degrees Kelvin.

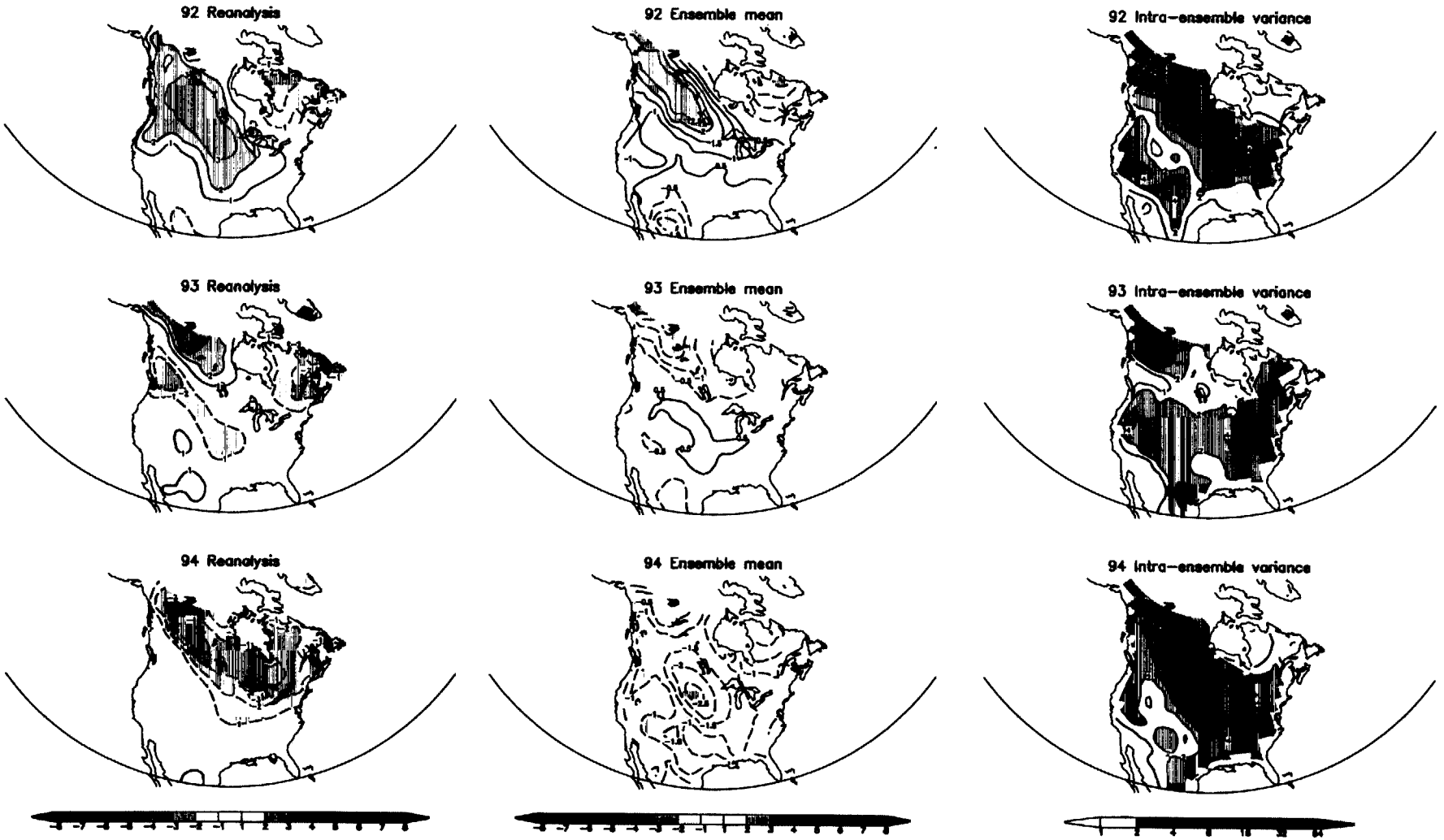

Figure 52: JFM surface temperature for 1992, 1993, and 1994. Left panels are the anomalies from the reanalysis; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. Units of temperature are degrees Kelvin.

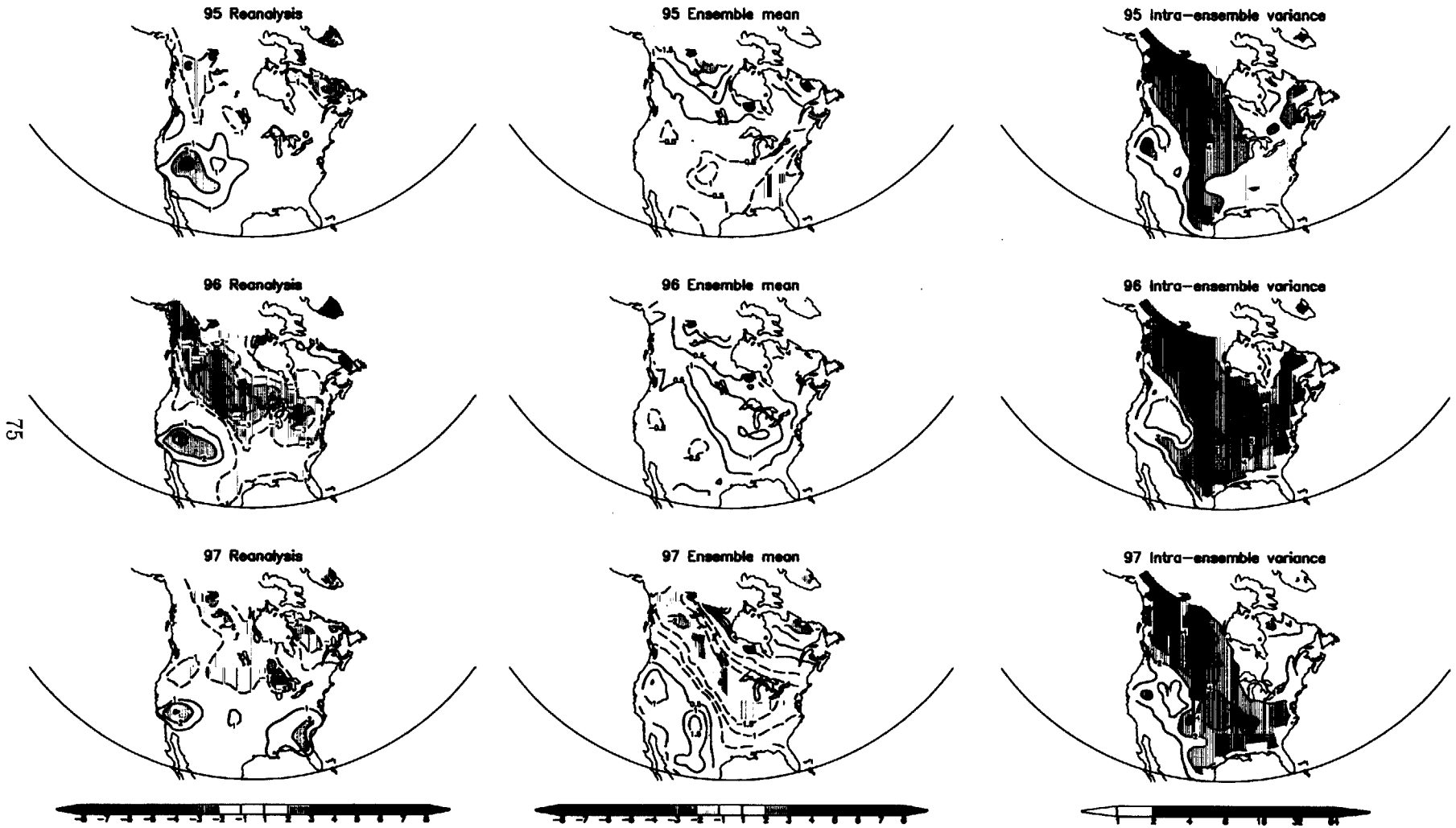

Figure 53: JFM surface temperature for 1995, 1996, and 1997. Left panels are the anomalies from the reanalysis; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. Units of temperature are degrees Kelvin.

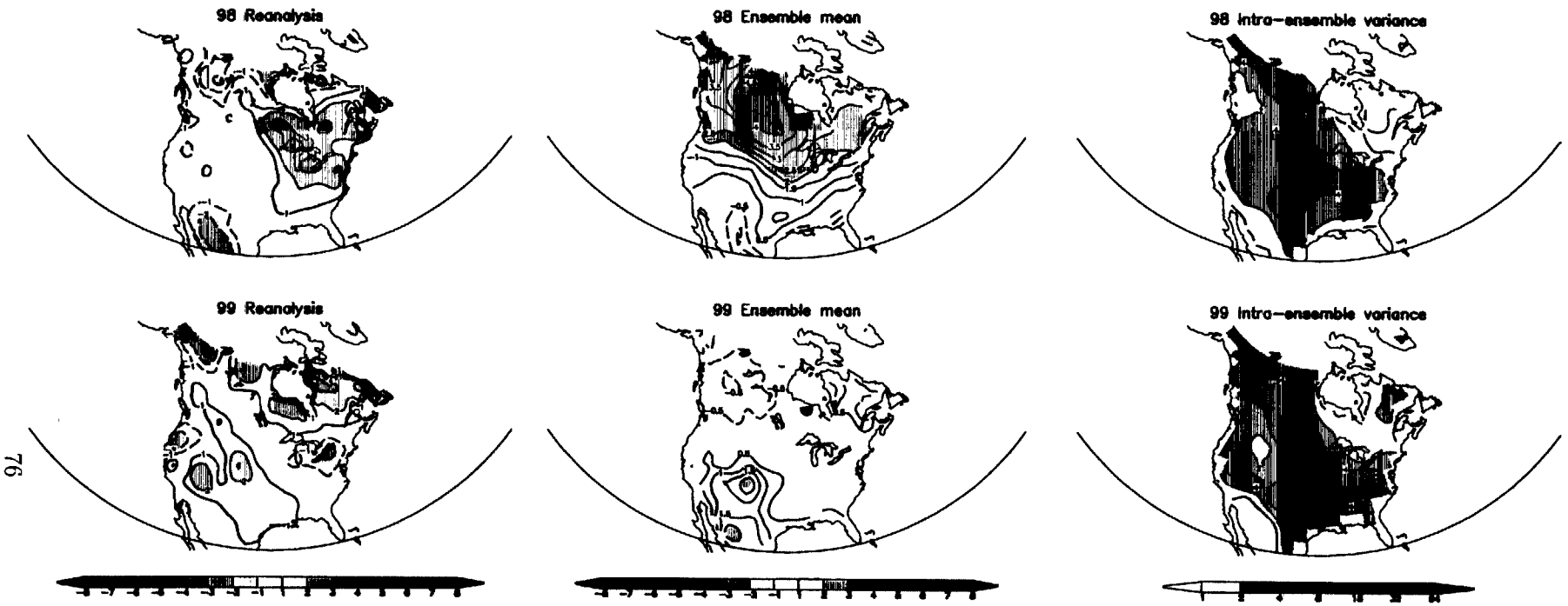

Figure 54: JFM surface temperature for 1998 and 1999. Left panels are the anomalies from the reanalysis; center panels are the ensemble-mean anomalies for the model; and right panels are the model's intra-ensemble variance. For the ensemble mean, colored regions are significant at the two-sided 5% level. Units of temperature are degrees Kelvin.

LARGE ENSEMBLE COMPARISONS

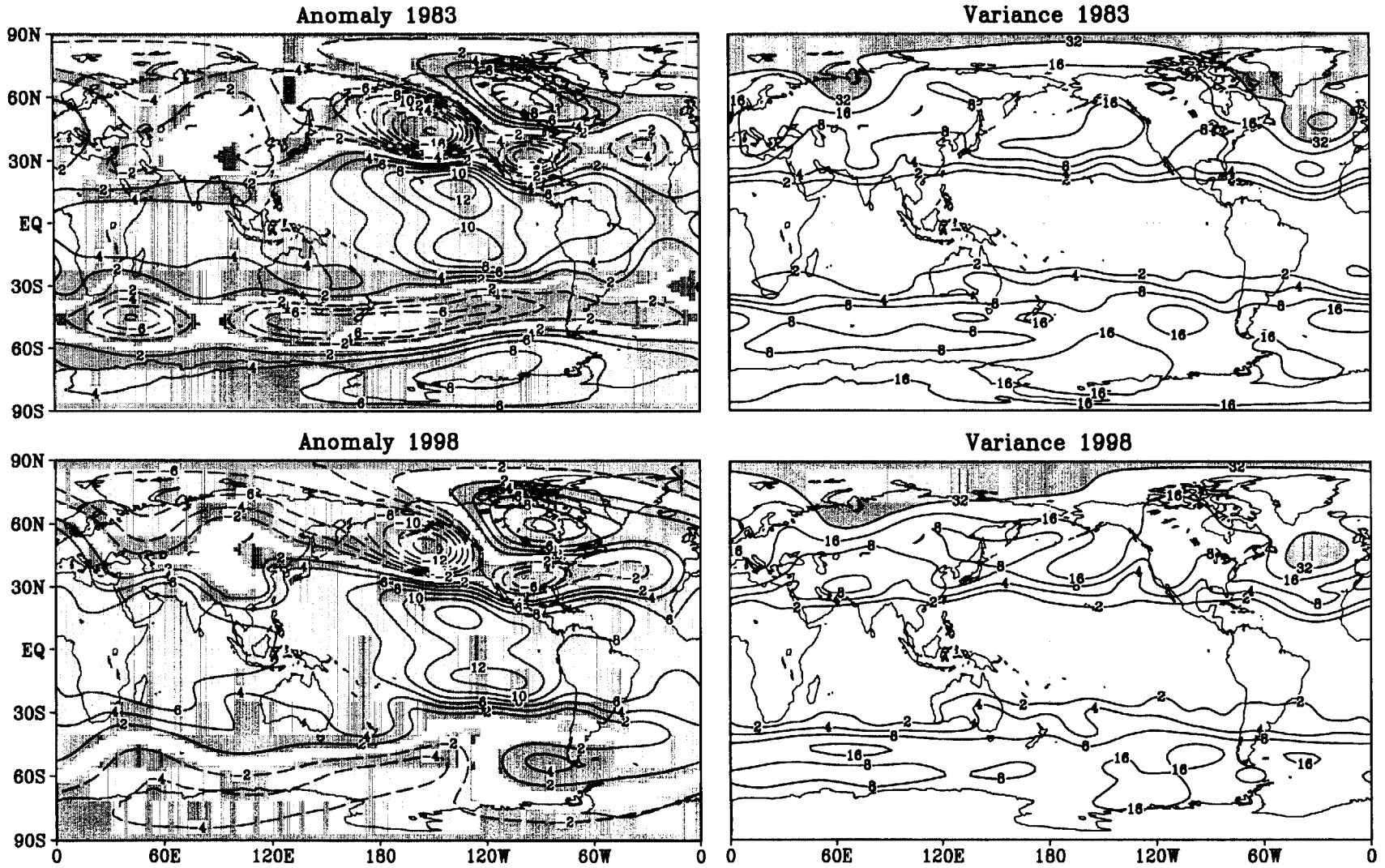

Figure 55: Model results for 1998 and 1983 for the JFM geopotential height at 200 mb. Both ensembles consist of 36 members. The left panels are the ensemble mean anomalies (shading indicates values significant at the two-sided 5% level). The contour interval is 20 m. Right panels show the intra-ensemble variance. The contour levels are 2 4 8 16 32 $\times 10^2$ m^2 and values greater than 3200 m^2 are shaded.

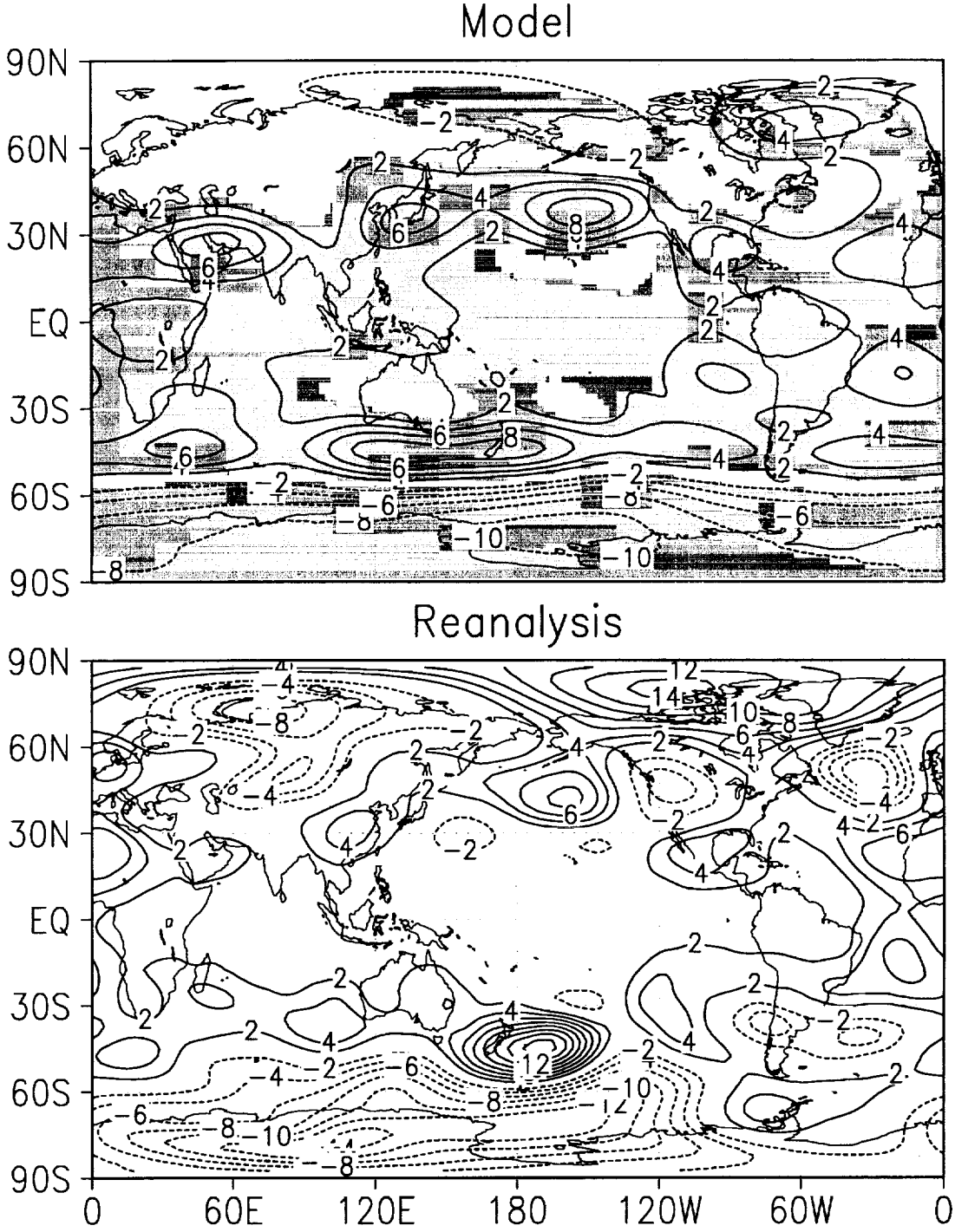

Figure 56: Difference in geopotential height at 200 mb (1998 minus 1983). Top panel is for the model's ensemble mean. Lower panel is for the observations. For the model, only values that pass a two-side 5% confidence test are shaded. Contour interval is 20 m. Both ensembles consist of 36 members.

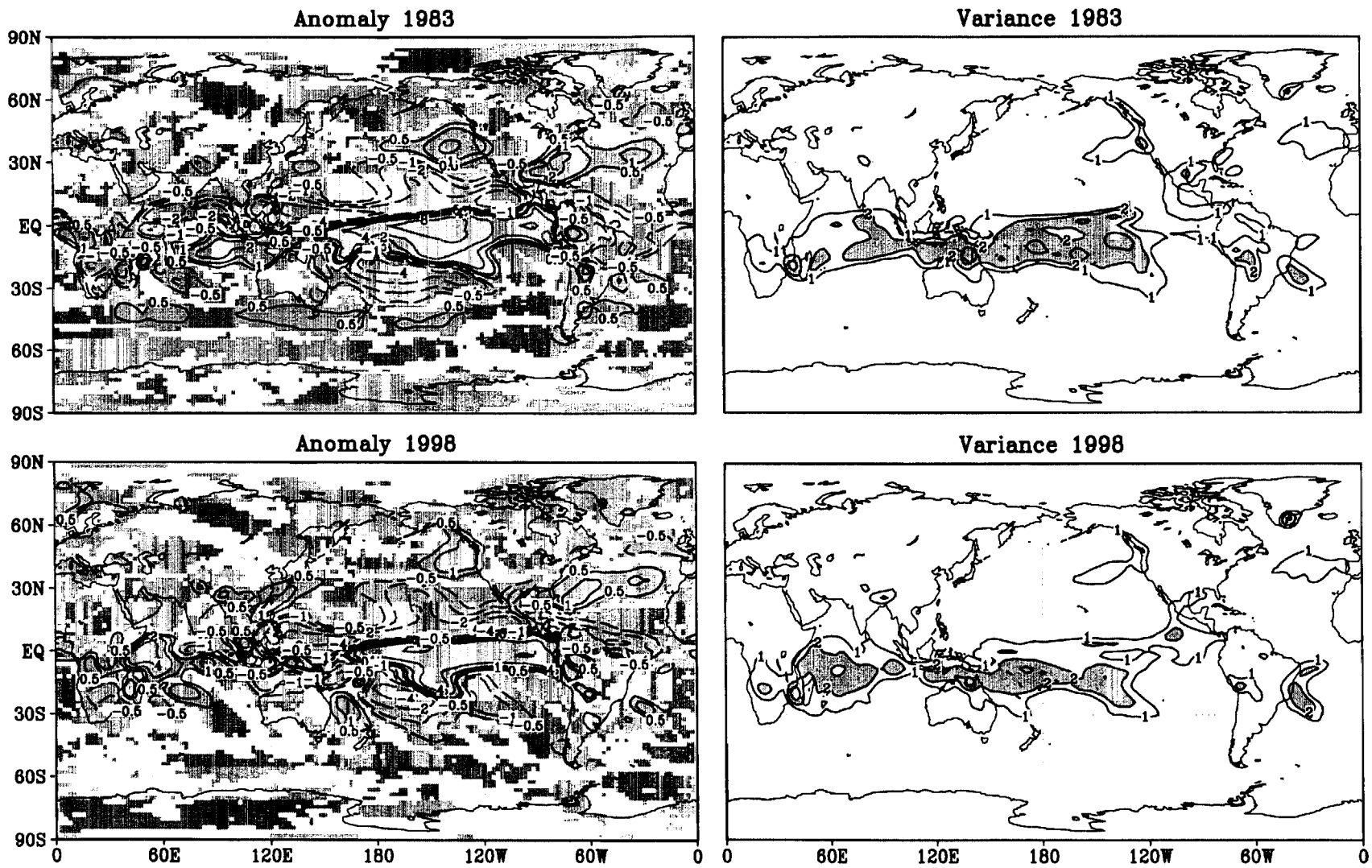

Figure 57: Model results for 1998 and 1983 for the JFM precipitation . Both ensembles consist of 36 members. The left panels are the ensemble mean anomalies (shading indicates values significant at the two-sided 5% level). The contour levels are ± 0.5 1 2 4 8 16 (mm/day). Right panels show the intra-ensemble variance. The contour levels are 1 2 4 8 (mm/day)^2 and values greater than 2 (mm/day)^2 are shaded.

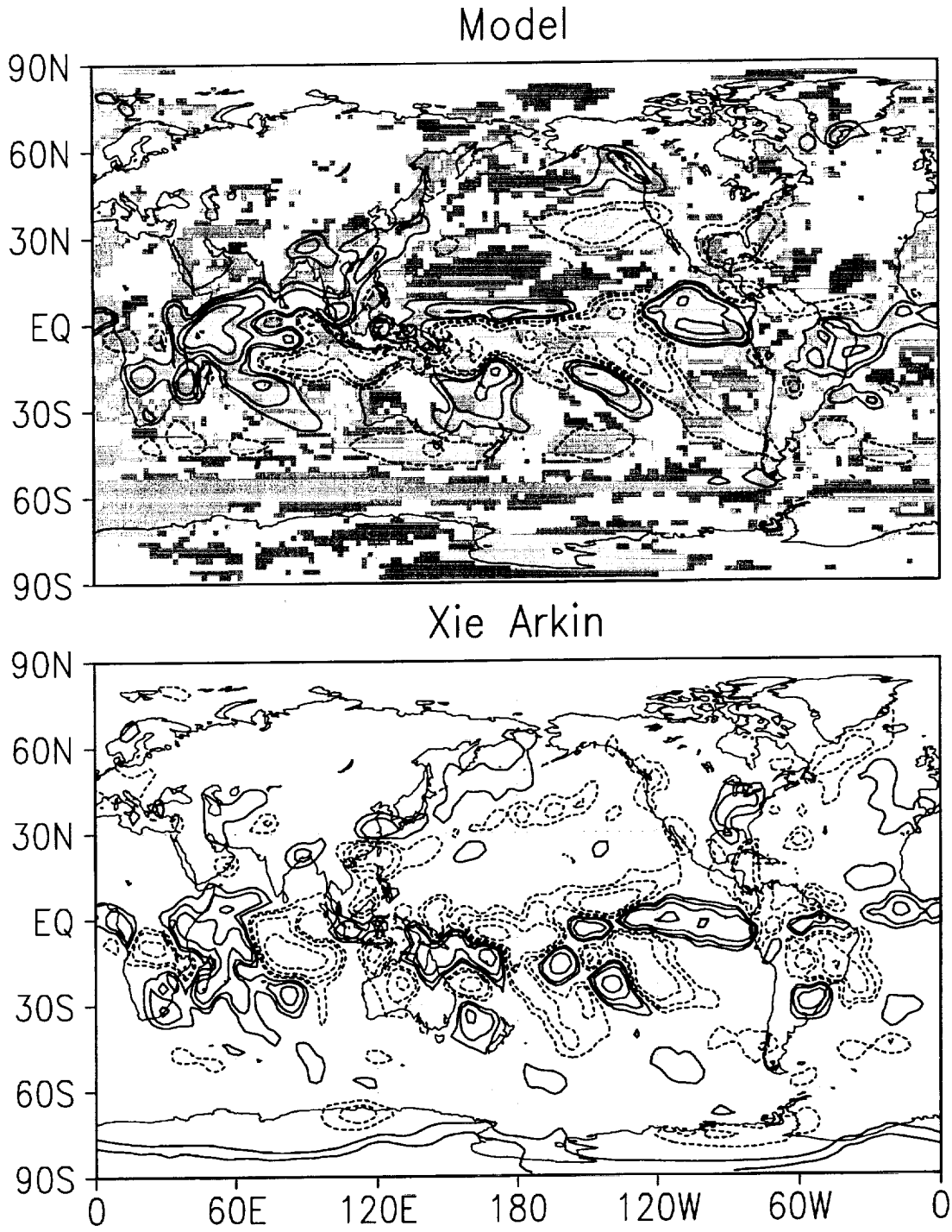

Figure 58: Difference in precipitation (1998 minus 1983). Top panel is for the model's ensemble mean. Lower panel is for the observations. For the model, only values that pass a two-side 5% confidence test are shaded. Contour levels are ± 0.5 1 2 4 8 16 (mm/day). Both ensembles consist of 36 members.

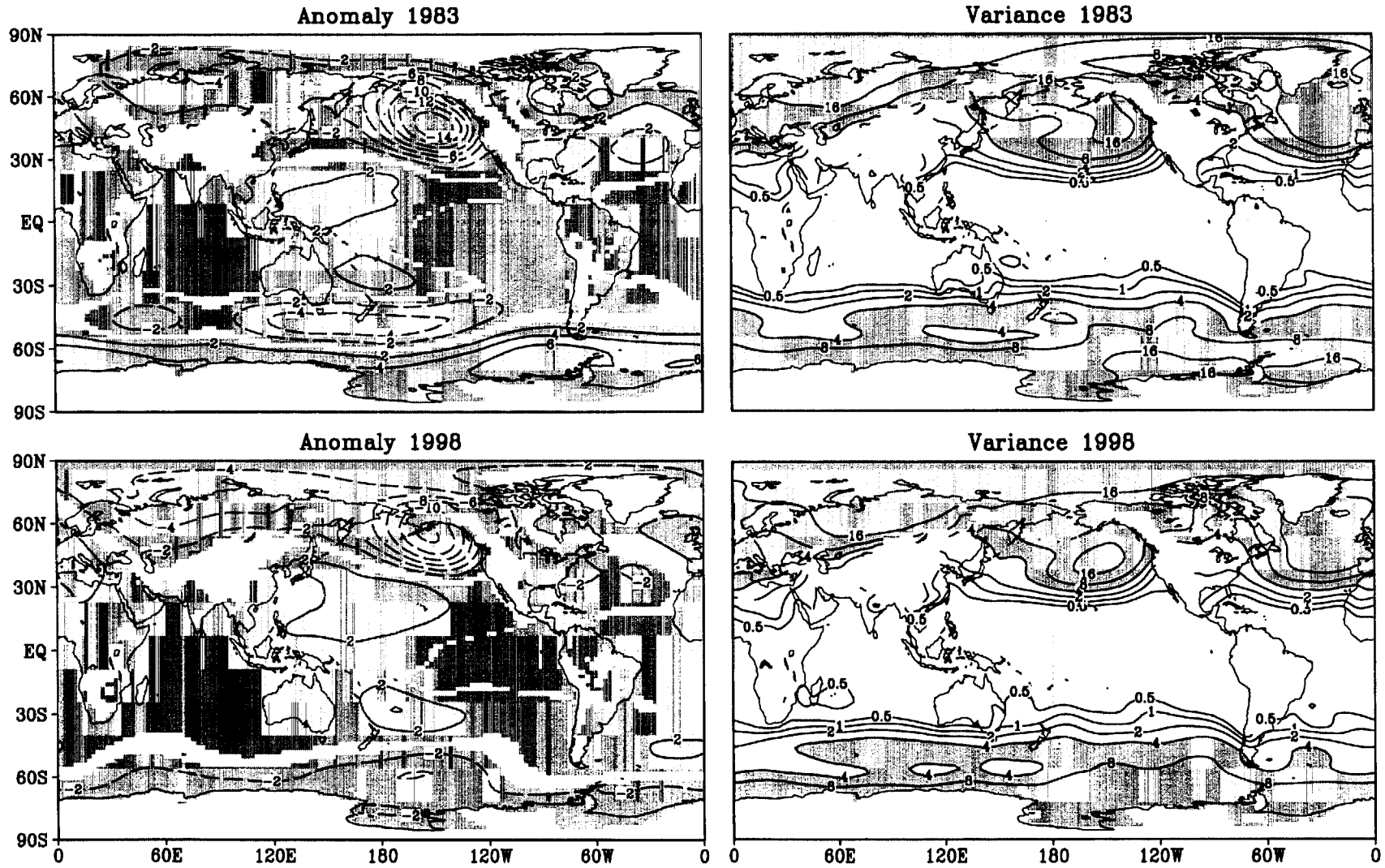

Figure 59: Model results for 1998 and 1983 for the JFM sea-level pressure . Both ensembles consist of 36 members. The left panels are the ensemble mean anomalies (shading indicates values significant at the two-sided 5% level). The contour interval is 2 mb. Right panels show the intra-ensemble variance. The contour levels are 0.5 1 2 4 8 16 mb^2 and values greater than 4 mb^2 are shaded.

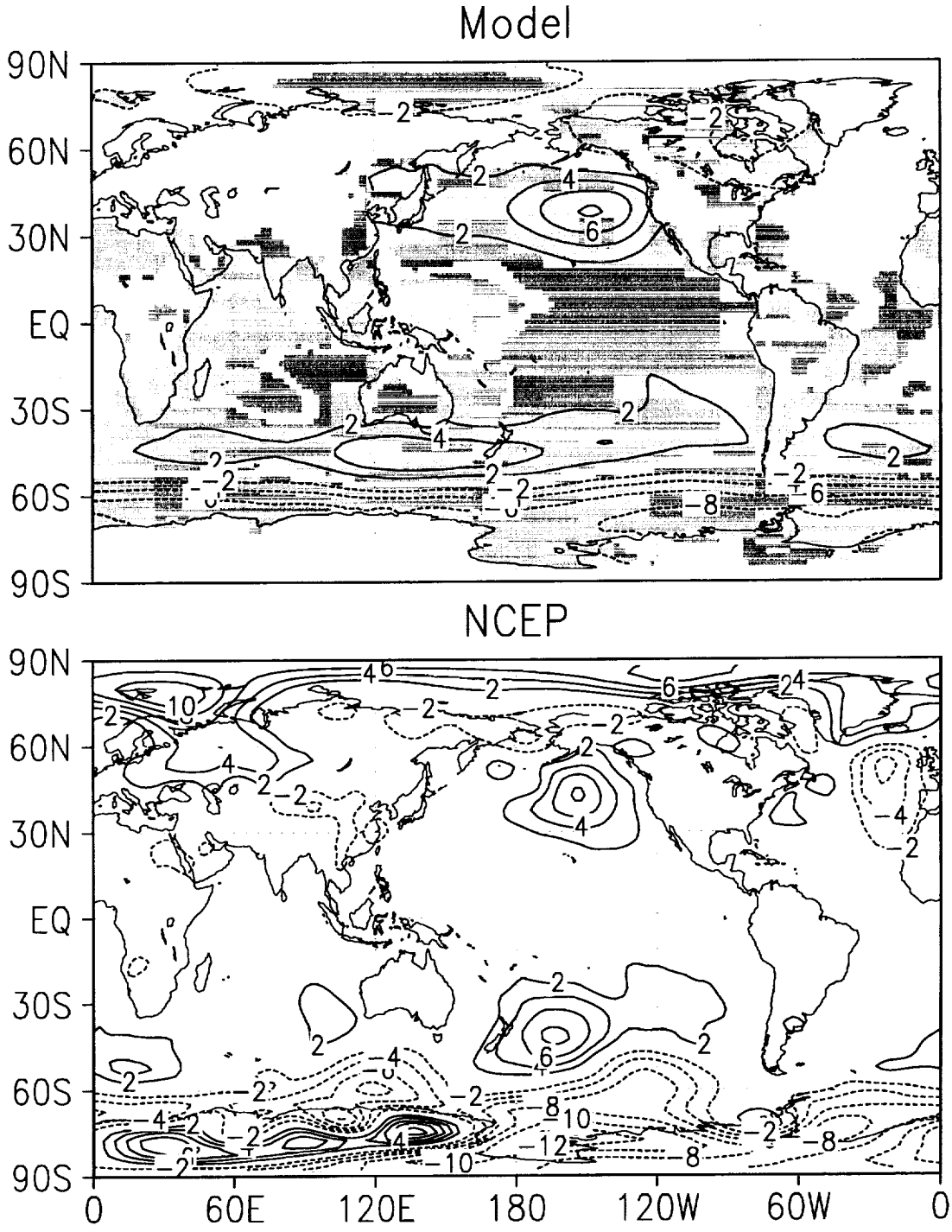

Figure 60: Difference in sea-level pressure (1998 minus 1983). Top panel is for the model's ensemble mean. Lower panel is for the observations. For the model, only values that pass a two-side 5% confidence test are shaded. Contour interval is 2 mb. Both ensembles consist of 36 members.

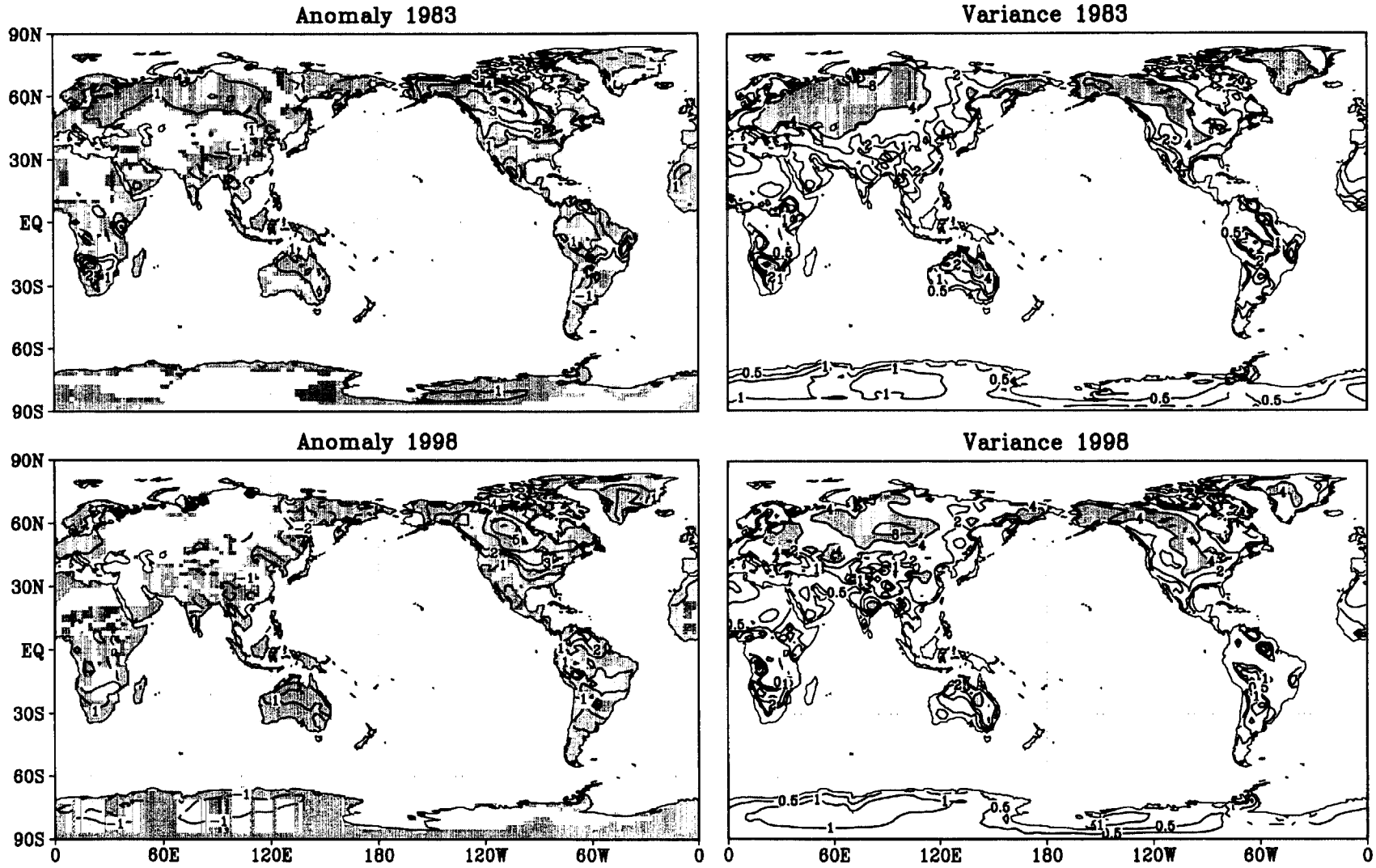

Figure 61: Model results for 1998 and 1983 for the JFM surface temperature . Both ensembles consist of 36 members. The left panels are the ensemble mean anomalies (shading indicates values significant at the two-sided 5% level). The contour interval is 1 K and the zero contour is omitted.. Right panels show the intra-ensemble variance. The contour levels are 1 2 4 8 K^2 and values greater than 4 K^2 are shaded.

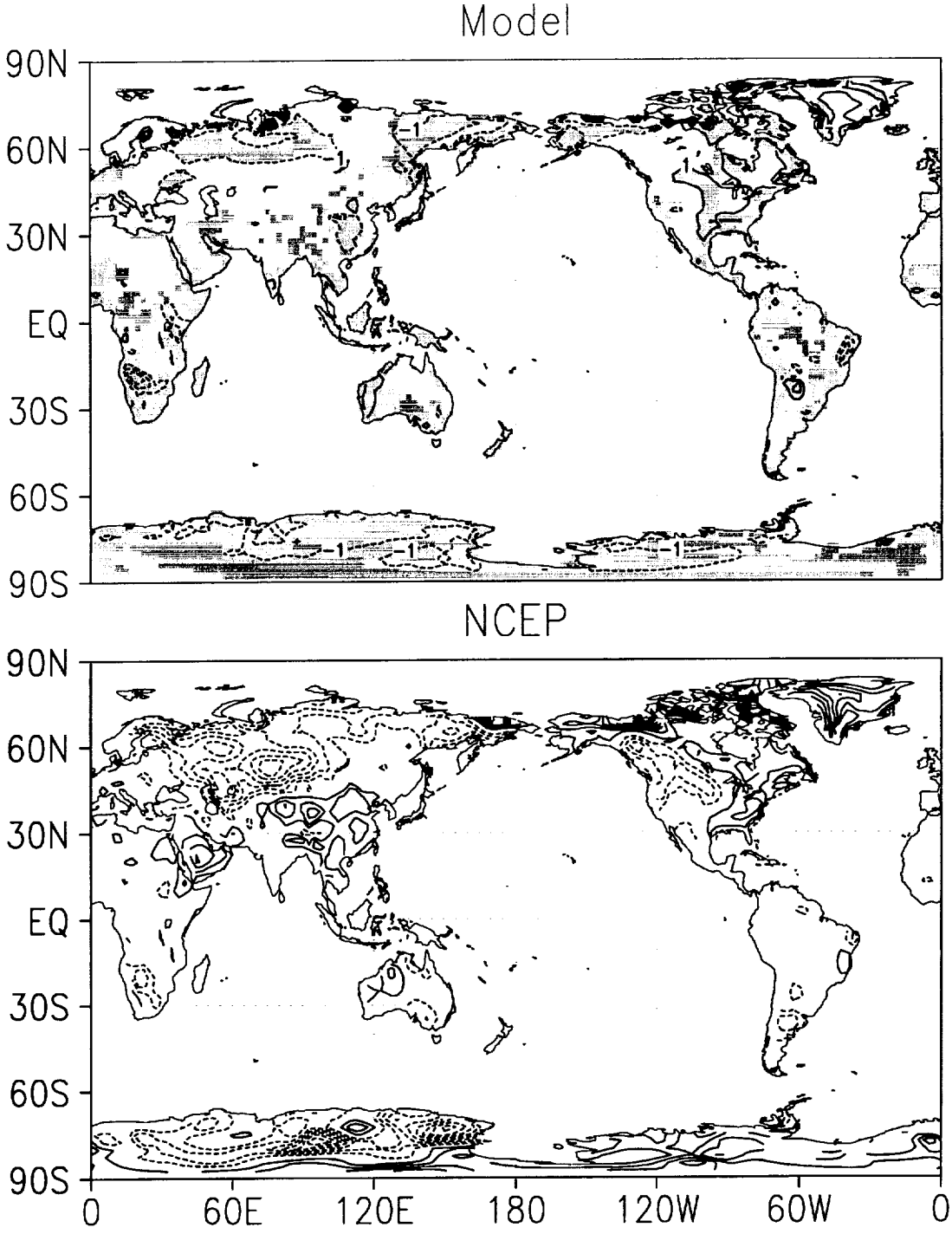

Figure 62: Difference in surface temperature (1998 minus 1983). Top panel is for the model's ensemble mean. Lower panel is for the observations. For the model, only values that pass a two-side 5% confidence test are shaded. Contour interval is 1 K and the zero contour is omitted.. Both ensembles consist of 36 members.

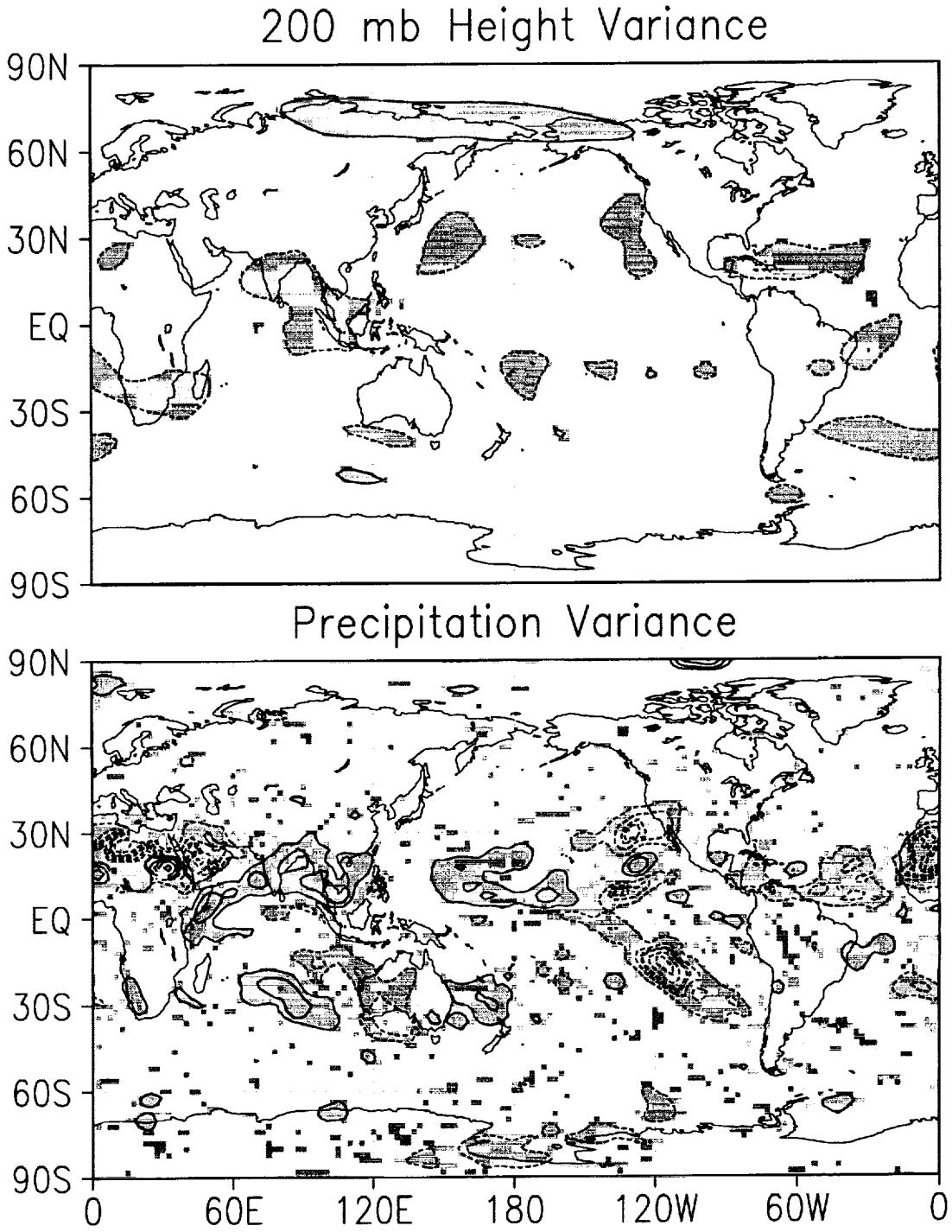

Figure 63: Ratio of the intra-ensemble variance of seasonal means of 200mb geopotential height and precipitation (1998 / 1983). Both ensembles consist of 36 members. Shading indicates values are significantly different from one at the two-sided 5% level based on an F test. Contour levels are ... $\frac{1}{8}\frac{1}{4}\frac{1}{2}$ 2 4 8.... Contours less than 1 are dashed. Note the 1 contour is omitted.

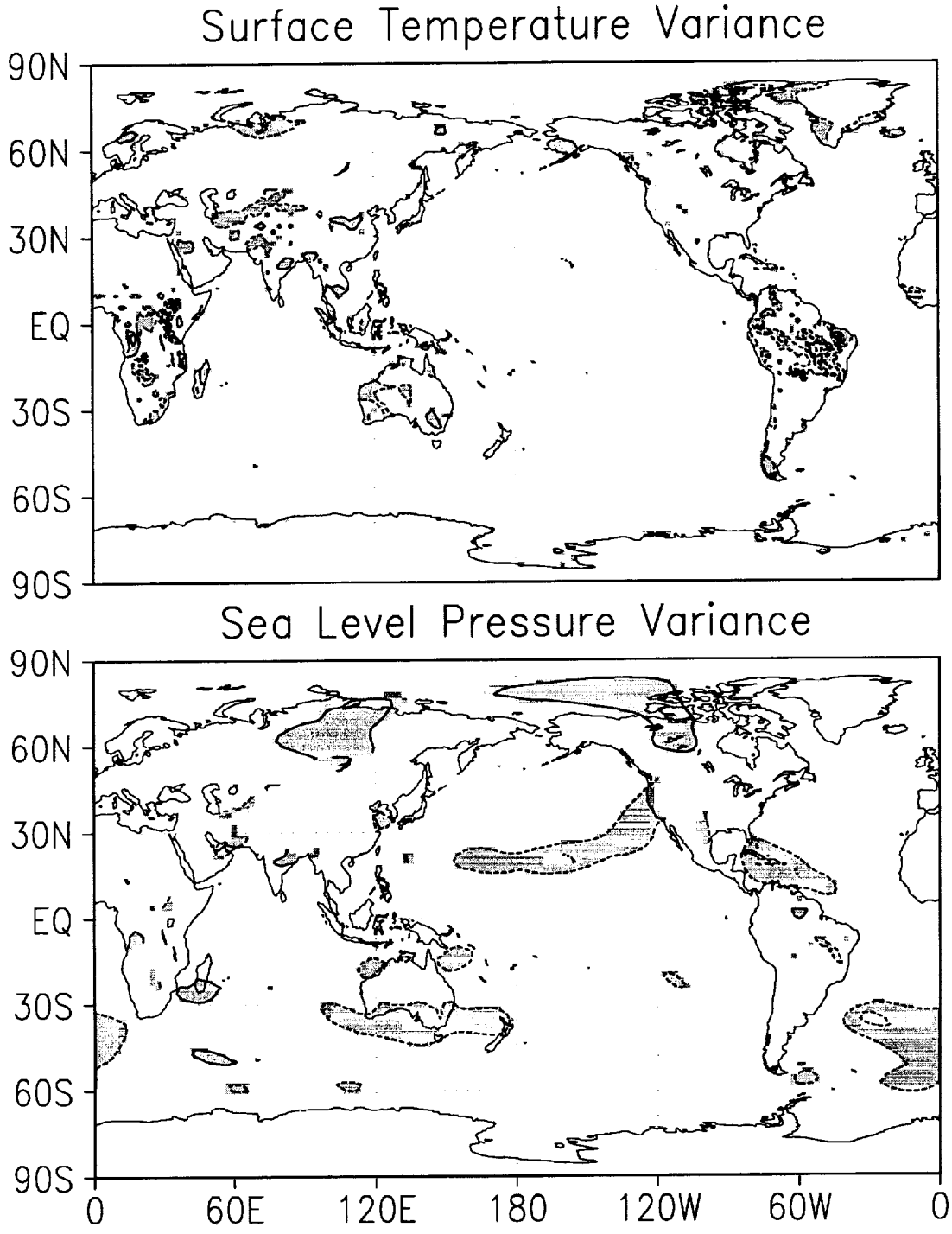

Figure 64: Ratio of the intra-ensemble variance of seasonal means of surface temperature and sea-level pressure (1998 / 1983). Both ensembles consist of 36 members. Shading indicates values are significantly different from one at the two-sided 5% level based on an F test. Contour levels are ... $\frac{1}{8}\frac{1}{4}\frac{1}{2}$ 2 4 8.... Contours less than 1 are dashed. Note the 1 contour is omitted.

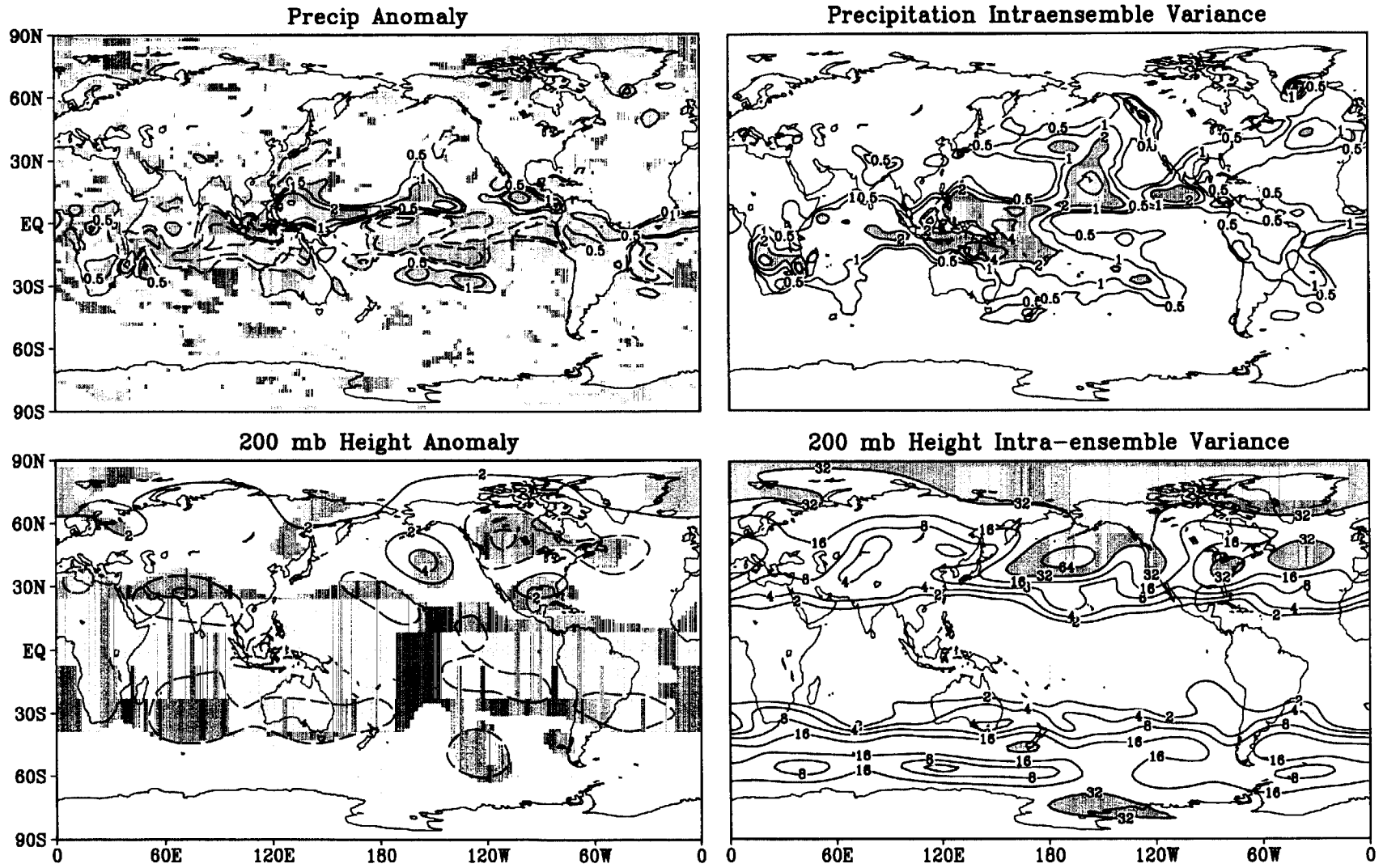

Figure 65: Model results for 1982 with 18 ensemble members for JFM-mean 200 mb height and precipitation. The left panels are the ensemble mean anomalies (shading indicates values significant at the two-sided 5% level). The contour levels are ± 0.5 1 2 4 (mm/day) for the precipitation and the contour interval is 20 meters for the height. Right panels show the intra-ensemble variance. The contour levels are 1 2 4 $(\text{mm/day})^2$ for the precipitation variance and 2 4 8 16 32 64 m^2 for the height variance. Regions of large variance are shaded.

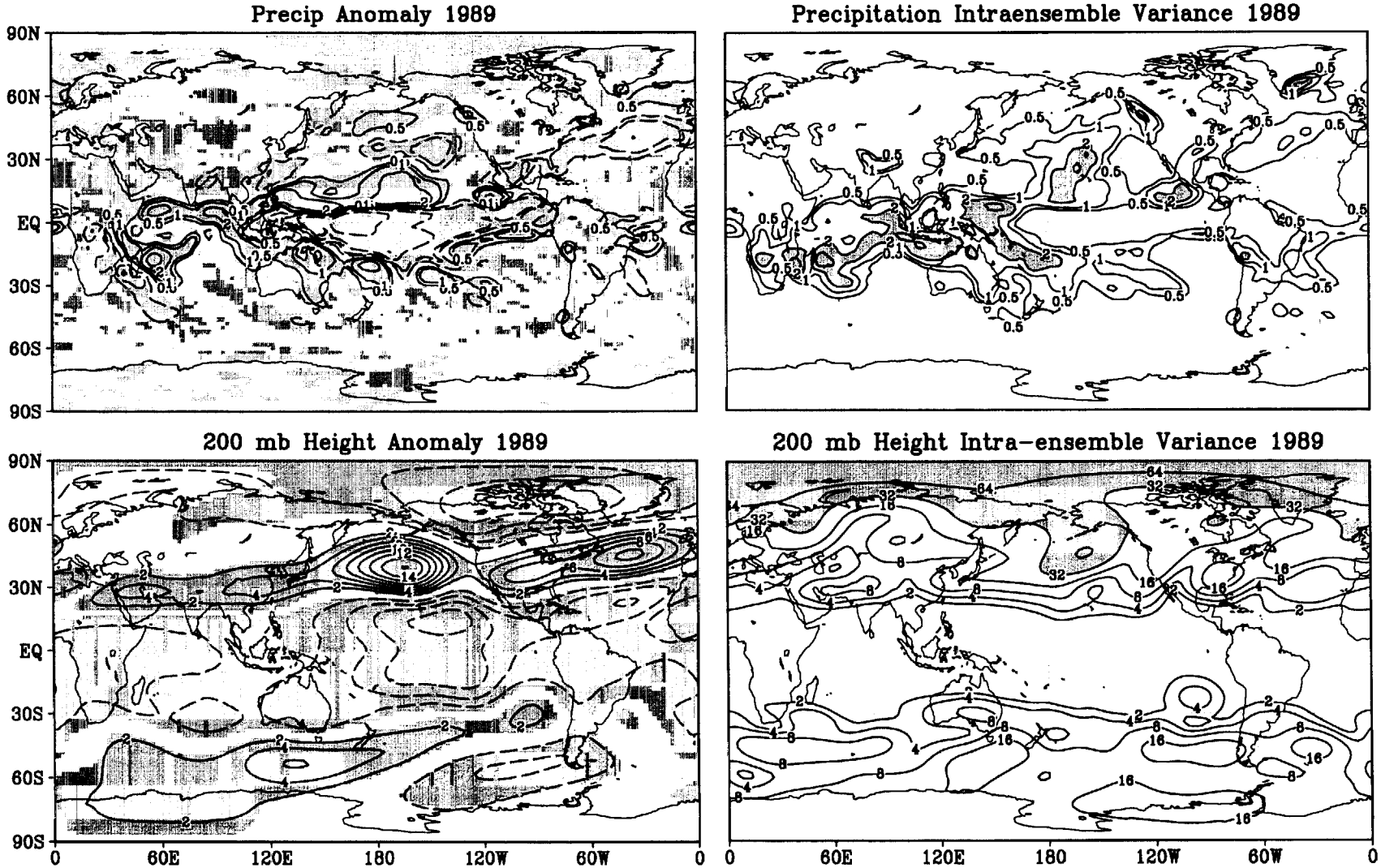

Figure 66: Model results for 1989 with 18 ensemble members for JFM-mean 200 mb height and precipitation. The left panels are the ensemble mean anomalies (shading indicates values significant at the two-sided 5% level). The contour levels are ± 0.5 1 2 4 (mm/day) for the precipitation and the contour interval is 20 meters for the height. Right panels show the intra-ensemble variance. The contour levels are 1 2 4 (mm/day)^2 for the precipitation variance and 2 4 8 16 32 64 m^2 for the height variance. Regions of large variance are shaded.

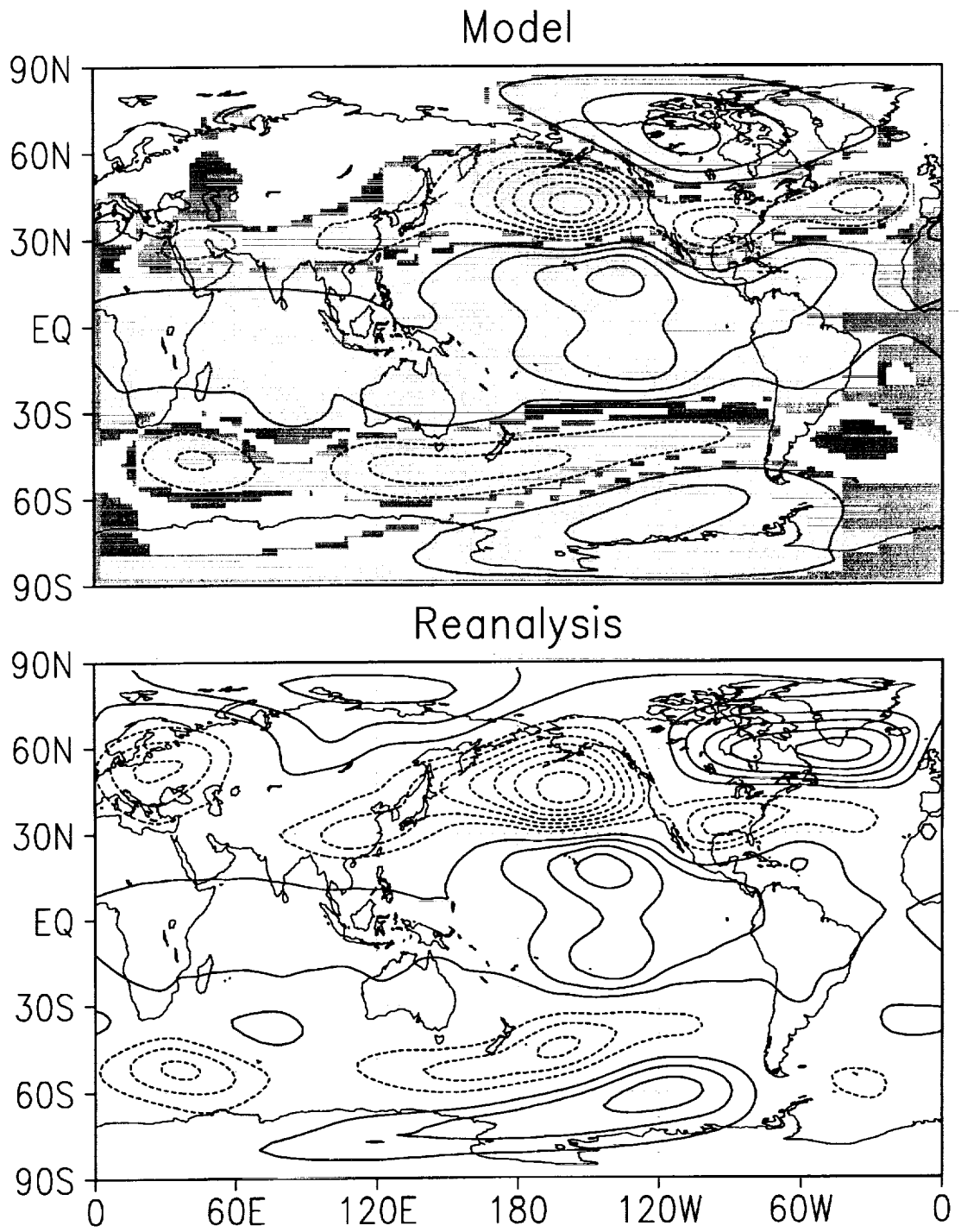

Figure 67: Difference in geopotential height at 200 mb (1983 minus 1989). Top panel is for the model's ensemble mean. Lower panel is for the reanalysis. The contour interval is 50 m; negative contours are dashed. For the model, only values that pass a 5% confidence test are shaded. The ensembles consist of 36 members for 1983 and 18 members for 1989. The statistical test takes into account the different ensemble sizes.

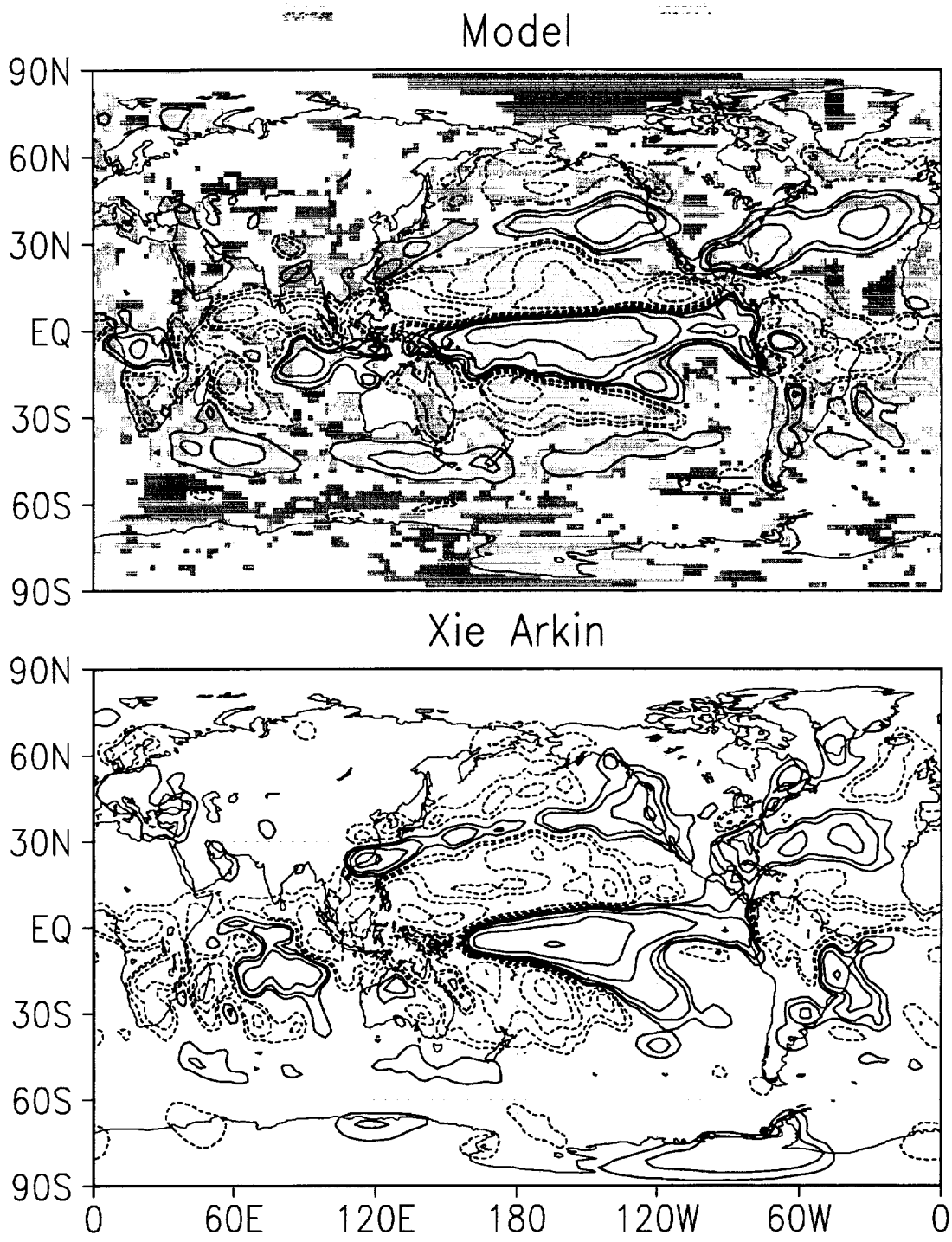

Figure 68: Difference in precipitation (1983 minus 1989). Top panel is for the model's ensemble mean. Lower panel is for the reanalysis. The contour levels are ± 0.5 1 2 4 8 16 mm/day; negative contours are dashed. For the model, only values that pass a 5% confidence test are shaded. The ensembles consist of 36 members for 1983 and 18 members for 1989. The statistical test takes into account the different ensemble sizes.

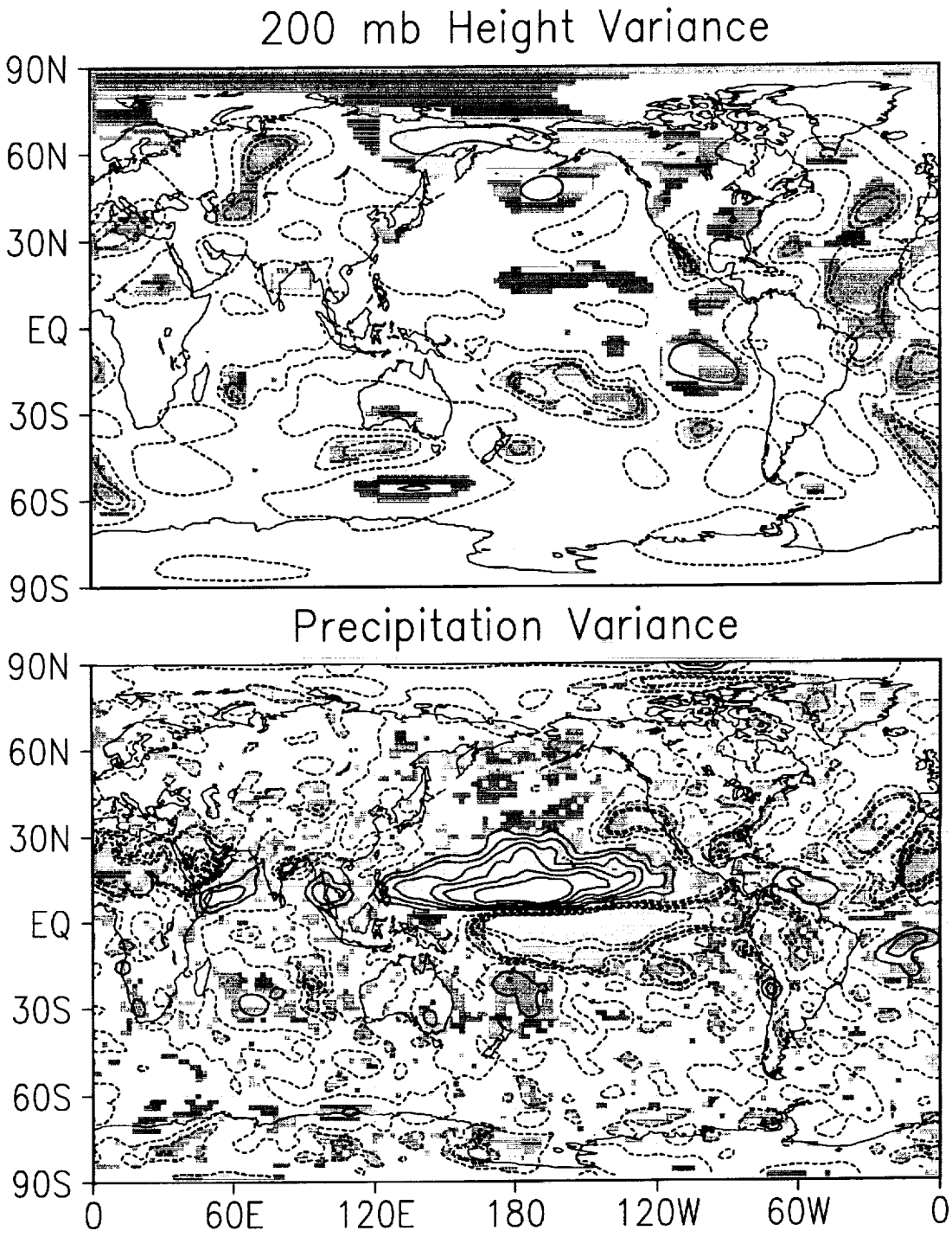

Figure 69: Ratio of the intra-ensemble variance of seasonal means of 200mb geopotential height and precipitation (1989 / 1983). The ensembles consist of 36 members for 1983 and 18 members for 1989. Shading indicates values are significantly different from one at the two-sided 5% level based on an F test. Contour levels are ... $\frac{1}{8}\frac{1}{4}\frac{1}{2}$ 2 4 8.... Contours less than 1 are dashed. Note the 1 contour is omitted.

PROBABILISTIC VERIFICATION

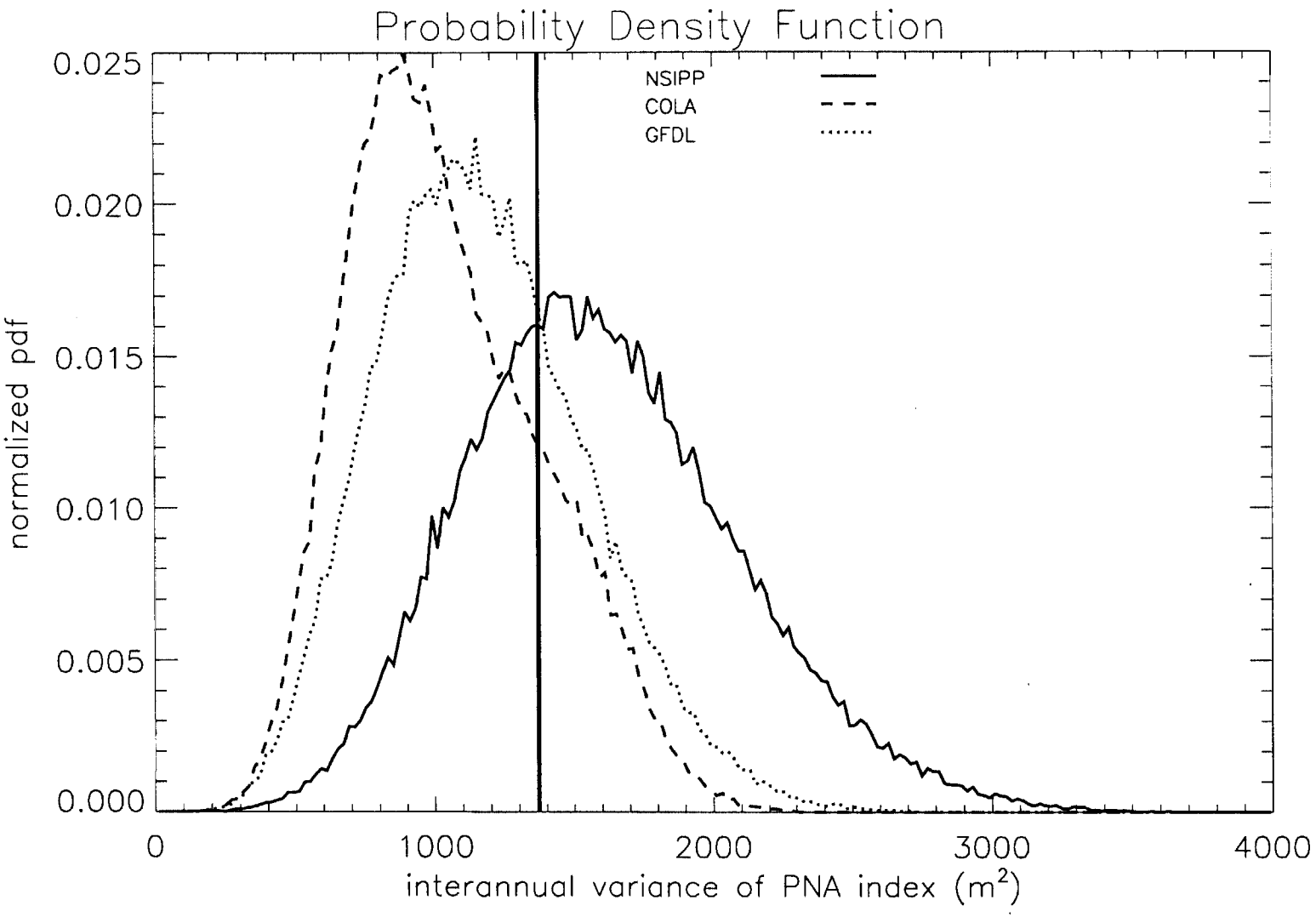

Figure 70: Probability distribution function of the variance of the PNA index as defined on page 4 . Each variance is computed from synthetic realizations generated by randomly sampling the ensembles for the 15-year period 1982 to 1996. Distributions are based on 100,000 of the 9^{15} possible combinations. The vertical line is the observed variance for this period.

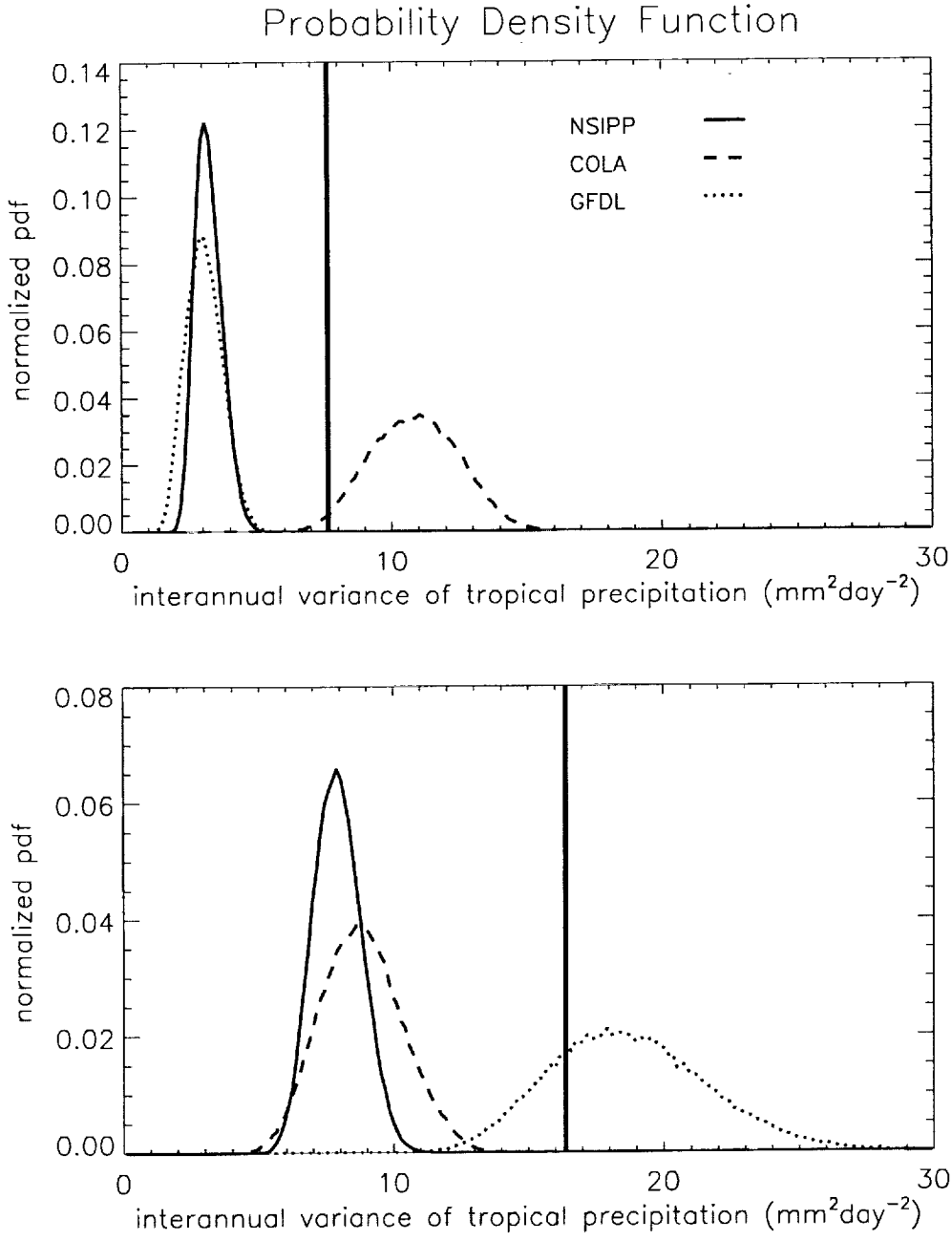

Figure 71: Probability distribution function of the variance of the area-mean precipitation in the tropical Pacific. The top figure is for the region (0-8° N, 160-200° E) and the bottom figure for (0-8° S, 160-200° E). Each variance is computed from synthetic realizations generated by randomly sampling the ensembles for the 15-year period 1982 to 1996. Distributions are based on 100,000 of the 9^{15} possible combinations. The vertical line is the observed variance for this period.

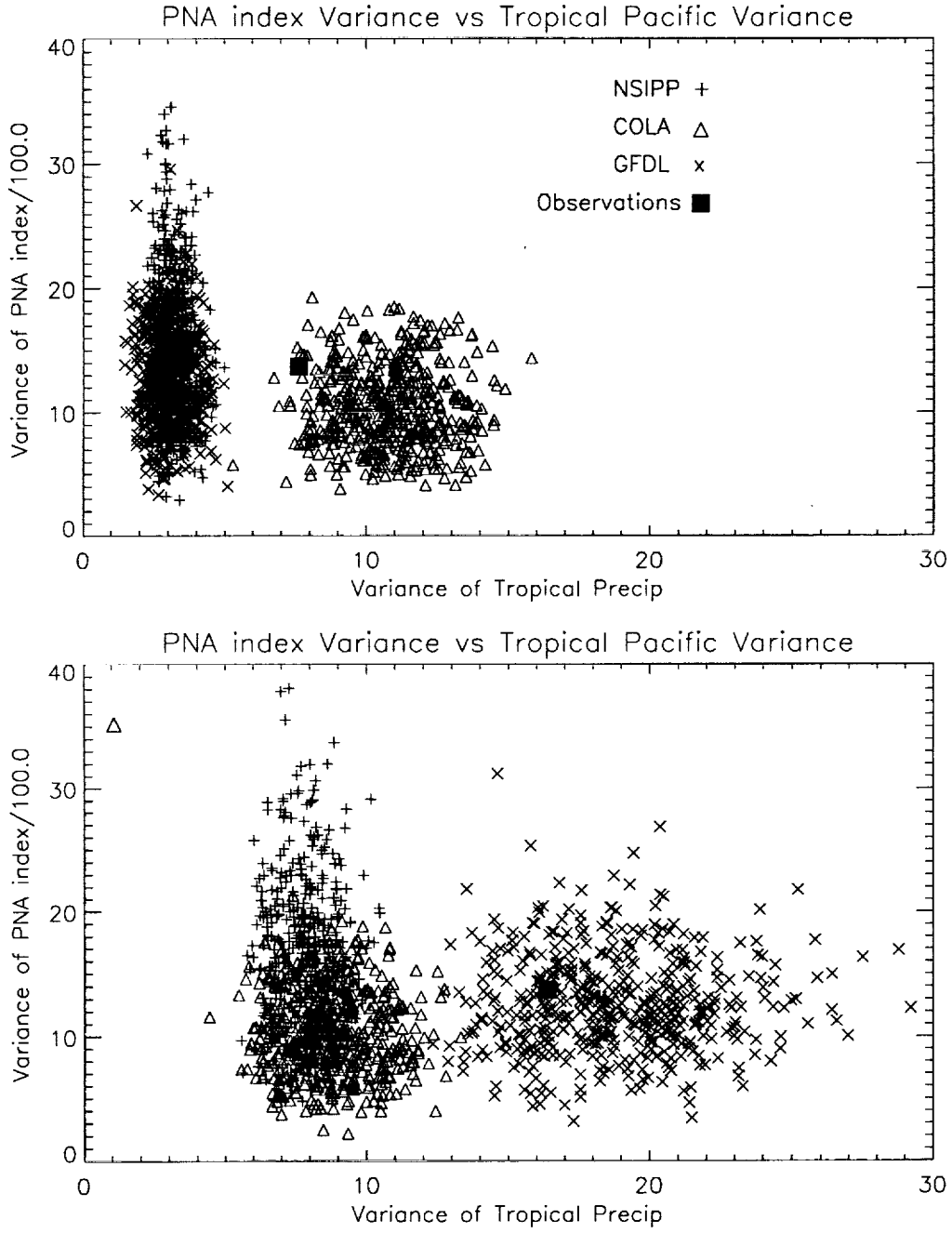

Figure 72: Scatter plot of precipitation and PNA index variance using sub-sampled synthetic realizations, as in the pdfs shown in the two previous figures. Each point represents the precipitation and PNA index variance for the same realization. The solid square marks the variances of the observations. Results are shown for the three models, as in the earlier figures.

THIS PAGE INTENTIONALLY LEFT BLANK

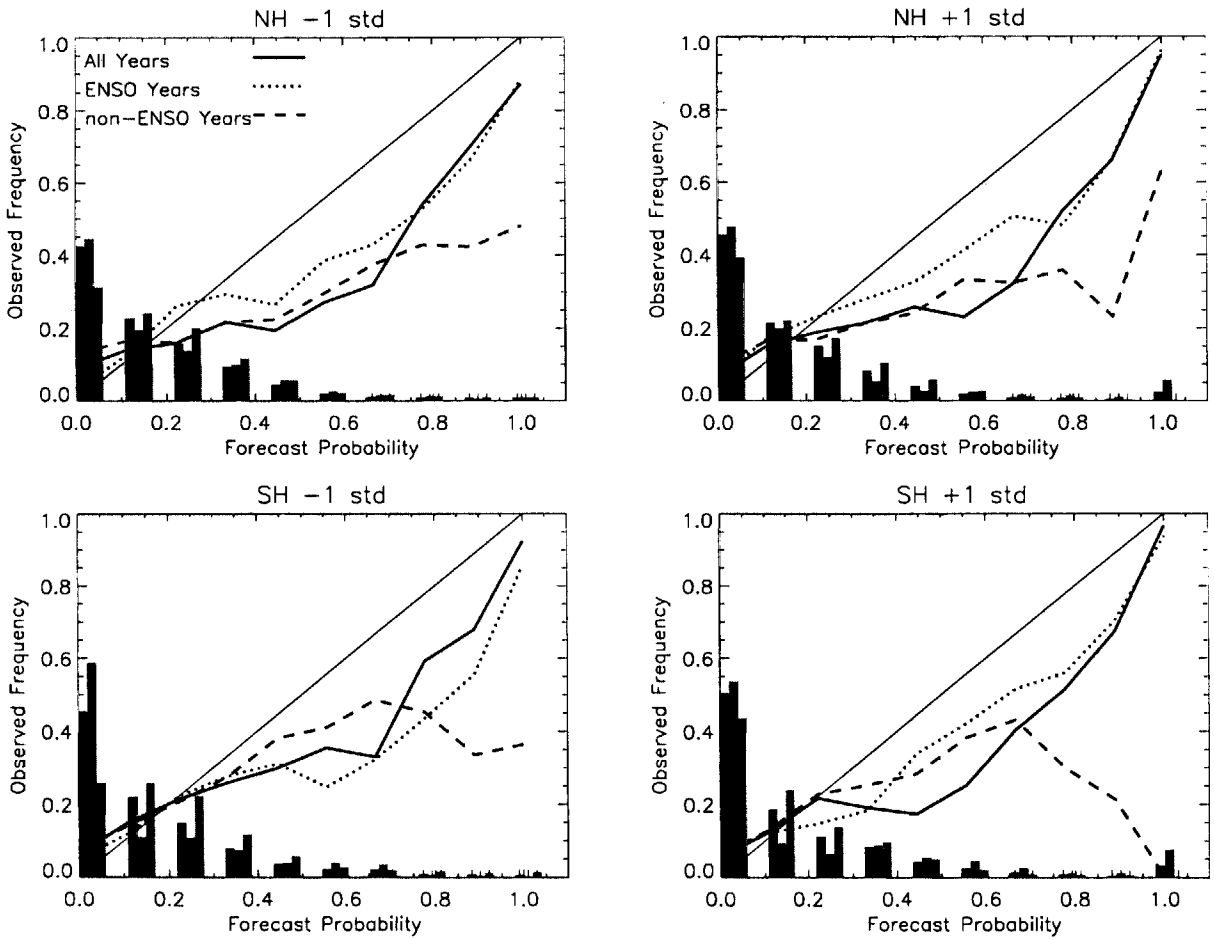

Figure 73: Reliability diagram over the two hemispheres for geopotential height at 200 mb. Right (Left) panels are for positive (negative) anomalies outside one standard deviation. The bars indicate the relative frequency of each of the probability forecasts. The darkest bars are for all years, the next lighter are for ENSO years and the lightest are for non-ENSO years.

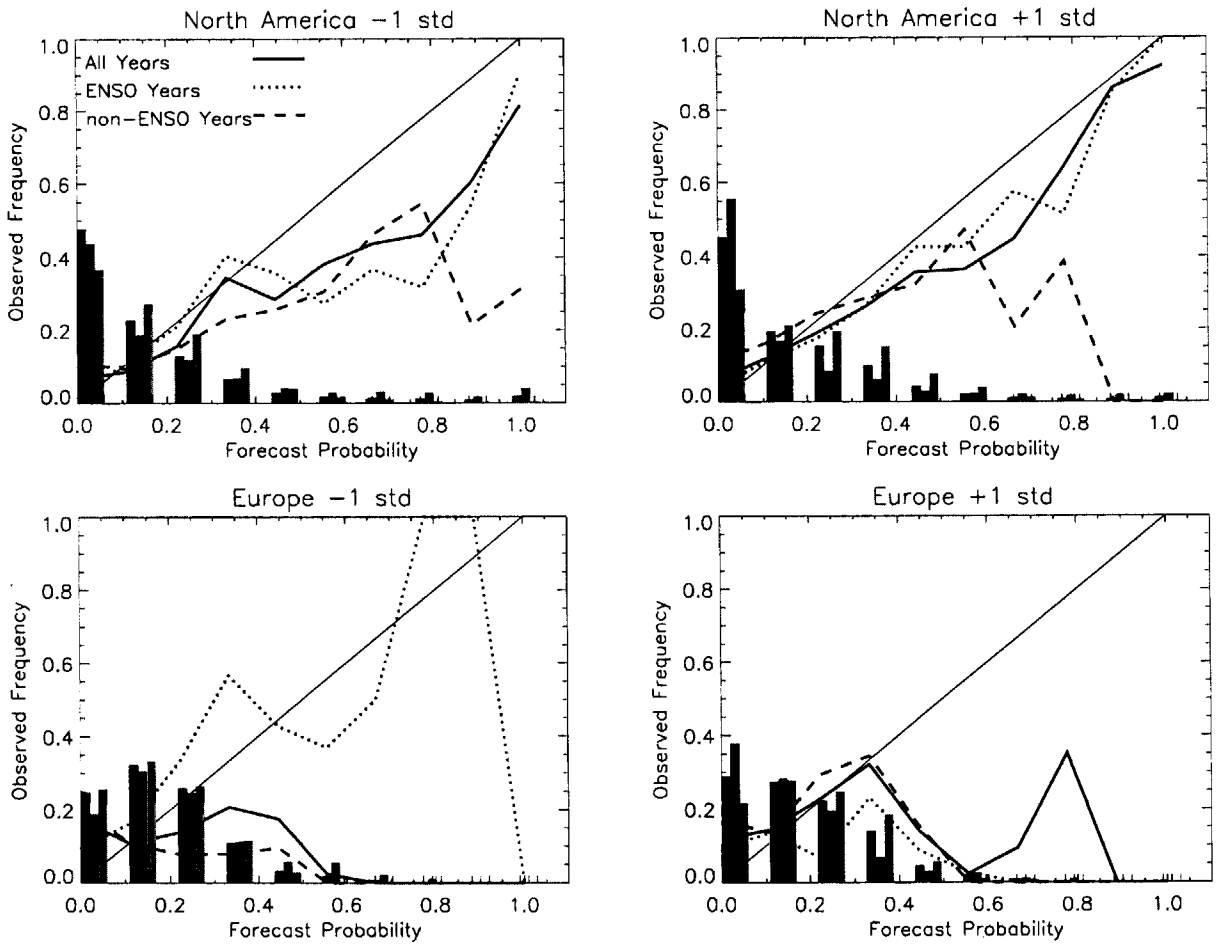

Figure 74: Reliability diagram over North America and Europe for geopotential height at 200 mb. See facing figure for explanation.

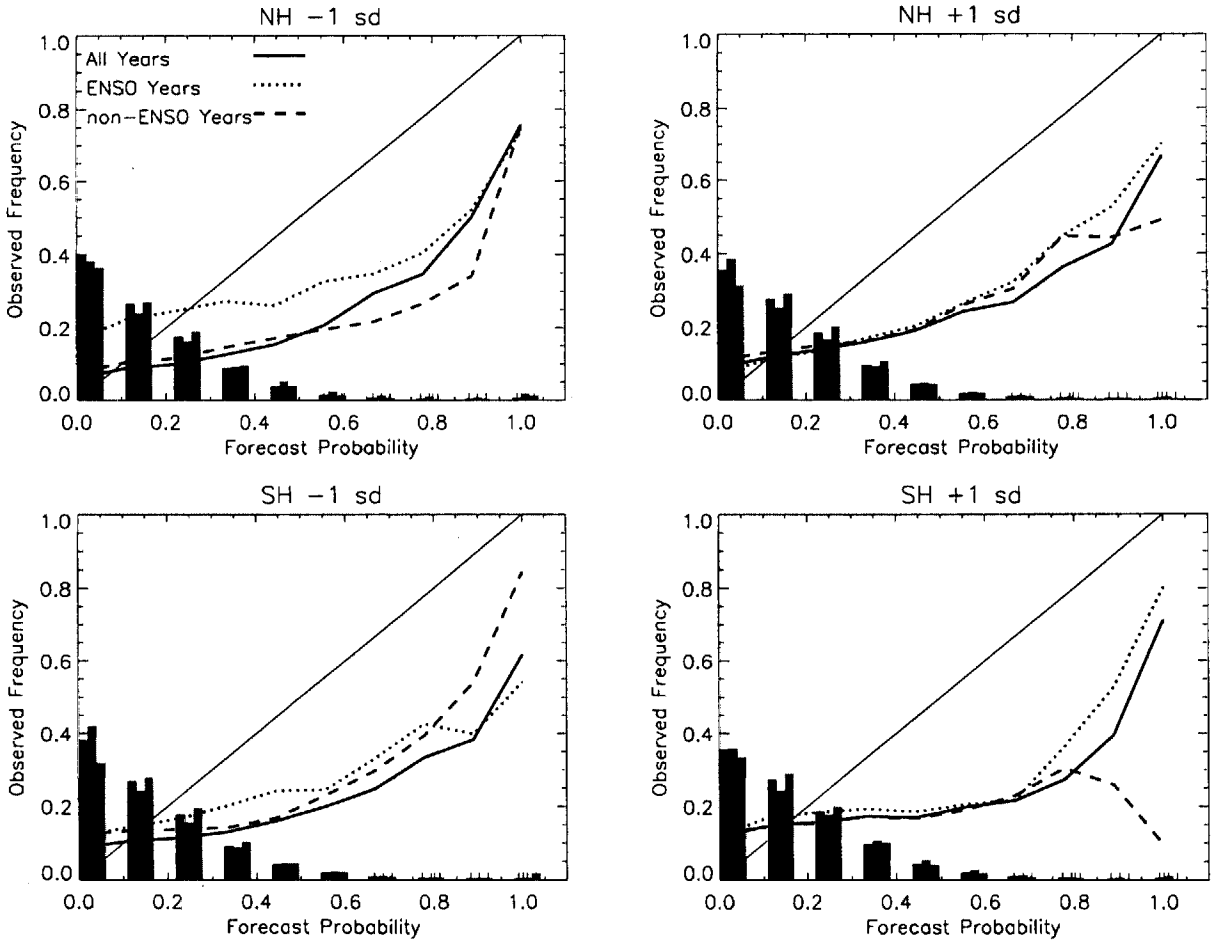

Figure 75: Reliability diagram over the two hemispheres for precipitation. Right (Left) panels are for positive (negative) anomalies outside one standard deviation. The bars indicate the relative frequency of each of the probability forecasts. The darkest bars are for all years, the next lighter are for ENSO years and the lightest are for non-ENSO years.

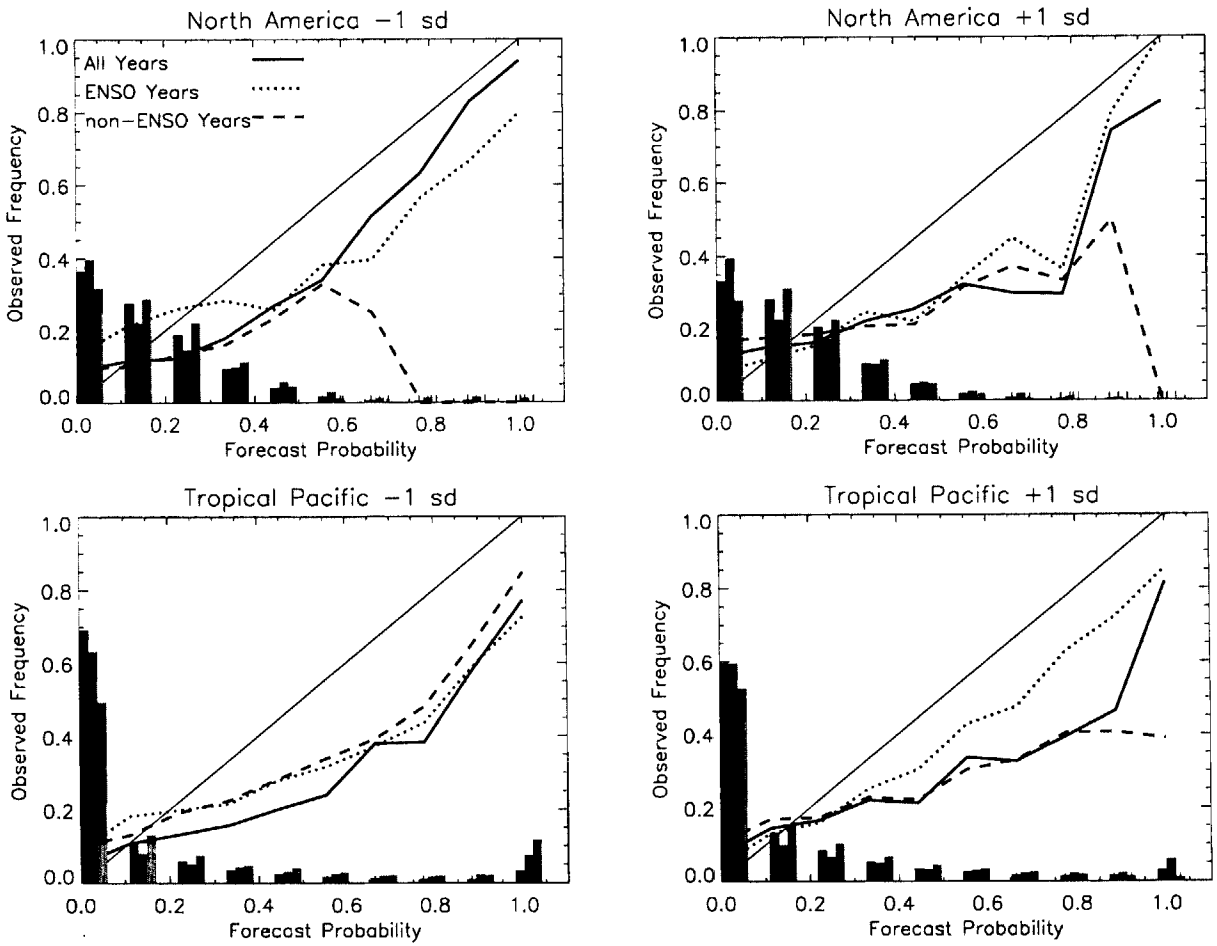

Figure 76: Reliability diagram over North America and tropical Pacific for precipitation. See facing figure for explanation.

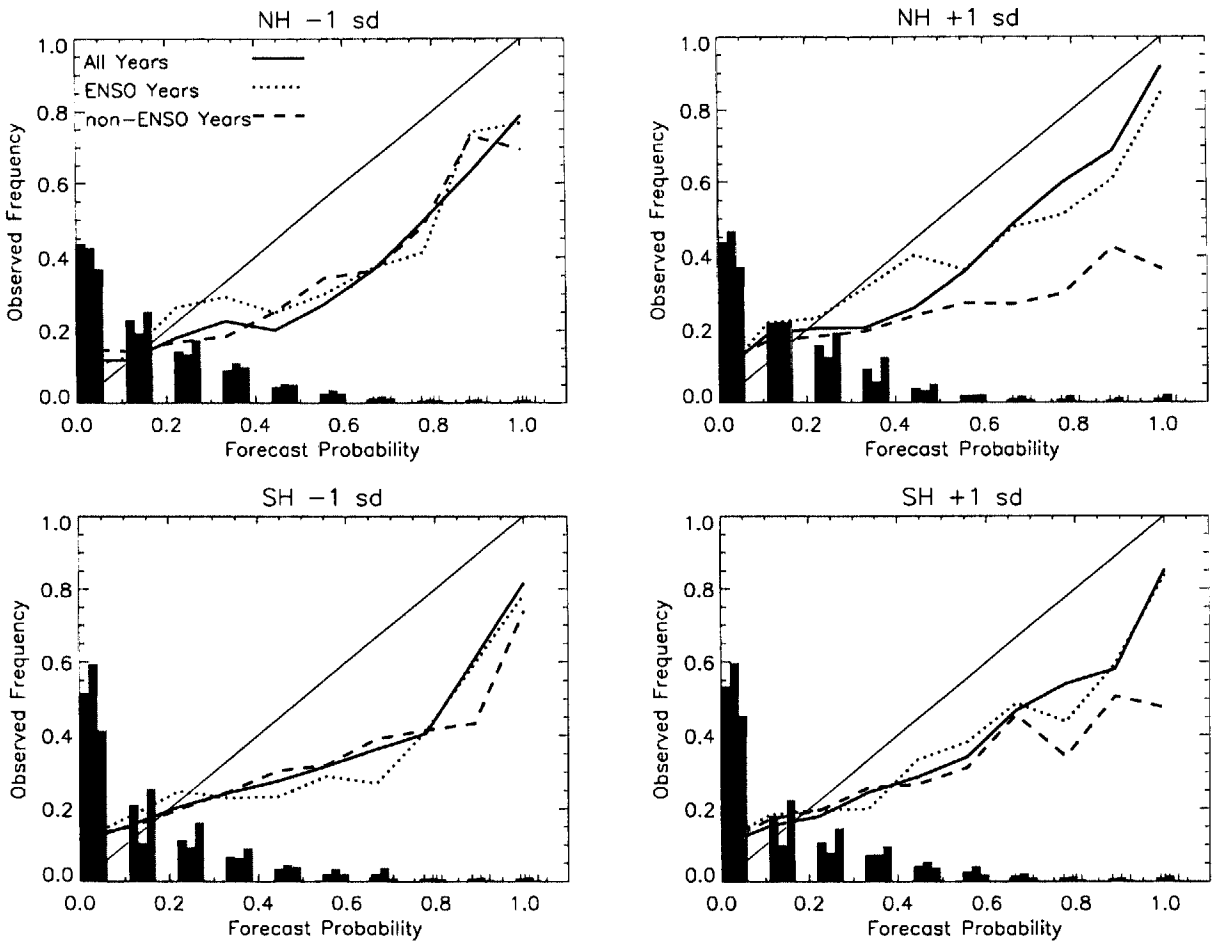

Figure 77: Reliability diagram over the two hemispheres for sea-level pressure. Right (Left) panels are for positive (negative) anomalies outside one standard deviation. The bars indicate the relative frequency of each of the probability forecasts. The darkest bars are for all years, the next lighter are for ENSO years and the lightest are for non-ENSO years.

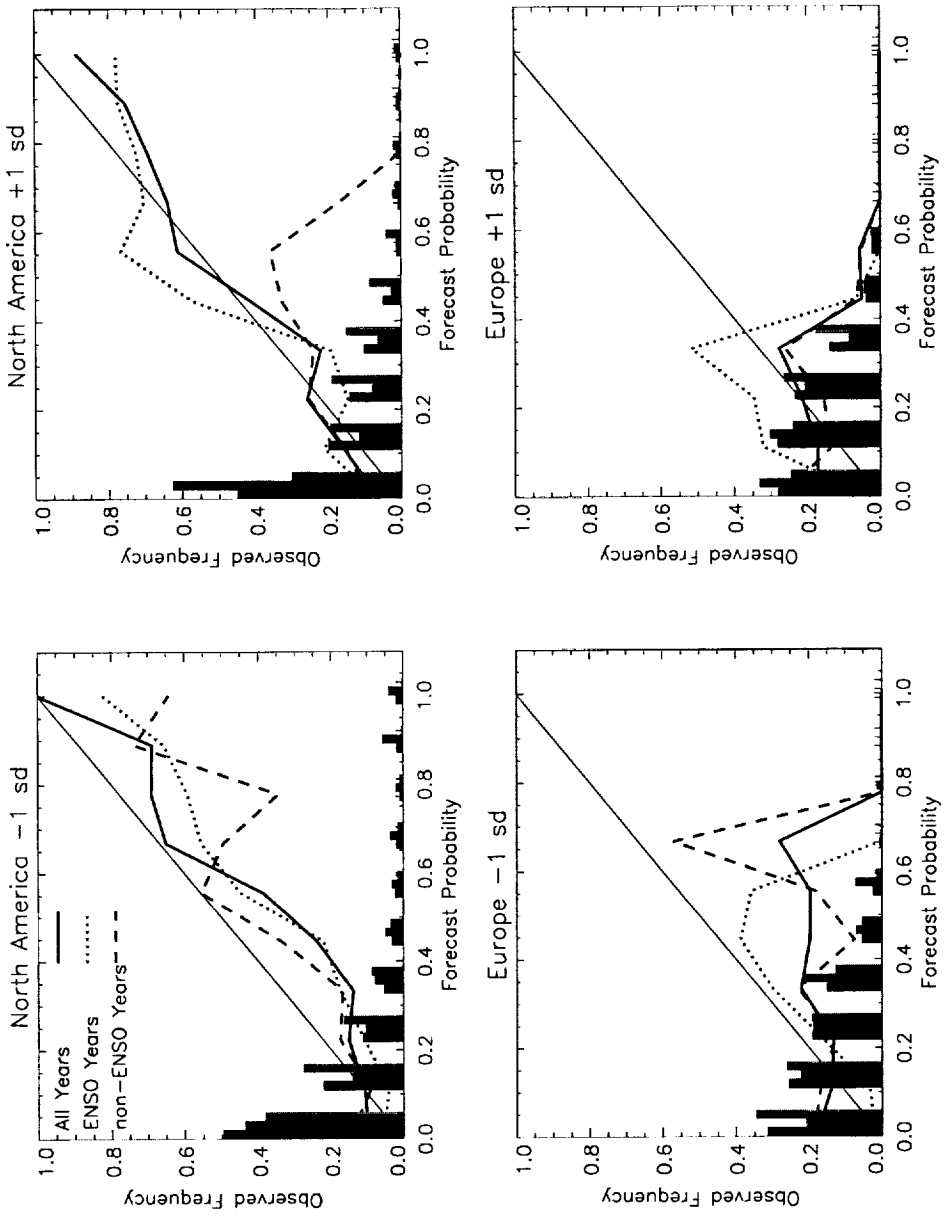

Figure 78: Reliability diagram over North America and Europe for sea-level pressure. See facing figure for explanation.

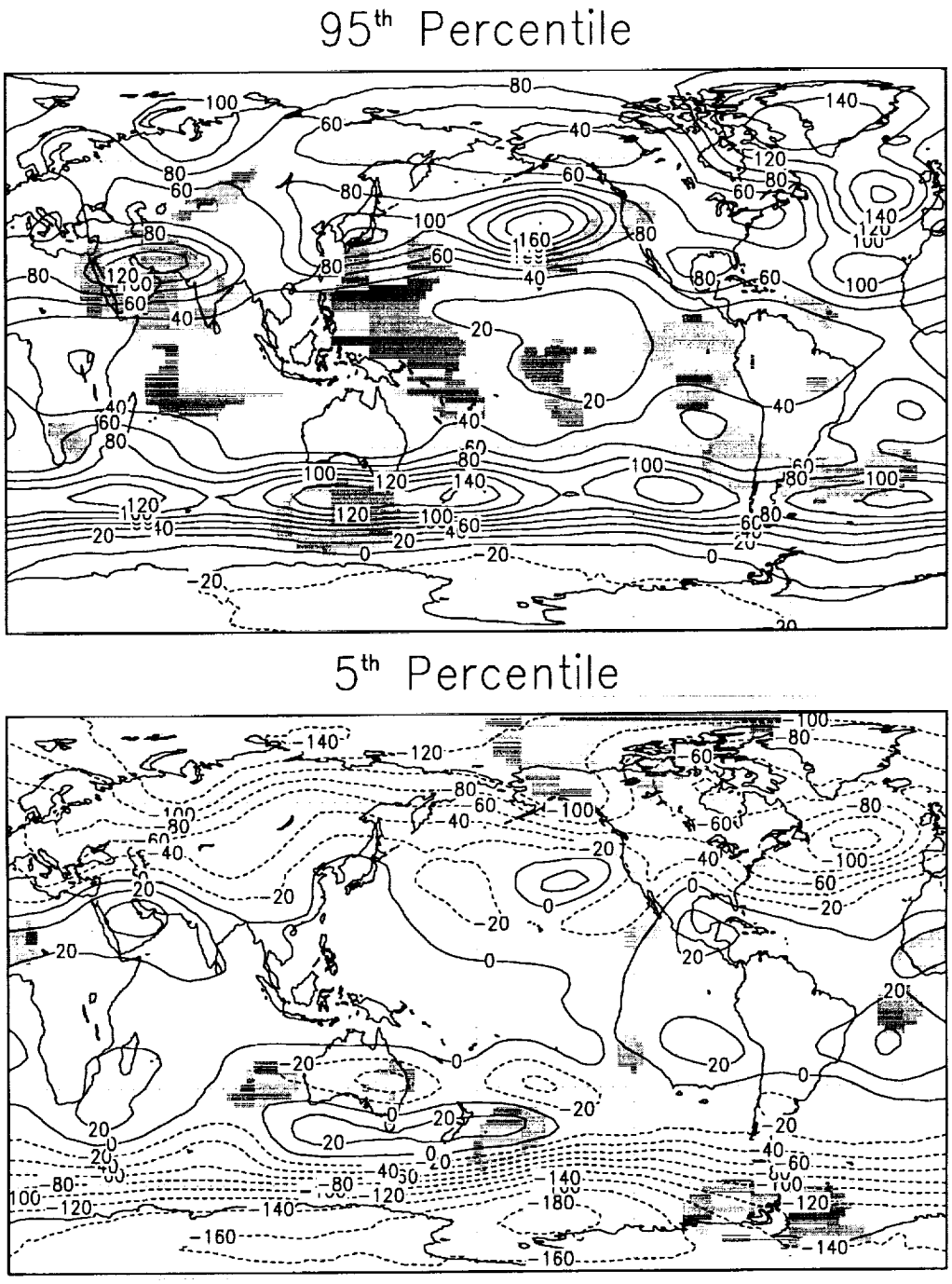

Figure 79: The 5th and 95th percentiles of the local distribution of the difference of geopotential height at 200 mb (1998 minus 1983). The contour interval is 20 m. The shading indicated regions where the observations fall outside these percentiles. Observations are above the 95th percentile over 14.7% of the globe, and below the 5th percentile over 2.4%.

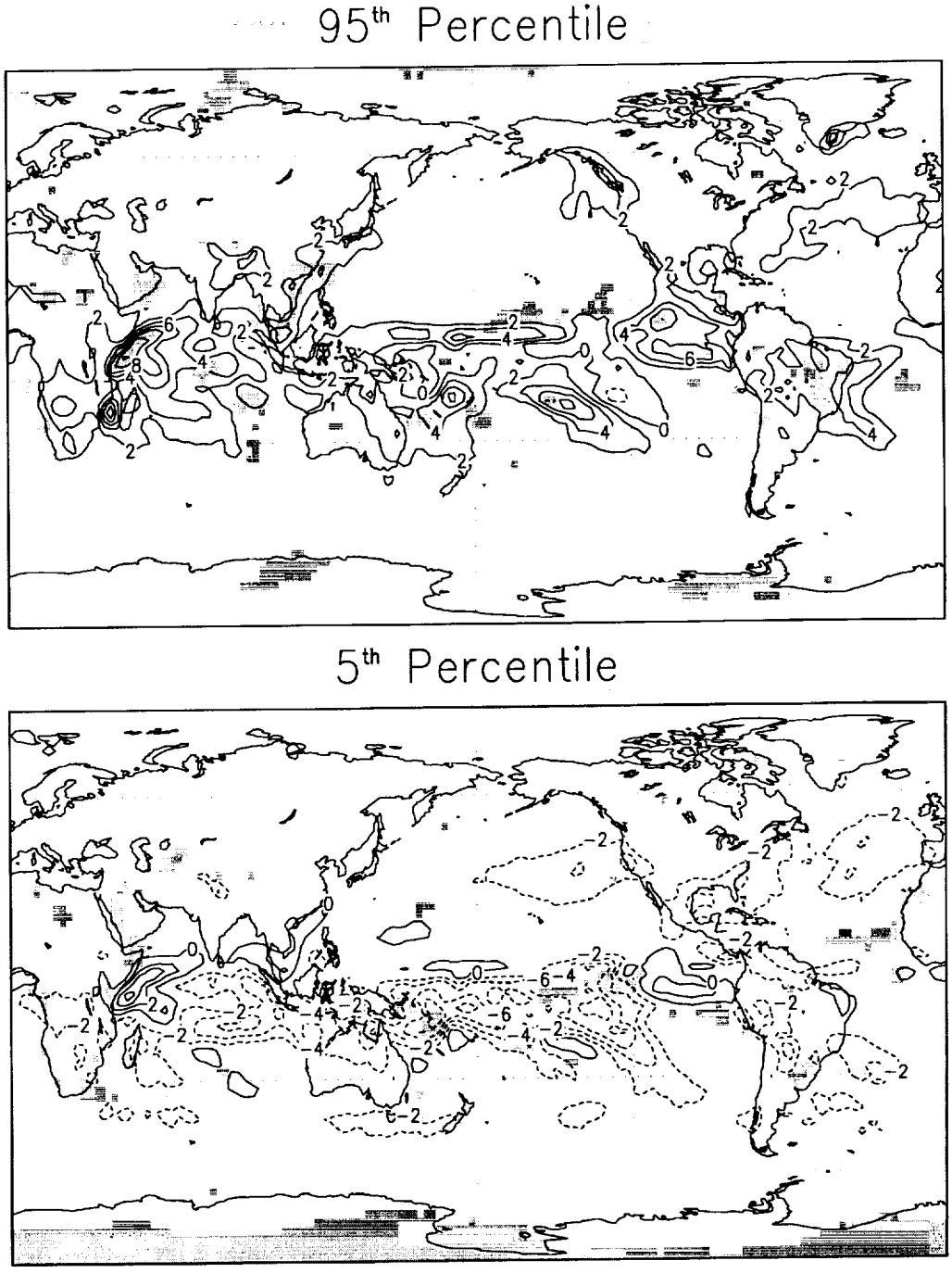

Figure 80: The 5th and 95th percentiles of the local distribution of the difference of precipitation (1998 minus 1983). The contour interval is 2 mm/day. The shading indicated regions where the observations fall outside these percentiles. Observations are above the 95th percentile over 3.61% of the globe, and below the 5th percentile over 3.88%.

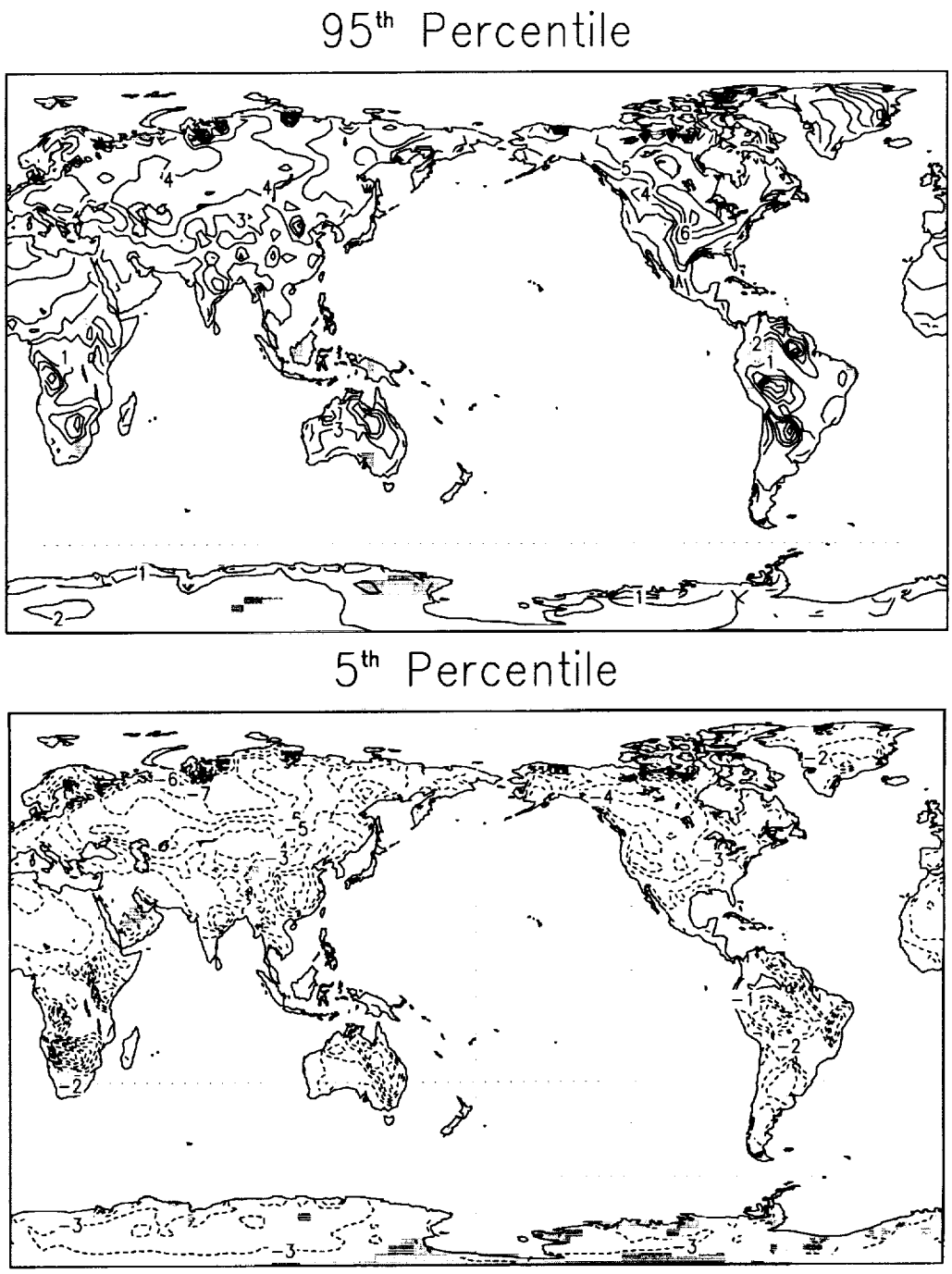

Figure 81: The 5th and 95th percentiles of the local distribution of the difference of surface temperature (1998 minus 1983). The contour interval is 1 K. The shading indicated regions where the observations fall outside these percentiles. Observations are above the 95th percentile over 2.35% of the land area, and below the 5th percentile over 2.2%.

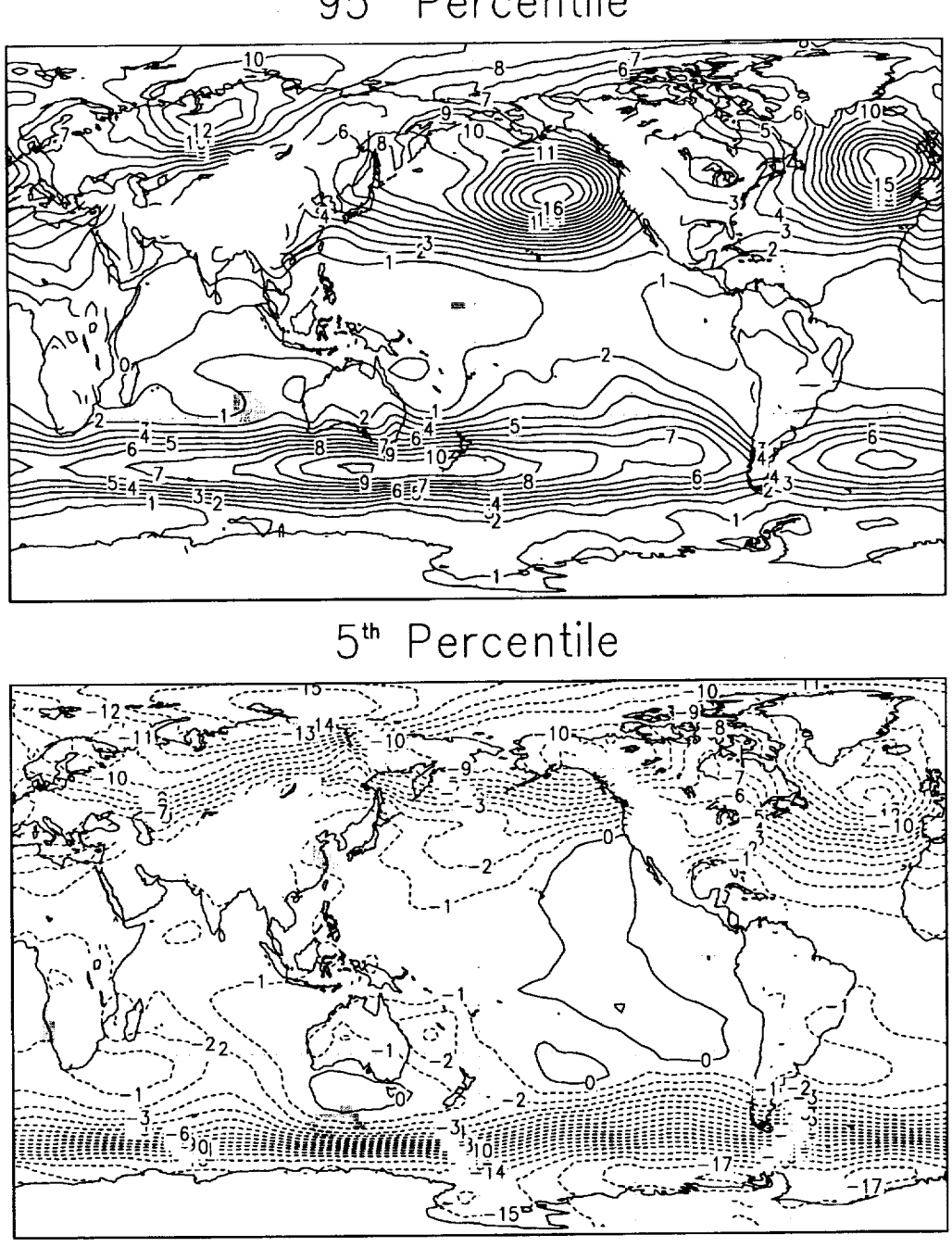

Figure 82: The 5th and 95th percentiles of the local distribution of the difference of sea-level pressure (1998 minus 1983). The contour interval is 2 mb. The shading indicated regions where the observations fall outside these percentiles. Observations are above the 95th percentile over 0.26% of the regions below a kilometer, and below the 5th percentile over 0.62%.

Previous Volumes in This Series

Volume 1 *September 1994*	Documentation of the Goddard Earth Observing System (GEOS) general circulation model - Version 1 **L.L. Takacs, A. Molod, and T. Wang**
Volume 2 *October 1994*	Direct solution of the implicit formulation of fourth order horizontal diffusion for gridpoint models on the sphere **Y. Li, S. Moorthi, and J.R. Bates**
Volume 3 *December 1994*	An efficient thermal infrared radiation parameterization for use in general circulation models **M.-D. Chou and M.J. Suarez**
Volume 4 *January 1995*	Documentation of the Goddard Earth Observing System (GEOS) Data Assimilation System - Version 1 **James Pfaendtner, Stephen Bloom, David Lamich, Michael Seablom, Meta Sienkiewicz, James Stobie, and Arlindo da Silva**
Volume 5 *April 1995*	Documentation of the Aries-GEOS dynamical core: Version 2 **Max J. Suarez and Lawrence L. Takacs**
Volume 6 *April 1995*	A Multiyear Assimilation with the GEOS-1 System: Overview and Results **Siegfried Schubert, Chung-Kyu Park, Chung-Yu Wu, Wayne Higgins, Yelena Kondratyeva, Andrea Molod, Lawrence Takacs, Michael Seablom, and Richard Rood**
Volume 7 *September 1995*	Proceedings of the Workshop on the GEOS-1 Five-Year Assimilation **Siegfried D. Schubert and Richard B. Rood**
Volume 8 *March 1996*	Documentation of the Tangent Linear Model and Its Adjoint of the Adiabatic Version of the NASA GEOS-1 C-Grid GCM: Version 5.2 **Weiyu Yang and I. Michael Navon**

Volume 9 *March 1996*	Energy and Water Balance Calculations in the Mosaic LSM **Randal D. Koster and Max J. Suarez**
Volume 10 *April 1996*	Dynamical Aspects of Climate Simulations Using the GEOS General Circulation Model **Lawrence L. Takacs and Max J. Suarez**
Volume 11 *May 1997*	Documentation of the Tangent Linear and its Adjoint Models of the Relaxed Arakawa-Schubert Moisture Parameterization Package of the NASA GEOS-1 GCM (Version 5.2) **Weiyu Yang I. Michael Navon, and Ricardo Todling**
Volume 12 *August 1997*	Comparison of Satellite Global Rainfall Algorithms **Alfred T.C. Chang and Long S. Chiu**
Volume 13 *December 1997*	Interannual Variability and Potential Predictability in Reanalysis Products **Wie Ming and Siegfried D. Schubert**
Volume 14 *August 1998*	A Comparison of GEOS Assimilated Data with FIFE Observations **Michael G. Bosilovich and Siegfried D. Schubert**
Volume 15 *June 1999*	A Solar Radiation Parameterization for Atmospheric Studies **Ming-Dah Chou and Max J. Suarez**
Volume 16 *November 1999*	Filtering Techniques on a Stretched Grid General Circulation Model **Lawrence Takacs, William Sawyer, Max J. Suarez, and Michael S. Fox-Rabinowitz**
Volume 17 *July 2000*	Atlas of Seasonal Means Simulated by the NSIPP-1 Atmospheric GCM **Julio T. Bacmeister, Philip J. Pegion, Siegfried D. Schubert, and Max J. Suarez**

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 2000	Technical Memorandum

4. TITLE AND SUBTITLE
Technical Report Series on Global Modeling and Data Assimilation
Volume 18—An Assessment of the Predictability of Northern Winter Seasonal Means with the NSIPP 1 AGCM

5. FUNDING NUMBERS
Code 971
971-622-24-47-26

6. AUTHOR(S)
Philip J. Pegion, Siegfried D. Schubert, and Max J. Suarez

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS (ES)
Oceans and Ice Branch
Goddard Space Flight Center
Greenbelt, Maryland 20771

8. PEFORMING ORGANIZATION REPORT NUMBER
2000–

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS (ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
TM-2000-104606, Vol. 18

11. SUPPLEMENTARY NOTES
P.J. Pegion: General Sciences Corporation, Laurel, Maryland

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Unclassified–Unlimited
Subject Category: 46
Report available from the NASA Center for AeroSpace Information,
7121 Standard Drive, Hanover, MD 21076-1320. (301) 621-0390.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This atlas assesses the predictability of January-February-March (JFM) means using version 1 of the NASA Seasonal-to-Interannual Prediction Project Atmospheric General Circulation Model (the NSIPP 1 AGCM). The AGCM is part of the NSIPP coupled atmosphere-land-ocean model. For these results, the atmosphere was run uncoupled from the ocean, but coupled with an interactive land model. The results are based on 20 ensembles of nine JFM hindcasts for the period 1980-1999, with sea surface temperature (SST) and sea ice specified from observations. The model integrations were started from initial atmospheric conditions (taken from NCEP/NCAR reanalyses) centered on December 15. The analysis focuses on 200 mb height, precipitation, surface temperature, and sea-level pressure. The results address issues of both predictability and forecast skill. Various signal-to-noise measures are computed to demonstrate the potential for skillful prediction on seasonal time scales under the assumption of a perfect model and perfectly known oceanic boundary forcings. The results show that the model produces a realistic ENSO response in both the tropics and extratropics.

14. SUBJECT TERMS
Predictability, Atmospheric Models, Seasonal Forecasts, Interannual Variability, Signal to Noise Ratios

15. NUMBER OF PAGES
117

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL