NASA Technical Memorandum 104606, Vol. 6

Technical Report Series on Global Modeling and Data Assimilation

Max J. Suarez, Editor

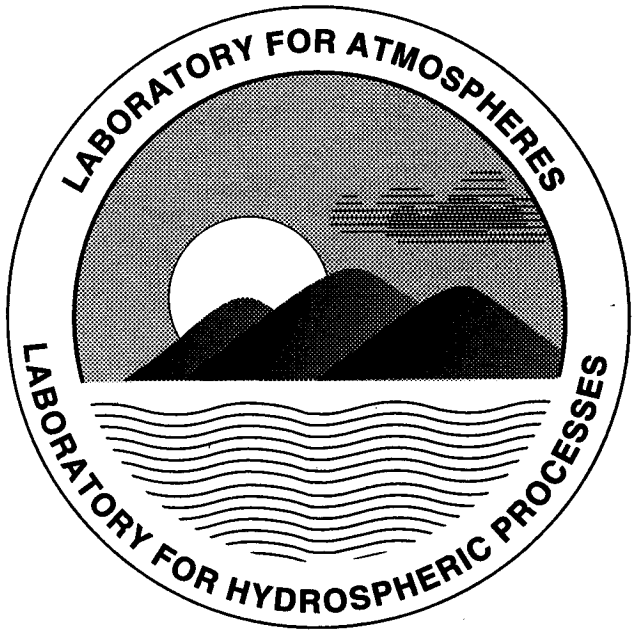
Volume 6

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

N95-26423
Unclas
G3/46 0048317
(NASA-TM-104606-Vol-6) TECHNICAL REPORT SERIES ON GLOBAL MODELING AND DATA ASSIMILATION. VOLUME 6: A MULTIYEAR ASSIMILATION WITH THE GEOS-1 SYSTEM: OVERVIEW AND RESULTS (NASA. Goddard Space Flight Center) 201 p

April 1995



DATA ASSIMILATION OFFICE

NASA Technical Memorandum 104606, Vol. 6

Technical Report Series on Global Modeling and Data Assimilation

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

Volume 6

A Multiyear Assimilation With the GEOS-1 System: Overview and Results

Siegfried Schubert
Richard Rood
Goddard Space Flight Center
Greenbelt, Maryland

Chung-Kyu Park
Joint Center for Earth System Science
University of Maryland
College Park, Maryland

Chung-Yu Wu
Yelena Kondratyeva
Andrea Molod
Lawrence Takacs
Michael Seablom
General Sciences Corporation
Laurel, Maryland

Wayne Higgins
Climate Analysis Center
NMC/NWS/NOAA
Washington, D. C.

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland
1995

This publication is available from the NASA Center for AeroSpace Information,
800 Elkridge Landing Road, Linthicum Heights, MD 21090-2934, (301) 621-0390.

Abstract

The Data Assimilation Office (DAO) at Goddard Space Flight Center has produced a multi–year global assimilated data set with version 1 of the Goddard Earth Observing System Data Assimilation System (GEOS–1 DAS). One of the main goals of this project, in addition to benchmarking the GEOS–1 system, was to produce a research quality data set suitable for the study of short–term climate variability. The output, which is global and gridded, includes all prognostic fields and a large number of diagnostic quantities such as precipitation, latent heating, and surface fluxes. Output is provided four times daily with selected quantities available eight times per day. Information about the observations input to the GEOS–1 DAS is provided in terms of maps of spatial coverage, bargraphs of data counts and tables of all time periods with significant data gaps.

The purpose of this document is to serve as a users guide to NASA's first multi–year assimilated data set and to provide an early look at the quality of the output. Documentation is provided on all the data archives, including sample read programs and methods of data access. Extensive comparisons are made with the corresponding operational European Center for Medium–Range Weather Forecasts analyses, as well as, various *in situ* and satellite observations. This document is also intended to alert users of the data about potential limitations of assimilated data in general, and the GEOS–1 data in particular. Results are presented for the period March 1985–Feb 1990.

Contents

List of Figures vii

1 Introduction 1

2 Assimilated Data: The Role of the Model 4

3 Overview of GEOS–1 DAS 6

3.1 The OI Scheme 6

3.2 The GEOS–1 Model 7

3.3 The Processing System 9

4 Results 12

4.1 Comparison Datasets 12

4.2 Climate Mean Quantities 14

4.2.1 Zonal Mean Prognostic 14

4.2.2 Global Prognostic 27

4.2.3 Surface Energy Balance 46

4.2.4 Precipitation and the U.S. Moisture Budget 63

4.3 Intraseasonal Variability 80

4.3.1 Global RMS Differences 80

4.3.2 The Tropics 85

4.3.3 Correlations 92

4.4 Interannual Variability 97

4.4.1 Space–Time 97

PRECEDING PAGE BLANK NOT FILMED

4.4.2 Global Distribution (1988–1987) 102
4.4.3 Precipitation Anomalies 112
4.5 Concluding Remarks 125

5 Input Observations **127**
5.1 Spatial Coverage 127
5.2 Inventory 132
5.2.1 Sample Bargraphs of Data Counts 132
5.2.2 Tables of Significant Data Gaps 137

6 Datasets and Data Access **145**
6.1 NCCS Datasets 146
6.1.1 Overview 146
6.1.2 Details of Assimilation Output 149
6.1.3 NCCS Unitree Directory Structure 163
6.1.4 Sample Read Program (Phoenix Format) 164
6.2 GDAAC Time Series Datasets 167
6.2.1 Description of the Datasets 168
6.2.2 Menu of Available Quantities 170
6.3 Datasets Available from DAO Server 174

Appendix: History of Output Problems **177**

References **179**

List of Figures

1 A schematic of the variability found in operational analyses. The spurious climate signals produced by changes in the data assimilation systems are found primarily at the seasonal and longer time scales. 2

2 The main elements of the processing system and data flow of the GEOS–1 DAS. 11

3 (a) Seasonal cycle of the 200 mb zonal mean u–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. 16

4 (a) Seasonal cycle of the 200 mb zonal mean v–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 0.5 m/sec. Negative regions are shaded. 17

5 (a) Seasonal cycle of the 300 mb zonal mean height from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 100.0 m in (a) and 10.0 m in (b). The values greater than 9700.0 m are shaded in (a) and negative regions are shaded in (b). 18

6 (a) DJF climatology of zonal mean u–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. The top level is 30 mb . . . 19

7 (a) JJA climatology of zonal mean u–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. The top level is 30 mb . . . 20

8 (a) DJF climatology of zonal mean v–wind from the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 0.5 m/sec. Negative regions are shaded. The top level is 30 mb 21

9 (a) JJA climatology of zonal mean v–wind from the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 0.5 m/sec. Negative regions are shaded. The top level is 30 mb 22

10 (a) DJF climatology of zonal mean temperature from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 $°C$ in (a) and 1.0 $°C$ in (b). Negative regions are shaded. The top level is 30 mb . . . 23

11 (a) JJA climatology of zonal mean temperature from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 $°C$ in (a) and 1.0 $°C$ in (b). Negative regions are shaded. The top level is 30 mb . . . 24

12 DJF climatology of zonal mean vertical–component of wind from (a) the GEOS–DAS and (b) from the ECMWF analysis. Contour interval is 40.0 nb/sec. Negative regions are shaded. 25

13 JJA climatology of zonal mean vertical–component of wind from (a) the GEOS–DAS and (b) from the ECMWF analysis. Contour interval is 40.0 nb/sec. Negative regions are shaded. Negative regions are shaded. 26

14 (a) DJF climatology of 200 mb u–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 m/sec in (a) and 2.0 m/sec in (b). Negative regions are shaded 28

15 (a) DJF climatology of 850 mb u–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 3.0 m/sec in (a) and 2.0 m/sec in (b). Negative regions are shaded. 29

16 (a) JJA climatology of 200 mb u–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 m/sec in (a) and 2.0 m/sec in (b). Negative regions are shaded. 30

17 (a) JJA climatology of 850 mb u–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 3.0 m/sec in (a) and 2.0 m/sec in (b). Negative regions are shaded. 31

18 (a) DJF climatology of 200 mb v–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 2.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. 32

19 (a) DJF climatology of 850 mb v–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 2.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. 33

20 (a) JJA climatology of 200 mb v–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 2.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. 34

21 (a) JJA climatology of 850 mb v–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 2.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. 35

22 (a) DJF climatology of 200 mb eddy streamfunction from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0×10^6 m^2/sec in (a) and 2.0×10^6 m^2/sec in (b). The values greater than 10.0×10^6 m^2/sec and less than -10.0×10^6 m^2/sec are shaded in (a) and the values greater than 2.0×10^6 m^2/sec and less than -2.0×10^6 m^2/sec are shaded in (b). 36

23 (a) JJA climatology of 200 mb eddy streamfunction from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0×10^6 m^2/sec in (a) and 2.0×10^6 m^2/sec in (b). The values greater than 10.0×10^6 m^2/sec and less than -10.0×10^6 m^2/sec are shaded in (a) and the values greater than 2.0×10^6 m^2/sec and less than -2.0×10^6 m^2/sec are shaded in (b). 37

24 (a) DJF climatology of 200 mb velocity potential from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 1.0×10^6 m^2/sec. The values greater than 5.0×10^6 m^2/sec and less than -5.0×10^6 m^2/sec are shaded in (a) and the values greater than 3.0×10^6 m^2/sec and less than -3.0×10^6 m^2/sec are shaded in (b). 38

25 (a) DJF climatology of 850 mb velocity potential from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 1.0×10^6 m^2/sec. The values greater than 5.0×10^6 m^2/sec and less than -5.0×10^6 m^2/sec are shaded in (a) and the values greater than 2.0×10^6 m^2/sec and less than -2.0×10^6 m^2/sec are shaded in (b). 39

26 (a) JJA climatology of 200 mb velocity potential from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 1.0×10^6 m^2/sec. The values greater than 5.0×10^6 m^2/sec and less than -5.0×10^6 m^2/sec are shaded in (a) and the values greater than 3.0×10^6 m^2/sec and less than -3.0×10^6 m^2/sec are shaded in (b). 40

27 (a) JJA climatology of 850 mb velocity potential from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 1.0×10^6 m^2/sec. The values greater than 5.0×10^6 m^2/sec and less than -5.0×10^6 m^2/sec are shaded in (a) and the values greater than 2.0×10^6 m^2/sec and less than -2.0×10^6 m^2/sec are shaded in (b). 41

28 (a) DJF climatology of sea–level pressure from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour intervals are 4.0 mb. Values greater than 1020 mb are shaded in (a). Negative regions are shaded in (b). 42

29 (a) JJA climatology of sea–level pressure from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 4.0 *mb*. Values greater than 1020 mb are shaded in (a). Negative regions are shaded in (b). 43

30 (a) DJF climatology (Dec 1987–Feb 1990) of total precipitable water from the GEOS–DAS and (b) the difference (GEOS–SSM/I) over the oceans. Contour interval is 0.5 gm/cm^2. The values greater than 4.0 gm/cm^2 are shaded in (a) and negative values are shaded in (b). 44

31 (a) JJA climatology of total precipitable water (Jul 1987–Aug 1989) from the GEOS–DAS and (b) the difference (GEOS–SSM/I) over the oceans. Contour interval is 0.5 gm/cm^2. The values greater than 4.0 gm/cm^2 are shaded in (a) and negative values are shaded in (b). 45

32 The latent heat flux for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Contour interval is 30 W/m^2. 50

33 The latent heat flux for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Contour interval is 30 W/m^2. 51

34 The sensible heat flux for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour intervals are -5 0 5 10 20 30 60 90 120 150 180 in (a) and every 5 in (b). 52

35 The sensible heat flux for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour intervals are -5 0 5 10 20 30 60 90 120 150 180 in (a) and every 5 in (b). 53

36 The net surface shortwave radiation for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Contour interval is 30 W/m^2. 54

37 The net surface shortwave radiation for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Contour interval is 30 W/m^2. 55

38 The net surface longwave radiation for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour interval is 15 in (a) and 10 in (b). 56

39 The net surface longwave radiation for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour interval is 15 in (a) and 10 in (b). 57

40 The net surface heat flux for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour interval is 60 in (a) and 30 in (b). . . 58

41 The net surface heat flux for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour interval is 60 in (a) and 30 in (b). . . 59

42 The zonal mean (oceans only) sensible and latent heat flux for the GEOS–DAS, GEOS–GCM, and COADS. Units:W/m^2. 60

43 The zonal mean (oceans only) surface shortwave and longwave flux for the GEOS–DAS, GEOS–GCM, and COADS. Units:W/m^2. 61

44 The zonal mean (oceans only) net surface radiative flux and net surface heat flux for the GEOS–DAS, GEOS–GCM, and COADS. Units:W/m^2. 62

45 The GEOS–DAS precipitation climatologies for (a) DJF and (b) JJA. Units: mm/day. 64

46 The climatological mean difference (GEOS–DAS minus station observations) of precipitation for (a) DJF and (b) JJA. The contour interval is 2 mm/day. Negative values are dashed. The values greater than 1 mm/day and less than -1 mm/day are shaded. 65

47 The boxes contain the land areas over which the averages are done for the precipitation line plots. 66

48 The seasonal cycle of the regional mean precipitation over land from GEOS and the station data. Units: mm/day. 67

49 The seasonal cycle of the regional mean precipitation over land from GEOS and the station data. Units: mm/day. 68

50 Components of the vertically–integrated moisture budget from the GEOS–1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a–c–d–e). Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded. 71

51 Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a−c−d−e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 72

52 Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a−c−d−e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 73

53 Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a−c−d−e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 74

54 Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a−c−d−e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 75

55 Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a−c−d−e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 76

56 Annual cycle of vertically-integrated moisture budget terms averaged over the continental United States (30°N–50°N). The CD refers to the climate division precipitation. Units are *mm/day*. 77

57 Moisture transport into the continental United States for January through June. LLJ (Low Level Jet) refers to the region between 92.5°W and 102.5°W, and below about 850 mb. Units are 10^6 *kg/sec*. To obtain units of *cm/month* divide by 29×10^6. 78

58 Moisture transport into the continental United States for July through December. LLJ (Low Level Jet) refers to the region between 92.5°W and 102.5°W, and below about 850 mb. Units are 10^6 kg/sec. To obtain units of $cm/month$ divide by 29×10^6. 79

59 (a) DJF and (b) JJA 200 mb zonal wind rms (non–systematic) difference between the GEOS–DAS and the ECMWF analysis based on daily time–series. Contour intervals are 1 m/sec. Values greater than 6 m/sec are shaded. 82

60 (a) DJF and (b) JJA 200 mb v–wind rms (non–systematic) difference between the GEOS–DAS and the ECMWF analysis based on daily time–series. Contour intervals are 1 m/sec. Values greater than 5 m/sec are shaded. . . 83

61 (a) DJF and (b) JJA 500 mb height rms (non–systematic) difference between GEOS–DAS and ECMWF analysis based on daily time–series. Contour intervals are 10 m. Values greater than 40 m are shaded. 84

62 Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 20–70 days (1986). Contour interval is 2×10^6 m^2/sec. The negative values are shaded. 86

63 Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 20–70 days (1987). Contour interval is 2×10^6 m^2/sec. The negative values are shaded. 87

64 Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 20–70 days (1988). Contour interval is 2×10^6 m^2/sec. The negative values are shaded. 88

65 Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 20–70 days (1989). Contour interval is 2×10^6 m^2/sec. The negative values are shaded. 89

66 Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 10–20 days (1987). Contour interval is 1×10^6 m^2/sec. The negative values are shaded. 90

67 Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis at 10N. The anomalies are departures from the seasonal cycle filtered to retain periods of 2–10 days. Contour interval is 1×10^6 m^2/sec and contours between -1×10^6 m^2/sec and 1×10^6 m^2/sec are not shown. 91

68 Correlation between the 200 mb velocity potential anomalies from the GEOS–DAS and the ECMWF analysis (Mar 1986–Feb 1990), filtered for the periods of (a) 2–10 days, (b) 10–20 days, and (c) 20–70 days. Contour intervals are 0.1. The shading represents the correlations greater than 0.7 in (a), 0.8 in (b), and 0.9 in (c). 93

69 Correlation between the 200 mb velocity potential anomalies from the GEOS–DAS and the ECMWF analysis (Mar 1986–Feb 1990), filtered for the periods of 2–10 days, 10–20 days, and 20–70 days at (a) 40N, (b) the equator, and (c) 40S. 94

70 Correlation between the 500 mb height anomalies from the GEOS–DAS and the ECMWF analysis, filtered for the periods of (a) 2–10 days, (b) 10–20 days, and (c) 20–70 days. Contour intervals are 0.1. The shading represents the correlations greater than 0.9 95

71 Correlation between the 500 mb height anomalies, filtered for the periods of 2–10 days, 10–20 days, from the GEOS–DAS and the ECMWF analysis and 20–70 days at (a) 40N, (b) the equator, and (c) 40S. 96

72 Departure from the seasonal cycle for the 200 mb zonal mean u-wind from (a) the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 2.0 m/sec. Zero contours are not drawn. 98

73 Departure from the seasonal cycle for the 200 mb zonal mean v-wind from (a) the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 0.2 m/sec. Zero contours are not drawn. The values greater than 0.4 m/sec and less than -0.4 m/sec are shaded. 99

74 Departure from the seasonal cycle for the 300 mb zonal mean height from (a) the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 20.0 m. Zero contours are not drawn. The values greater than 20 m and less than -20 m are shaded. 100

75 Longitude-time section of the outgoing longwave radiation anomaly [5S-5N] from NOAA satellite measurements, the GEOS-DAS and the GEOS-GCM simulation. Units: W/m^2 101

76 The 200 mb streamfunction difference (1988-87) for (a) May, (b) June, and (c) July from the GEOS–DAS. Contour intervals are 5.0×10^6 m^2/sec. The values greater than 10.0×10^6 m^2/sec and less than -10.0×10^6 m^2/sec are shaded. 104

77 The 200 mb streamfunction difference (1988-87) for (a) May, (b) June, and (c) July from the ECMWF analysis. Contour intervals are 5.0×10^6 m^2/sec. The values greater than 10.0×10^6 m^2/sec and less than -10.0×10^6 m^2/sec are shaded. 105

78 The 200 mb velocity potential difference (1988-87) for (a) May, (b) June, and (c) July from the GEOS–DAS. Contour intervals are 1.0×10^6 m^2/sec. The values greater than 5.0×10^6 m^2/sec and less than -5.0×10^6 m^2/sec are shaded. 106

79 The 200 mb velocity potential difference (1988-87) for (a) May, (b) June, and (c) July from the ECMWF analysis. Contour intervals are 1.0×10^6 m^2/sec. The values greater than 5.0×10^6 m^2/sec and less than -5.0×10^6 m^2/sec are shaded. 107

80 The total precipitable water difference (88-87) between July 1987 and July 1988 from (a) the GEOS–DAS and (b) SSM/I. Contour intervals are 0.5 gm/cm^2. The zero contours are not drawn. The values greater than 0.5 gm/cm^2 and less than -0.5 gm/cm^2 are shaded. 108

81 The outgoing longwave radiation difference (1988-87) for May, June, July, and August from the GEOS-DAS. Units: W/m^2 109

82 The outgoing longwave radiation difference (1988-87) for May, June, July, and August from the NOAA satellite observations. Units: W/m^2 110

83 The outgoing longwave radiation difference (1988-87) for May, June, July, and August from the GEOS-GCM simulation. Units: W/m^2 111

84 The regional precipitation anomalies over land from the GEOS analysis and the station observations. The anomalies are departure from the seasonal cycle. Units are *mm/day*. 113

85 The regional precipitation anomalies over land from the GEOS analysis and the station observations. The anomalies are departure from the seasonal cycle. Units are *mm/day*. 114

86 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1985. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 115

87 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1985, and February 1986. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 116

88 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1986. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 117

89 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1986, and February 1987. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 118

90 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1987. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 119

91 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1987, and February 1988. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 120

92 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1988. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 121

93 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1988, and February 1989. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 122

94 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1989. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 123

95 The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1989, and February 1990. Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded. 124

96 A summary of known deficiencies in the GEOS–1 DAS output based on the analysis of the first 5 years (March 1985–February 1990). 126

97 Locations of rawinsonde (upper panel) and special upper air (lower panel) observations for May 5th, 00Z. The latter include pilot balloon, dropwinsonde and reconnaissance aircraft and radar winds. 128

98 Locations of land (upper panel) and ship (lower panel) reports for May 5th, 00Z. 129

99 Locations of buoy (upper panel) and aircraft (lower panel) reports for May 5th, 00Z. 130

100 Locations of satellite wind (upper panel) and NESDIS temperature retrievals (lower panel) under clear (path A), partly cloudy (path B) and cloudy (path C) conditions for May 5th, 00Z. 131

101 Number of TOVS clear (path A, upper panel) and TOVS partly cloudy (path B, lower panel) reports for May 1985. Stratospheric only soundings are defined as occurring above the 100 mb level. 133

102 Number of TOVS cloudy (path C, upper panel) and surface reports (lower panel) for May 1985. In (a) stratospheric only soundings are defined as occurring above the 100 mb level. 134

103 Number of rawinsonde (moisture only, upper panel) and rawinsonde (all, lower panel) reports for May 1985. 135

104 Number of upper air reports (upper panel), and satellite wind reports (lower panel) for May 1985. 136

1 Introduction

The Data Assimilation Office (DAO) at Goddard Space Flight Center has produced a multi–year global assimilated data set with version 1 of the Goddard Earth Observing System Data Assimilation System (GEOS–1 DAS; Schubert et al., 1993). One of the main goals of this project was to produce a research quality data set suitable for the study of general Earth science problems such as climate variability, atmospheric chemistry, stratosphere–troposphere exchange, and surface processes.

The DAO project is one of several efforts (see also Bengtsson and Shukla, 1988; Kalnay and Jenne, 1991; US Naval Research Laboratory) to re–analyze historical data with an unvarying analysis system. These "reanalyses" are motivated by the fact that operational data assimilation systems undergo frequent updates which introduce spurious climate signals in the analysis output. Figure 1 illustrates the dominant climate signals found in operational analyses at various frequencies. At weekly and shorter time scales the variability is dominated by synoptic–scale (weather) fluctuations. At somewhat longer time scales one finds the signature of middle latitude planetary wave fluctuations and blocking events. In the tropics, the Madden–Julian Oscillation is the dominant mode of intraseasonal variation; operational analyses appear to capture at least some aspects of this oscillation fairly well (see e.g., section 4.3.2). The seasonal time scale is dominanted in most regions by the annual cycle. Various teleconnections such as the Pacific/North American (PNA) pattern are also important components of the variability at this and longer time scales. It is at the seasonal time scale that spurious signals due to changes in the analysis system become quite evident. At interannual and longer time scales the problem gets worse, such that the signature of even the dominant climate events (e.g., the El Niño/Southern Oscillation, ENSO) is often masked by the spurious signals produced by analysis system changes. The bottom panel of Fig. 1 provides a somewhat subjective assessment of the quality of the analyses in each of the frequency bands. The strong analysis signal at the longest time scales makes them poorly suited for studying climate change. As we shall see in section 4.4, the signal of the analysis system changes is very strong in some quantities, while it is almost undetectable in others. Unfortunately, quantities tied to the hydrological cycle (e.g., convection, divergence, precipitation) are especially sensitive to analysis system changes. Section 4.4 provides examples of the ablility of the GEOS-DAS to estimate interannual variations in such key phenomena as the Hadley cell and tropical precipitation.

The current assimilation experiment is also considered a critical benchmark for further development of the GEOS system, which must become operational in time for the first launch of the Earth Observing System (EOS) platform scheduled for 1998. Thus it is important that a thorough evaluation of the quality and usefulness of the data be done in a timely manner. While an important part of the evaluation is carried out "in–house", the potential applications of the data are quite broad, and feedback from the general Earth Science community is deemed vital. This feedback is critical to the DAO development effort since, unlike data assimilation for Numerical Weather Prediction (NWP), which has as its

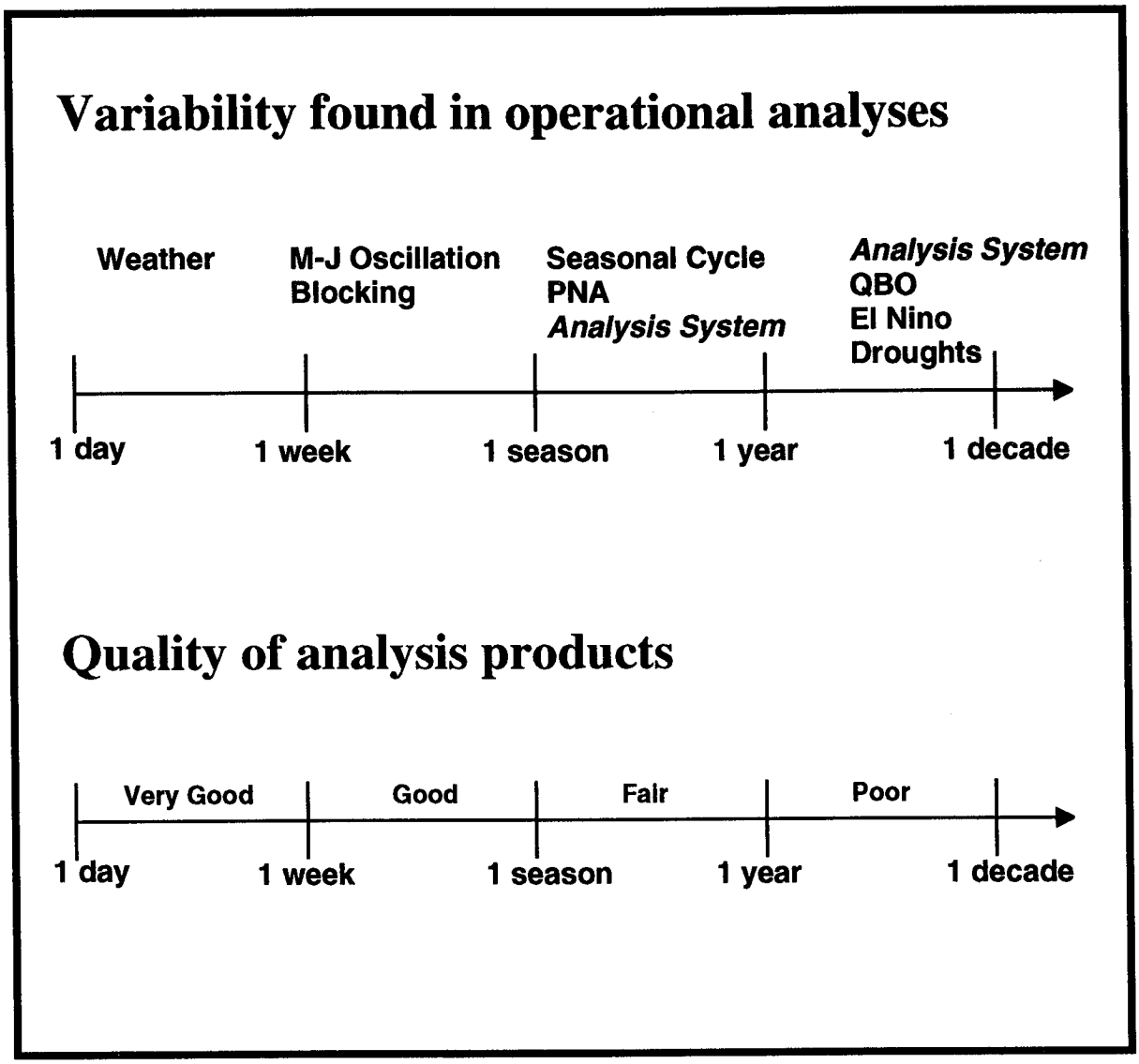


Figure 1: A schematic of the variability found in operational analyses. The spurious climate signals produced by changes in the data assimilation systems are found primarily at the seasonal and longer time scales.

central goal to produce initial conditions that maximize forecast skill, there is no clear cut measure of quality for climate data assimilation.

It is expected that at the turn of the century the GEOS–DAS will provide one of the primary, high–level data products from the Earth Observing System (EOS) program for the study of the Earth (atmosphere, ocean, and land) system. The quality and usefulness of these data products will depend very much on the veracity of the modeled physical processes and the extent to which the assimilation system encompasses the climatologically important components of the climate system. It should be kept in mind that the current version of the GEOS–DAS has a strong NWP heritage (and thus atmospheric focus); the scientific studies employing this first multi–year assimilated data set will help set priorities for increased emphasis in other areas.

The greatest potential benefit of assimilation systems for climate studies is that they can provide essentially time continuous global estimates of all the relevant parameters at the full resolution of the assimilating geophysical model. The danger is, of course, that users treat the assimilated data as truth, while in fact the estimates may not be very good. This is particularly true for highly derived quantities, which are only weakly constrained by the observations. Thus climate applications place new demands on the quality of the parameterized physical processes in the assimilating geophysical models. For example, accurate and consistent estimates of such quantities as precipitation, cloudiness, and surface fluxes require a degree of veracity in the physical parameterizations and a level of sophisticated in the analysis techniques that current systems are just beginning to achieve.

The purpose of this document is to provide potential users of the data a convenient guide to NASA's first multi–year assimilated data set and an early look at the quality of the data for studying the climate and its variability. This document is also intended to alert users to potential limitations of the GEOS–1 data, in particular, and of assimilated data, in general.

Section 2 discusses the possible impact of model deficiencies on the assimilated data. Section 3 gives an overview of the GEOS–1 DAS. Selected results from the assimilation are presented in Section 4. The results are organized as follows: Section 4.2 shows various climate mean fields, including differences from the operational European Center for Medium–Range Weather Forecasts (ECMWF) analyses and from various observational data sets described in Section 4.1. The surface energy balance is presented in section 4.2.3. Global and regional precipitation estimates, including the full vertically–integrated U.S. moisture budget, are presented in section 4.2.4. Sections 4.3–4.4 give examples of intraseasonal and interannual variability, respectively. Section 4.5 gives some general conclusions, and includes a list of known problems with the assimilated data. Section 5 provides information about the input observational data in the form of data coverage maps, data count histograms, and summaries of the major occurrences of unexpected/unusual data volumes. Section 6 provides information about the organization of the various data sets from the multi–year assimilation available from the NASA Center for Computational Science (NCCS), from the

Goddard Distributed Active Archive Center (GDAAC), and from a local (DAO) server via anonymous ftp.

2 Assimilated Data: The Role of the Model

The key advantage of modern global data assimilation systems is the incorporation of a geophysical model as an integral part of the data processing scheme. Currently these are primarily atmospheric general circulation models (AGCMs). It is the model which, in principle, provides a global and physically consistent framework for combining and interpreting the observations. An important goal of climate data assimilation is to increase the utility of assimilated data products by making available accurate estimates, of not just the prognostic fields, but of a host of derived fields based on the model's physical parameterizations. These estimates included climate processes and forcing functions (e.g., precipitation and heating rates), which should allow a better understanding of the mechanisms driving climate variability.

Data assimilation is basically a synthesis of two different estimates of nature: one provided by the observations and the other provided by a short forecast from a geophysical model. The accuracy and distribution of the observations, the veracity of the geophysical model, and the ability of the analysis technique to optimally combine the observations and the model forecast (the first guess) all influence the accuracy of the final assimilated fields.

Molod et al. (1995) have performed an extensive study of the impact of the model on the quality of the GEOS-1 assimilation, and the reader is referred to that study for specific examples. We present here a general overview of what users should be aware of when working with assimilated data.

Assimilated data products can be approximately grouped into two catagories: those (primarily prognostic) fields that are directly assimilated (e.g. winds and specific humidity) and those (primarily diagnostic) fields that are generated from the various physical parameterizations. The former are the quantities which are strongly constrained by the observations and, where these are available, are only marginally impacted by errors in the model (examples are shown in the section 4). The quality of the latter fields depends strongly both on the accuracy of the physical parameterizations and the quality of the observations. Of course, in regions where observations are sparse, all estimates will be dominated by the model's first guess field.

While the hope is that by constraining the prognostic fields with the observations the geophysical model will generate consistent and accurate estimates of the various diagnostic fields, there are two basic reasons why this has generally not been true. One reason is that the process of inserting observations into the model tends to generate imbalances in

the model; for example, convective precipitation tends to show very unrealistic fluctuations during the times at which data are inserted. It is for this reason that the operational centers currently provide precipitation estimates based on short term forecasts. Fortunately, much progress has been made over the past few years to dramatically reduce the severity of this problem. In the GEOS–1 DAS this problem has been virtually eliminated by the introduction of an Incremental Analysis Update (IAU) procedure (Bloom et al. 1995; see section 3), and the precipitation fields are those generated during the assimilation.

The other main reason for inconsistencies between the prognostic and diagnostic fields is that the physical parameterizations have deficiencies which result in significant model bias. Thus, for example, if the model prefers to have a drier tropical atmosphere than nature, the observations will systematically add water during the assimilation. This will be reflected in budget equations for the assimilation, which then include a mean moisture source from the observations. The general budget equation for the assimilated quantities is thus:

$$\dot{x} = dynamics + physics + \Delta x, \tag{1}$$

where the dynamics terms depends primarily on directly assimilated quantities (e.g. winds, temperature, moisture), and the physics terms are determined from the physical parameterizations, which depend only indirectly on the data. The Δx is a non–physical source/sink coming from the observations. Note that not all terms in the dynamics are strongly constrained by the observations. An obvious exception is the vertical motion field, which in general, is poorly observed.

Many applications of assimilated data have focused on the dynamics terms. The physics fields and data increments have rarely been made available, so that the total "forcing" is determined as a residual of the dynamics and tendency terms. This is an obvious and useful application of existing analyses since most of the effort has been directed at improving the estimates of the prognostic (and thus dynamics) fields. As the physics fields improve, assimilated data will provide a more complete picture of the processes driving the climate system. Thus the residual approach will be replaced with the direct calculation (from the GCM parameterizations) of all the relevant forcing (physics) terms.

It is as yet unclear whether current assimilation systems are of sufficient quality that the residual approach to estimating the forcing terms can give way to the direct approach of using the physical parameterizations. Some light can be shed on this by examing the mean increments (Δx), which provide an estimate of the model bias. Averaging (1) over a time period τ one obtains

$$\overline{\dot{x}}^{\tau} = \overline{dynamics}^{\tau} + \overline{physics}^{\tau} + \overline{\Delta x}^{\tau}. \tag{2}$$

The size of the bias must be considered relative to the physical forcing terms. For example, when employing such a budget to study nature's balance, one would like to have the ratio

$$\frac{\overline{\Delta x}^{\tau}}{\overline{physics}^{\tau}} \tag{3}$$

be much less than 1. If this is not the case, the budget terms are ambiguous, since a substantial part of the forcing cannot be attributed to real physical processes. For large τ, (and removing any systematic tendency such as that associated with the seasonal cycle) (2) may be written as

$$\overline{physics}^{\tau} + \overline{\Delta x}^{\tau} = -\overline{dynamics}^{\tau}, \tag{4}$$

which illustrates the traditional approach to estimating the forcing as a residual of the dynamics. Section 4.2.4 shows examples of the contributions of the mean analysis increments to the budget of the vertically–integrated moisture over the continental United States. The results show the increments are a substantial component of the mean budget, suggesting that considerable improvements still need to be made to the physical parameterizations to reduce the model bias. The results of section 4.4 suggest, however, that despite the mean climate bias, estimates of interannual forcing anomalies may be quite reasonable. This is apparently the result of the bias cancelling when the budgets (4) for, say, two different years are subtracted.

3 Overview of GEOS–1 DAS

The two primary components of the GEOS–1 DAS are the GEOS–1 GCM and an optimal interpolation (OI) analysis scheme. The GEOS–1 GCM is described in Takacs et al. (1994), Suarez and Takacs (1995) and Molod et al. (1995). The analysis scheme is described in Pfaendtner et al. (1995). For convenience, the GEOS–1 DAS is summarized below.

3.1 The OI Scheme

The GEOS–1 DAS employs an optimal interpolation (OI) analysis scheme, which for the multi–year assimilation described here, was carried out at a horizontal resolution of 2° latitude by 2.5° longitude at 14 upper–air pressure levels (20, 30, 50, 70, 100, 150, 200, 250, 300, 400, 500, 700, 850, 1000 mb) and at sea–level. The analysis increments are computed every 6 hours using observations from a +/− 3 hour data window centered on the analysis

times (00, 06, 12, and 18 UTC). The innovation vector (observation minus background forecast) used as input to the OI is computed using a single forecast valid at the analysis time.

The upper–air analyses of height, wind, and moisture incorporate the data from rawinsondes, dropwindsondes, rocketsondes, aircraft winds, cloud tracked winds, and thicknesses from the historical TOVS soundings produced by NOAA NESDIS. The satellite heights are computed using a reference level that depends on the analyzed sea–level pressure. The only bogus data used are 1000 mb height observations, which are generated above pressure reports from ships. These serve to further couple the surface and upper–air analyses.

The OI scheme is multivariate in geopotential height and winds and employs a damped cosine function for the horizontal correlation of model prediction error. The height–wind cross–correlation model is geostrophic and scaled to zero at the equator. The multivariate surface analysis scheme over the oceans adopts an Ekman balance for the pressure–wind analysis. The moisture analysis for mixing ratio employs only rawinsonde moisture data. All grid point analyses are done using up to 75 nearby observations from within a circular data–selection cylinder of 1600 km radius.

The assimilation system does not include an initialization scheme and relies on the damping properties of a Matsuno time differencing scheme to control initial imbalances generated by the insertion of observations. However, the initial imbalances and spinup have been greatly reduced over earlier versions by the introduction of an Incremental Analysis Update (IAU) procedure (Bloom et al., 1991, 1995). In the IAU procedure, standard OI analysis increments are computed at the analysis times; the increments are then inserted gradually into the AGCM by rerunning the forecast and adding a fraction of the increment at each model time step over the 6 hour period centered on the analysis time (see Pfaendtner et al. 1995 and Bloom et al. 1995 for details of the implementation and the statistical properties of the IAU approach). The assimilation thus effectively consists of a continuous AGCM forecast with additional heat, momentum, moisture and mass source terms updated every 6 hours from observations (see section 2).

3.2 The GEOS–1 Model

The tropospheric version of the GEOS–1 GCM uses the potential enstrophy and energy–conserving horizontal differencing scheme on a C–grid developed by Sadourny (1975), and further described by Burridge and Haseler (1977). An explicit leapfrog scheme is used for the time differencing, applying an Asselin (1972) time filter to damp out the computational mode. An 8th–order Shapiro filter (Shapiro, 1970) is applied to the wind, potential temperature, and specific humidity to avoid non–linear computational instability. The filter is applied at every step in such a way that the amplitude of the two–grid interval wave would be reduced by half in two hours. Applying the filter weakly at each time step eliminates the

shock that occurred in earlier assimilations by intermittent application of filter. The model also uses a polar Fourier filter to avoid linear instability due to violation of the CFL condition for the Lamb wave and internal gravity waves. This polar filter, however, is applied only to the tendencies of the winds, potential temperature, specific humidity, and surface pressure. The model's vertical finite differencing scheme is that of Arakawa and Suarez (1983). The above dynamics routines are organized into a plug–compatible module called the ARIES/GEOS "dynamical core" (Suarez and Takacs 1995).

The infrared and solar radiation parameterizations follow closely those described by Harshvardhan et al. (1987). In the longwave, water vapor absorption is parameterized as in Chou (1984), the 15 micron band of CO_2 as in Chou et al. (1983), and ozone absorption as in Rodgers (1968) with the modifications suggested by Rosenfield et al. (1987). The shortwave follows Davies (1982), as described in Harshvardhan et al. (1987). Shortwave absorption by water vapor uses a k–distribution approach as in Lacis and Hansen (1974). Cloud albedo and transmissivity for the model layers are obtained from specified single–scattering albedo and cloud optical thickness by using the delta–Eddington approximation (Joseph et al., 1976; King and Harshvardhan, 1986).

The penetrative convection originating in the boundary layer is parameterized using the Relaxed Arakawa–Schubert (RAS) scheme (Moorthi and Suarez, 1992), which is a simple and efficient implementation of the Arakawa–Schubert (1974) scheme. Unlike the Arakawa–Schubert scheme, which solves an adjustment problem by considering simultaneous interaction among all possible cloud types, RAS considers only one cloud at a time, and rather than adjusting fully every hour or two, it does a series of partial adjustments that tend to relax the state toward equilibrium. The AGCM also includes a parameterization that models the evaporation of falling convective rain as described in Sud and Molod (1988). Negative values of specific humidity produced by the finite–differenced advection are filled by borrowing from below.

The planetary boundary layer (PBL) is explicitly resolved in a 2 to 4 layer region. Wind, temperature, and humidity profiles in an "extended" surface layer (which can be up to 150m thick), and the turbulent fluxes of heat, moisture, and momentum at the surface are obtained from Monin–Obukov similarity theory by selecting similarity functions that approach the convective limit for unstable profiles and that agree with observations for very stable profiles. Surface roughness lengths are taken as functions of vegetation type over land and as a function of surface stress over water. Turbulent fluxes above the "extended" surfaced layer are computed using the second order closure model of Helfand and Labraga (1988). In this scheme, the turbulent kinetic energy is a prognostic variable and the remaining second order moments are diagnosed from it and from the atmospheric sounding.

For the multi–year assimilation, GEOS–1 was integrated on a 2° latitude by 2.5° longitude grid with 20 sigma levels. The sigma levels are distributed to provide enhanced resolution in the planetary boundary layer and at upper levels. The topography used in GEOS–1

was prepared from the 10 minute topography obtained from the Navy Fleet Numerical Oceanography Center in Monterey. The 2° latitude by 2.5° elevation values were obtained by averaging the high resolution values (areas with more than 60% water were considered water points), and then applying a Lanczos (1966) filter. The Lanczos filter was designed to remove small scale structure (it completely removes $2\Delta X$ waves) while minimizing the Gibbs phenomena.

This version of the AGCM is run without a land surface model (note that current development efforts include the incorporation of a land surface model). For the assimilation described here, soil moisture is computed off–line, based on a simple bucket model that used monthly mean observed surface air temperature and precipitation (Schemm et al., 1992). The snow line and surface albedo are prescribed and vary with the season (see Takacs et al. 1994). The sea–surface temperature is updated according to the observed monthly mean values provided by the Climate Analysis Center at NMC and the Center for Ocean, Land and Atmosphere (COLA).

3.3 The Processing System

Figure 2 presents an overview of the main processing elements of the GEOS–1 DAS. The primary computations are carried out at the Nasa Center for Computational Sciences (NCCS). All output is stored on the NCCS Unitree mass storage system which currently has a capacity of about 19.2 terabytes (TB) of near–online storage consisting of robotically–controlled silos.

The processing was begun on a Cray YMP with 6 processors and then moved to the Cray C90 (also with 6 processors) when it became available in August of 1993. The new computer provided a much improved turn around and allowed the completion of the first five years in just over one year (March 1993–May 1994). A substantial fraction of that time was devoted to trouble shooting, dealing with mass storage problems, and filling gaps in the observational database.

The processing is now fully automated. It is carried out in 5–10 day segments. After each segment is completed the output is temporarily placed on a silo attached directly to the Cray (current capacity of 2.4TB). After one month is completed, various post processing activities are carried out. These include the generation of monthly means, quality control statistics to look for outliers, and a re–organization of the data into a time series format (described in section 6). The time series are then copied to 8mm tape and given to the Goddard DAAC for ingest into their system.

Local workstations are used for pre–processing the conventional and satellite observations. This processing includes: reformatting, data thinning, consistency checks, and documenting the observational data coverage. Examples of the data coverage are shown in section 5.

A quick–look capability was developed to help keep track of the processing and to look for any obvious problems that might arise during production. The quick–look fields consist of a small subset of the output, which is automatically sent to the Cray silo to allow easy access and viewing on local workstations.

All visualization and much of the post–processing is carried out on local workstations. The GrADS display program (developed at the Center for Ocean–Land–Atmosphere, COLA), is used extensively for much of the in–house visualization and scientific analysis; the time series output has been organized to facilitate the use of this software (see section 6.2).

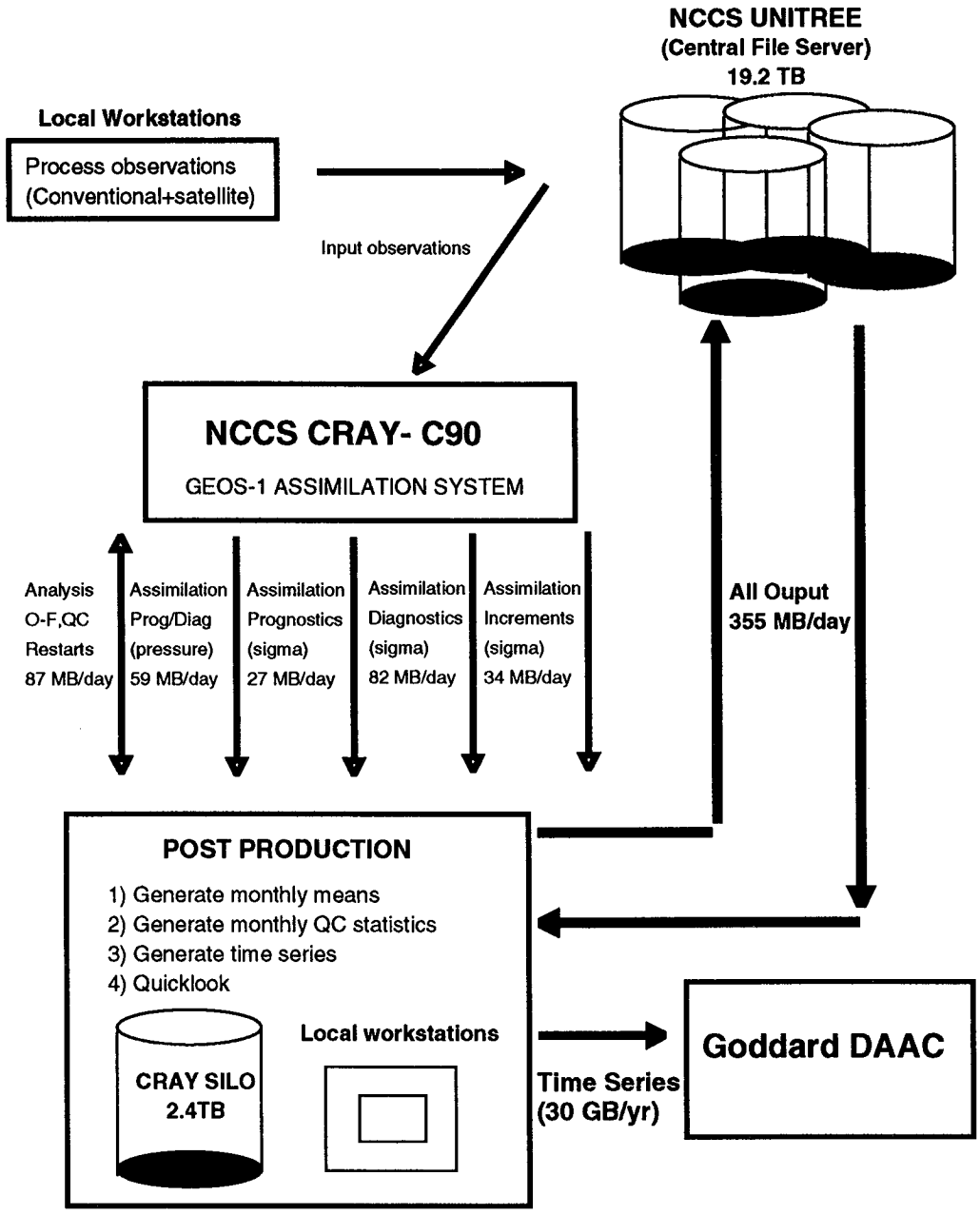


Figure 2: The main elements of the processing system and data flow of the GEOS–1 DAS.

4 Results

We present here some selected results from the first five years of the assimilation (March, 1985–February, 1990). Comparisons are made with ECMWF analyses, a parallel GEOS–1 GCM simulation forced by the same boundary conditions used in the assimilation, and other observational data sets described below. It is important to note that the ECMWF data are an operational product and therefore are impacted by changes introduced to the system during this time period (see e.g. Trenberth and Olson 1988; Arpe 1990). We hope in future reports to make further comparisons with the reanalysis products from both NMC and ECMWF. The assimilation currently extends to the end of 1993; it is planned that, by the end of 1995, the assimilation will cover the period (1979–1994).

Note that an output field generated from an assimilation is tradionally referred to as an analysis. This terminology is all right for those fields that are actually analyzed (e.g., with an optimal interpolation scheme), however, it is misleading for a diagnostic field (e.g., surface sensible heat flux), which is generated by the GCM during the assimilation but is never analyzed and, therefore, is not directly influenced by the observations. In this document we have attempted to be consistent and refer to the results as the "output from the assimilation" or simply as coming from the GEOS-DAS, though when comparing the prognostic fields with ECMWF results, we have also referred to the results as analyses.

The results presented here are not meant to be comprehensive, and the reader is referred to Takacs and Suarez (1995) and Molod et al. (1995) for further results on the circulation statistics and diagnostic fields. Extensive comparisons between the surface wind products from the GEOS–1 DAS and several other sources have been carried out over the North Pacific by Rienecker et al. (1995).

4.1 Comparison Datasets

ECMWF Analyses

The European Center for Medium–Range Weather Forecasts (ECMWF) analyses used for comparison consist of the operational "ECMWF/TOGA basic level III" data (ECMWF, 1989). The data were received in a packed format from ECMWF and consist of 00Z and 12Z uninitialized analyses on a 2.5° X 2.5° latitude/longitude grid beginning January 1, 1985. The data were unpacked and interpolated to a 2.0° X 2.5° latitude/longitude grid for comparison with the GEOS results. Periods with missing analyses (Schubert et. al., 1991) were avoided in the processing.

GEOS–1 GCM Simulation

A companion simulation was performed with the GEOS–1 model for the same five year period as the assimilation. The model simulation was forced by the same SST and soil moisture boundary conditions as the assimilation. The model is identical to that used in the assimilation, except that in simulation mode the model employs a leapfrog time–differencing scheme, while in assimilation mode it employs a Matsuno scheme (see Takacs et al. 1994).

NOAA OLR

The outgoing longwave radiation (OLR) data (kindly provide by K.–M. Lau) are from NOAA polar orbiting satelites (Gruber and Krueger, 1984). The data were received on a 2.5° X 2.5° latitude/longitude grid. The nighttime and daytime OLR values were averaged to obtain 5–day means. No data are available for March 1978–December 1978, most of November 1988, and March–April, 1989. The impact of satellite changes and data processing techniques on the quality of the data are discussed in Hurrel and Campbell (1992) and references therein.

SSM/I Total Precipitable Water

The Special Sensor Microwave Imager (SSM/I) total precipitable water is that generated by Wentz (1992). The radiative transfer algorithm uses three channels of microwave measurements (22V, 37V, 37H) and a model that accounts for absorption and emission in the atmosphere. The model uses a surface emissivity value over oceans appropriate for a wind–roughened sea surface. The model does not account for scattering by raindrops or by frozen hydrometers and is, therefore inaccurate for high rain rates. No calculation is done over land or sea ice, because of the complexity of the surface emissivity. These data are available starting in June 1987.

Comprehensive Ocean–Atmosphere Data Set

Estimates of the surface energy budget were obtained from the comprehensive Ocean–Atmosphere Data Set (COADS) processed by da Silva et al. (1995). The observations were corrected for systematic biases, and the derived quantities were interpolated to a $1° \times 1°$ grid, filling–in data gaps with a sinusoidal interpolation. Data quality control included the conversion of all wind measurements to an improved Beaufort Equivalent Scale and the rejection of nighttime cloud observations if the sky is too dark. The surface fluxes were calculated using a similarity theory parameterization that calculates the transfer coefficients

according to Large and Pond (1981). Surface radiation estimates were made using empirical formulas to approximate cloud attenuation, surface albedo, and atmospheric transmissivity. The objective analysis scheme used is essentially the same scheme described by Levitus (1982). This is a sucessive-correction scheme (Cressman 1959, Daley 1991) with a weight function developed by Barnes (1964).

Gridded Precipitation Station Data

The global monthly mean precipitation data are from the world monthly surface station climatology data received from W. Spangler and R. Jenne at the National Center for Atmospheric Research (NCAR). Most of the data are from the National Climate Data Center (NCDC). See Spangler and Jenne (1990) for additional information on data sources and characteristics. The station values were interpolated to a $2.0° \times 2.5°$ latitude/longitude grid by averaging station values within a 300 km radius of each grid point. The averaging weights are proportional to the inverse of the square of the distance to the station. The value at a gridpoint was set to undefined if there were no stations within the 300 km radius. Details of the processing may be found in Schemm et al. (1992)

The U.S. climate division precipitation data used here are a combination of about 300 National Weather Service (NWS) stations and approximately 2500 cooperative stations over the continental United States. These station data were averaged over each of the 344 United States climate divisions (see Cayan et. al., 1986). The data were then gridded as described above to a $2.0° \times 2.5°$ latitude/longitude grid. The data were kindly provided by Jae Schemm and Suranjana Saha of the National Meteorological Center.

4.2 Climate Mean Quantities

4.2.1 Zonal Mean Prognostic

The top panels of Figs. 3–5 show the seasonal cycles of the GEOS–DAS zonal mean zonal (u) and meridional (v) winds at 200 mb and the zonal mean 300 mb height field. The cycle is based on monthly data and is an average over the five years. The difference (GEOS−ECMWF) is shown in the bottom panels. The differences in the zonal (u) wind at 200 mb are generally less than 1 m/sec everywhere except over Antarctica. These differences are small compared to the magnitude of the u–wind itself. This is in contrast to the v–wind, which shows that the ECMWF boreal winter Hadley cell is almost twice as strong as that in the GEOS–DAS at this level. The differences in the height field are generally small except for much of the Antarctic, where the difference exceeds 100 meters.

Figures 6–13 show the five–year mean zonal mean cross sections of various quantities for

the two extreme seasons June–July–August (JJA) and December–January–February (DJF). The layout of the figures is such that the GEOS–DAS results are in the top panel and the differences with the ECMWF analyses (GEOS–ECMWF) are in the bottom panel.

Differences in the u–wind during DJF (Fig. 6) are largest in the tropics and Southern Hemisphere at upper levels. The GEOS–DAS tends to produce somewhat stronger westerlies (or weaker easterlies) at upper levels. Similar results are found for JJA (Fig. 7), except in the tropics above 100 mb where the GEOS analysis shows substantially stronger easterlies. The peculiar shape of the region of weak westerlies at the equator in the GEOS–DAS results during DJF appears to be the systematic signature of a problem with the GEOS-1 analysis scheme due to an improper treatment of the wind–wind correlations near the equator (see Appendix C.4 in Pfaendtner et al. 1995).

The differences in the DJF Hadley cell mentioned earlier are quite evident in Fig. 8. The ECMWF zonal mean meridional wind reaches a maximum of 3 m/sec at 200 mb, while the maximum in the GEOS–DAS is less than 2 m/sec and is placed somewhat higher, above 200 mb. In contrast, the JJA Hadley cells (Fig. 9) are more similar. The zonal mean temperature field (Figs. 10–11) show that the GEOS analyses are warmer than the ECMWF analyses in the tropical upper troposphere and cooler in the stratosphere during both seasons. The large differences near the surface are likely due to interpolation/extrapolation problems in regions of high topography. The zonal mean vertical velocity (Figs. 12–13) is remarkably similar in the two analyses. The largest differences occur near the South Pole. Also, the DJF Hadley cell appears to be stronger and better organized in the ECMWF analyses.

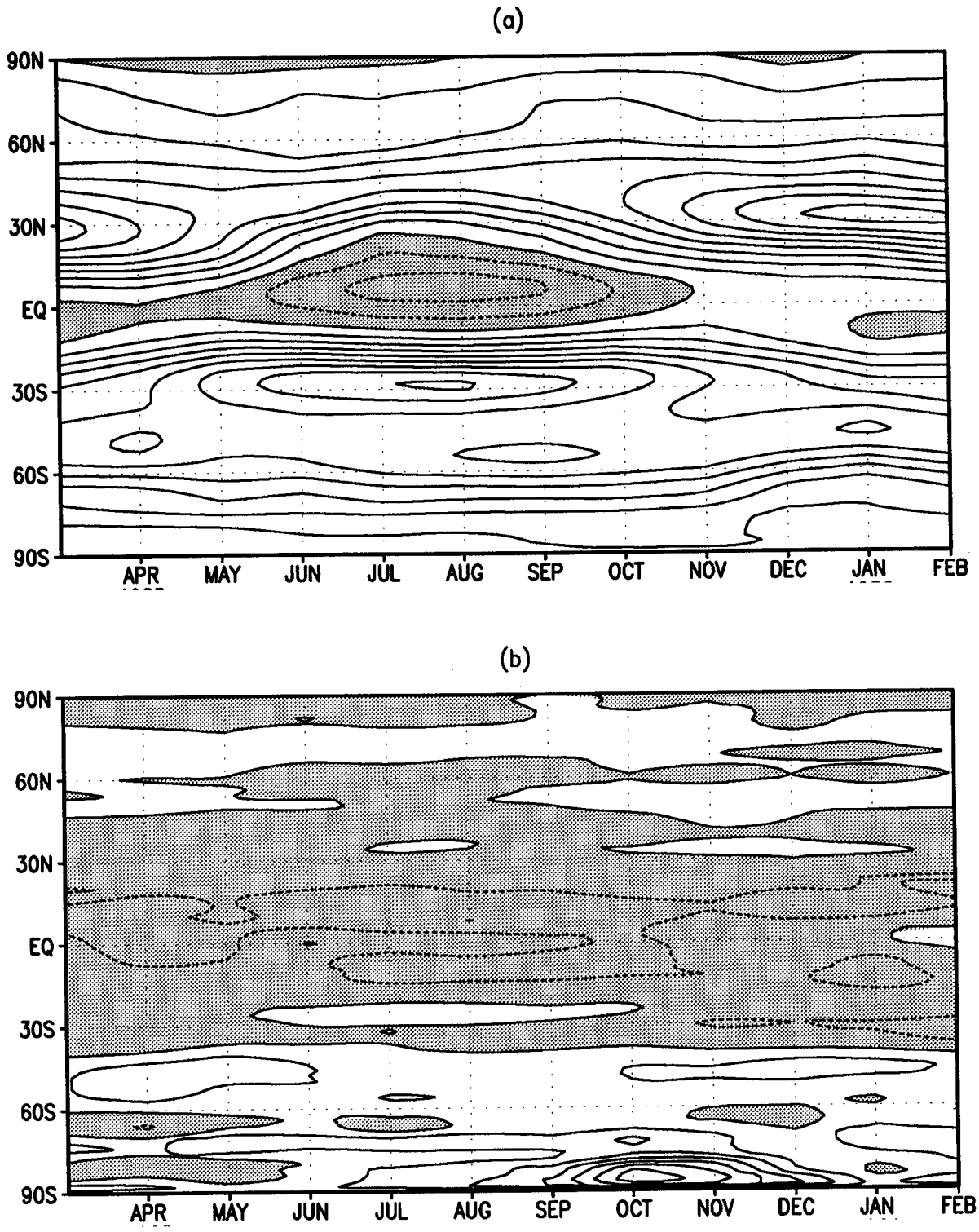


Figure 3: (a) Seasonal cycle of the 200 mb zonal mean u–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 5.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded.

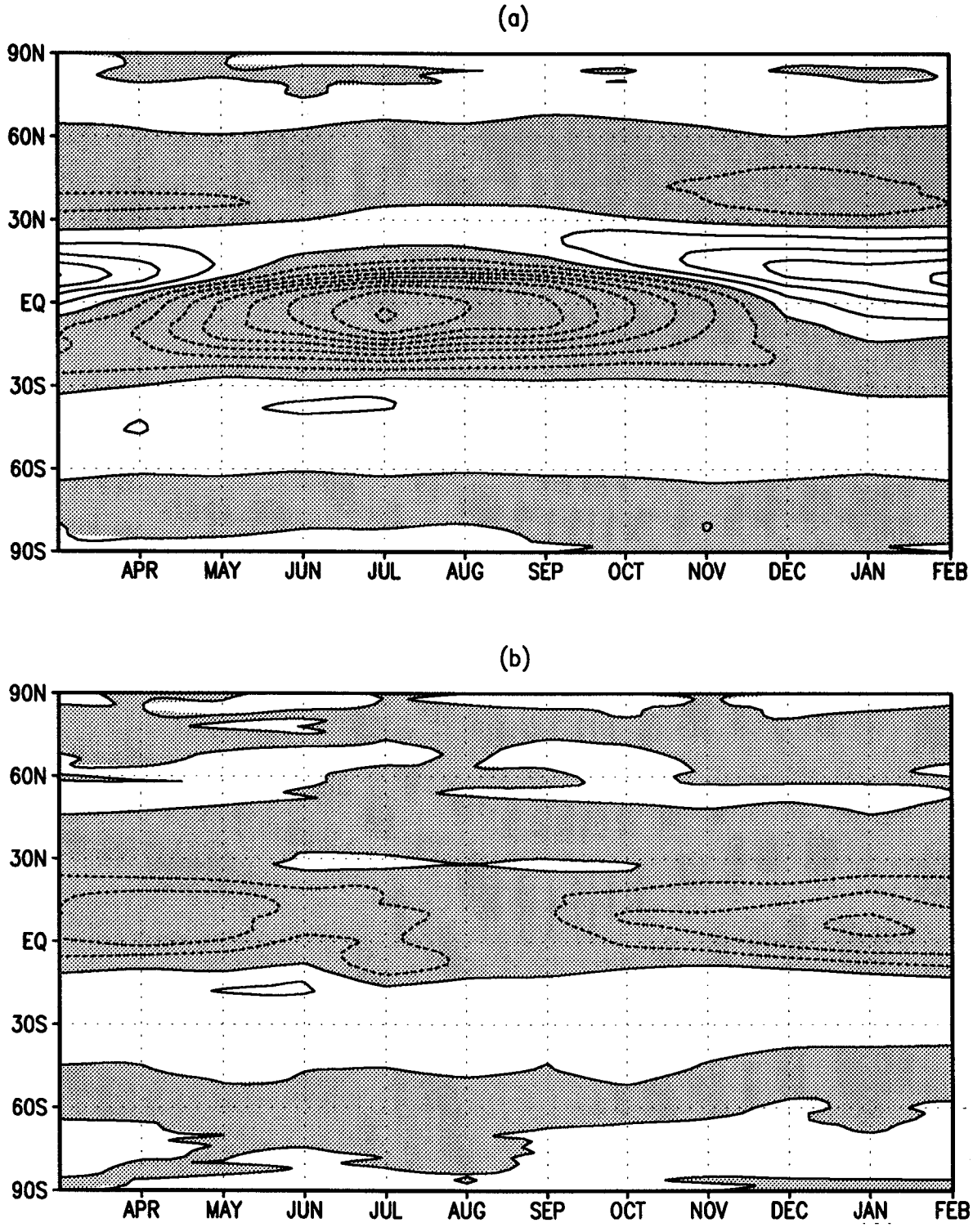


Figure 4: (a) Seasonal cycle of the 200 mb zonal mean v–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 0.5 m/sec. Negative regions are shaded.

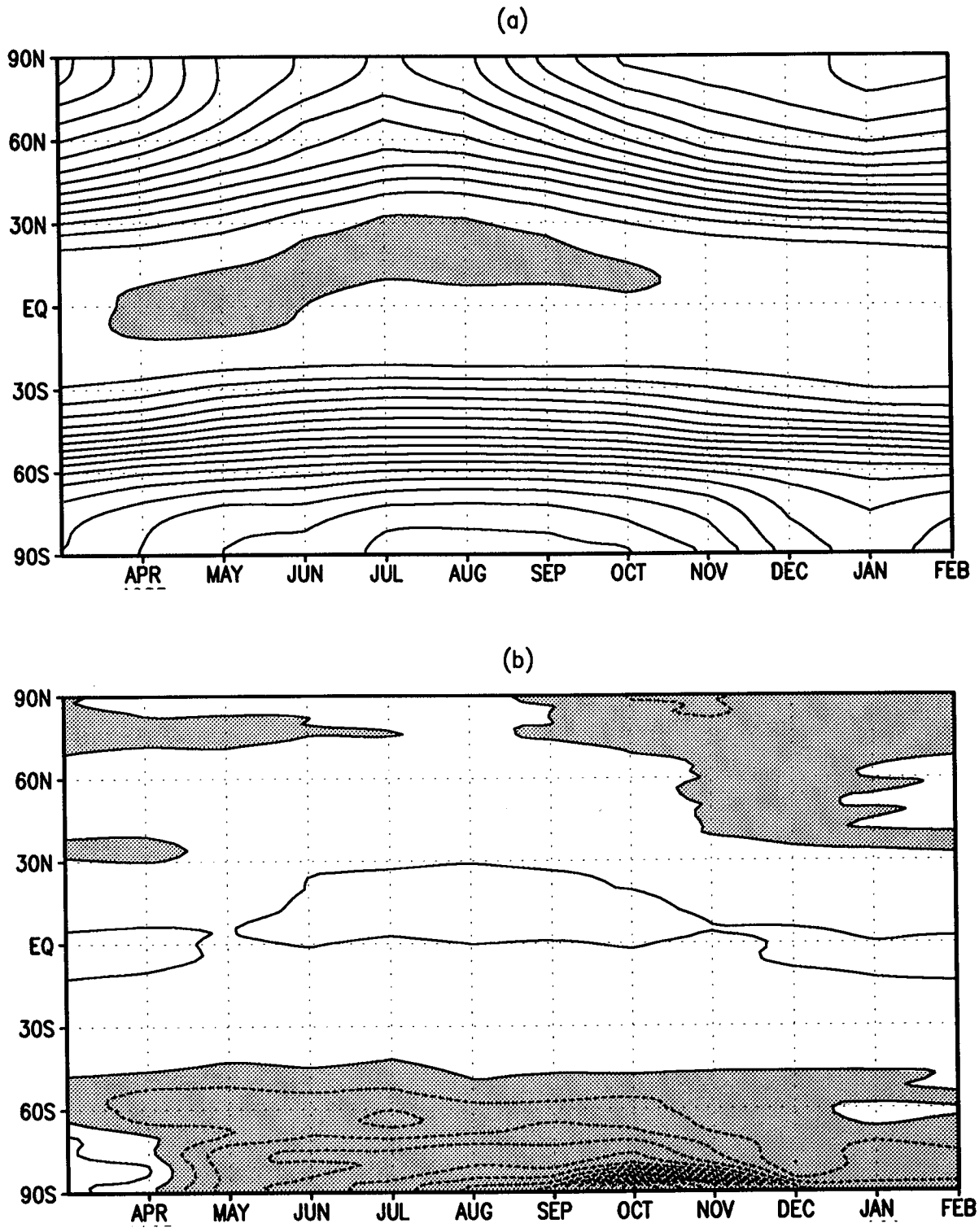


Figure 5: (a) Seasonal cycle of the 300 mb zonal mean height from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 100.0 m in (a) and 10.0 m in (b). The values greater than 9700.0 m are shaded in (a) and negative regions are shaded in (b).

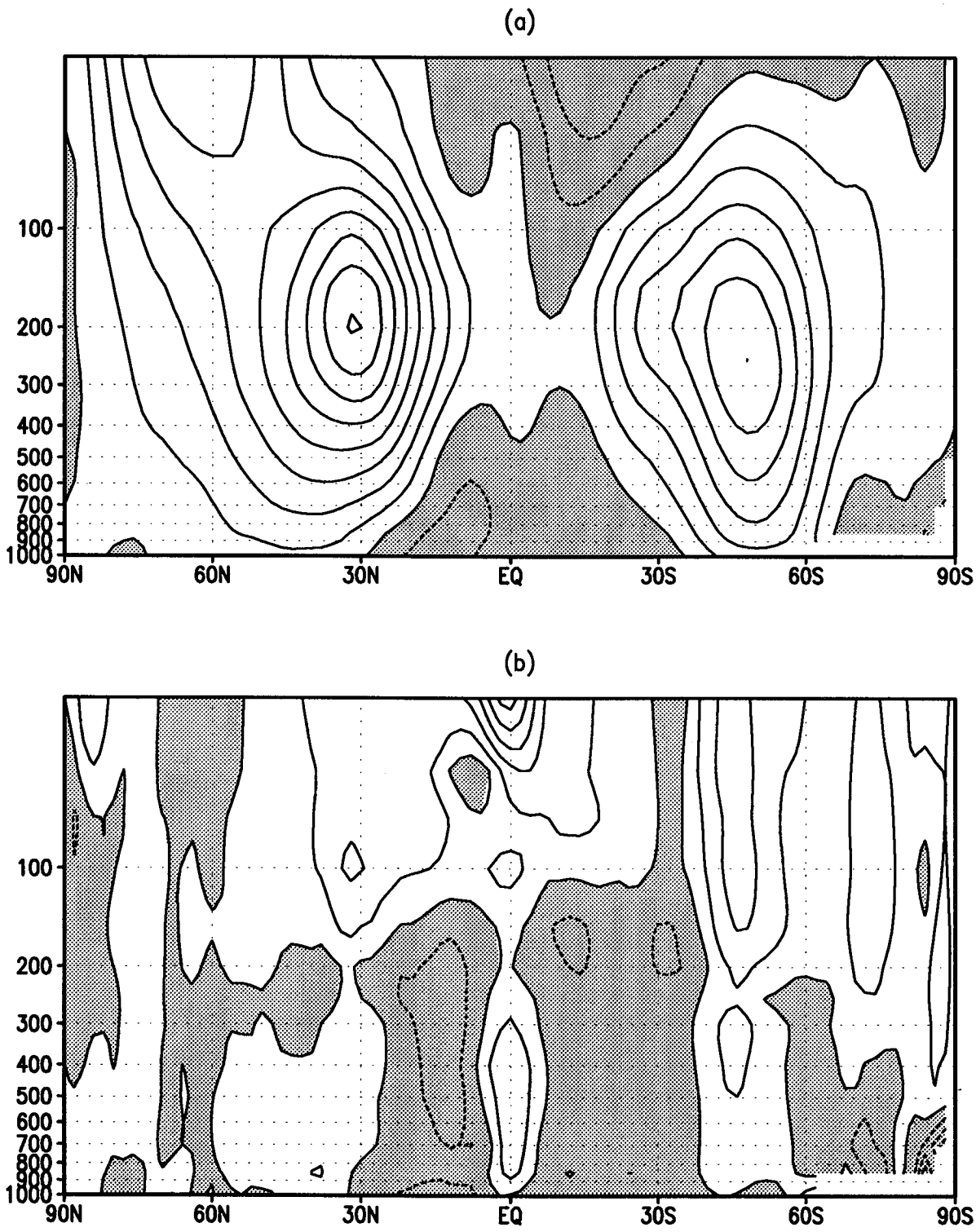


Figure 6: (a) DJF climatology of zonal mean u–wind from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. The top level is 30 mb

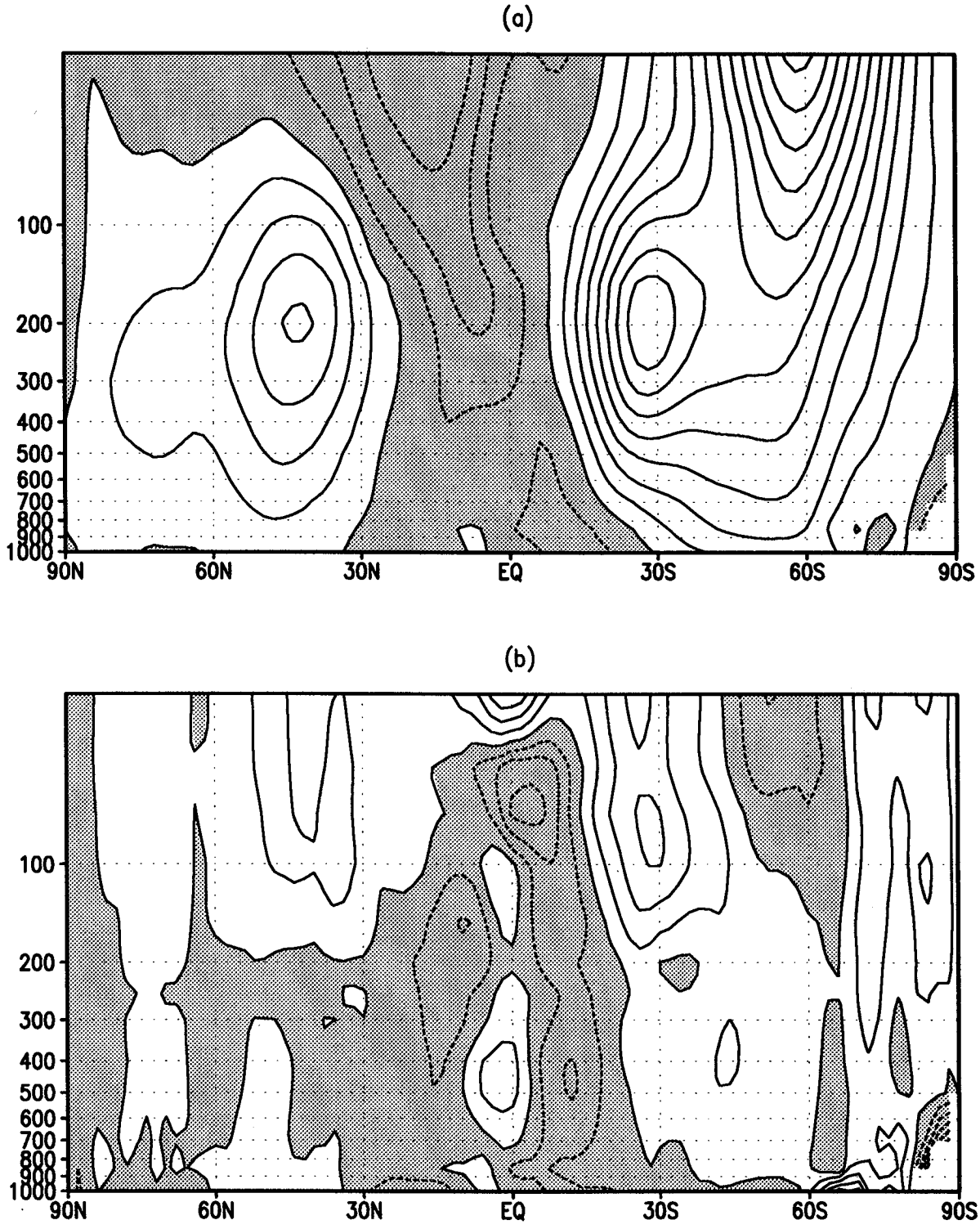


Figure 7: (a) JJA climatology of zonal mean u–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 5.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded. The top level is 30 mb

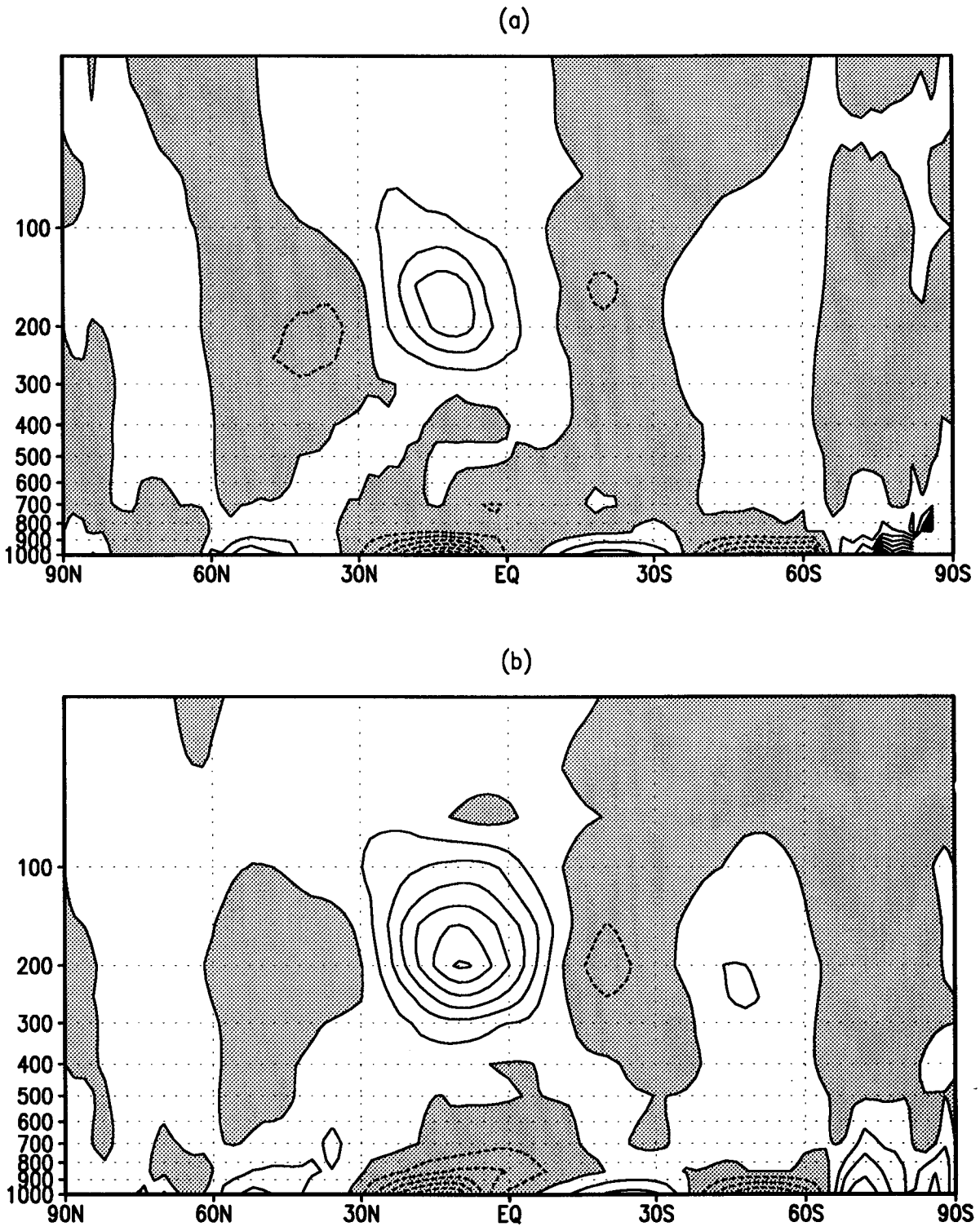


Figure 8: (a) DJF climatology of zonal mean v–wind from the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 0.5 m/sec. Negative regions are shaded. The top level is 30 mb

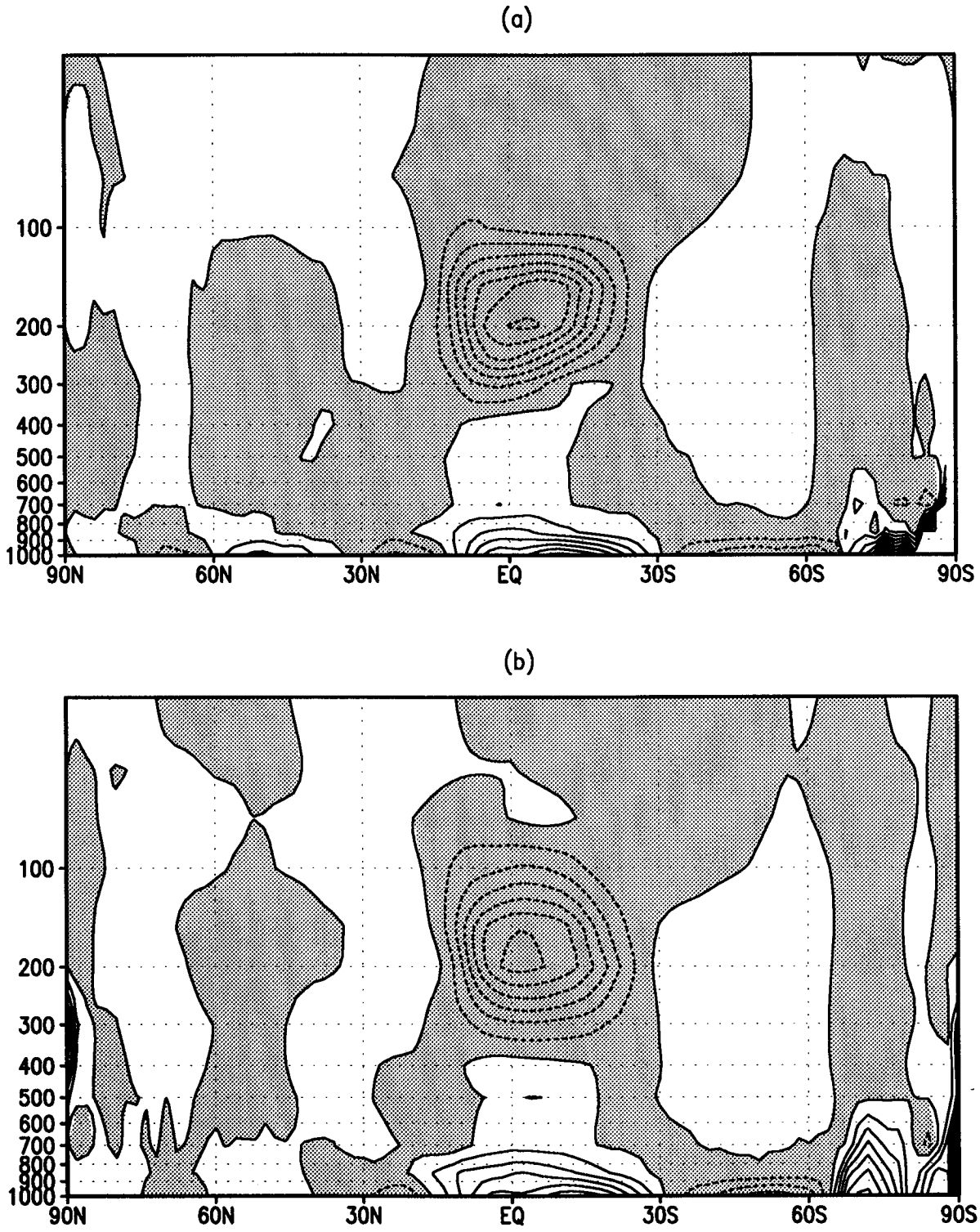


Figure 9: (a) JJA climatology of zonal mean v–wind from the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 0.5 m/sec. Negative regions are shaded. The top level is 30 mb

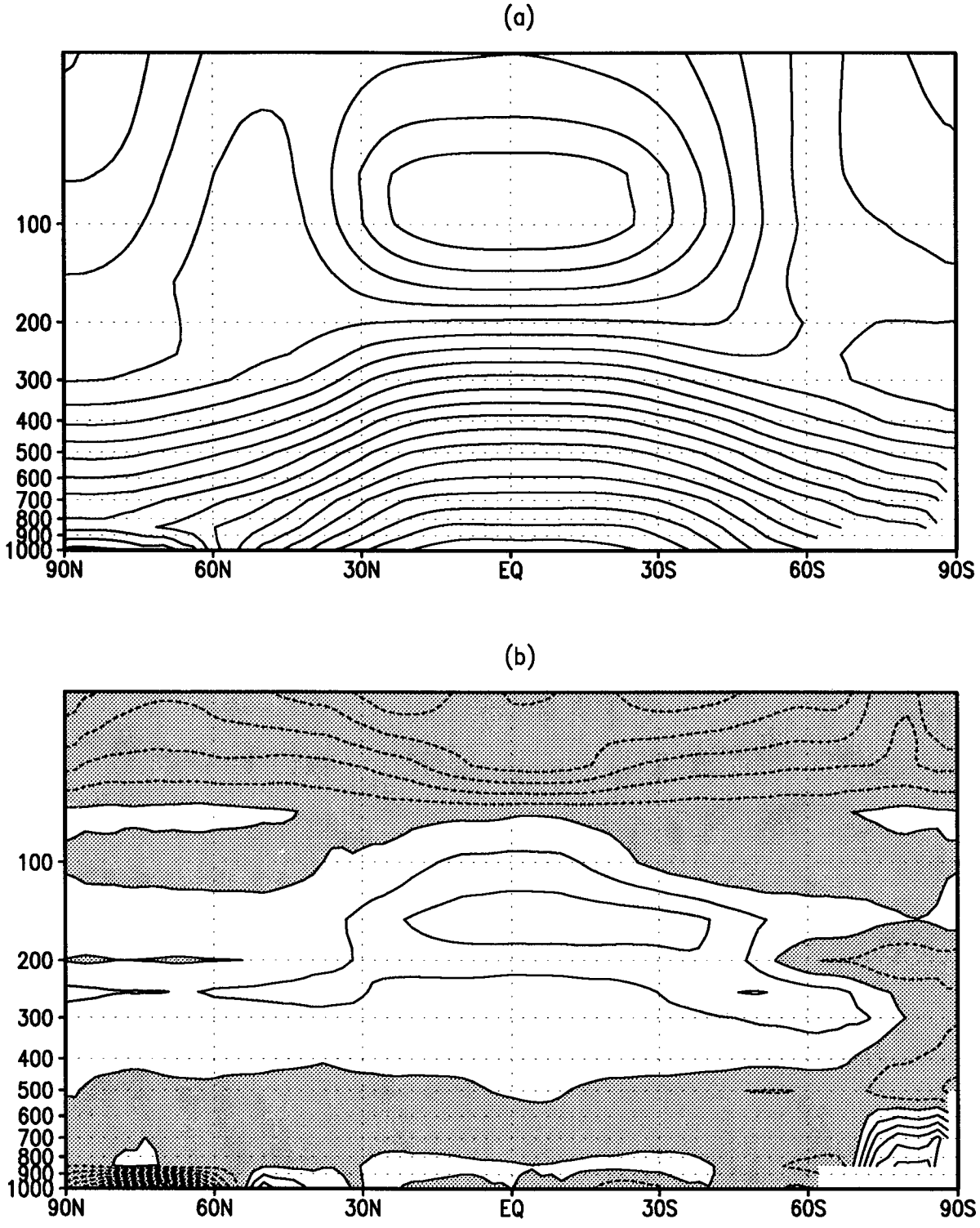


Figure 10: (a) DJF climatology of zonal mean temperature from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 5.0 $^{\circ}C$ in (a) and 1.0 $^{\circ}C$ in (b). Negative regions are shaded. The top level is 30 mb

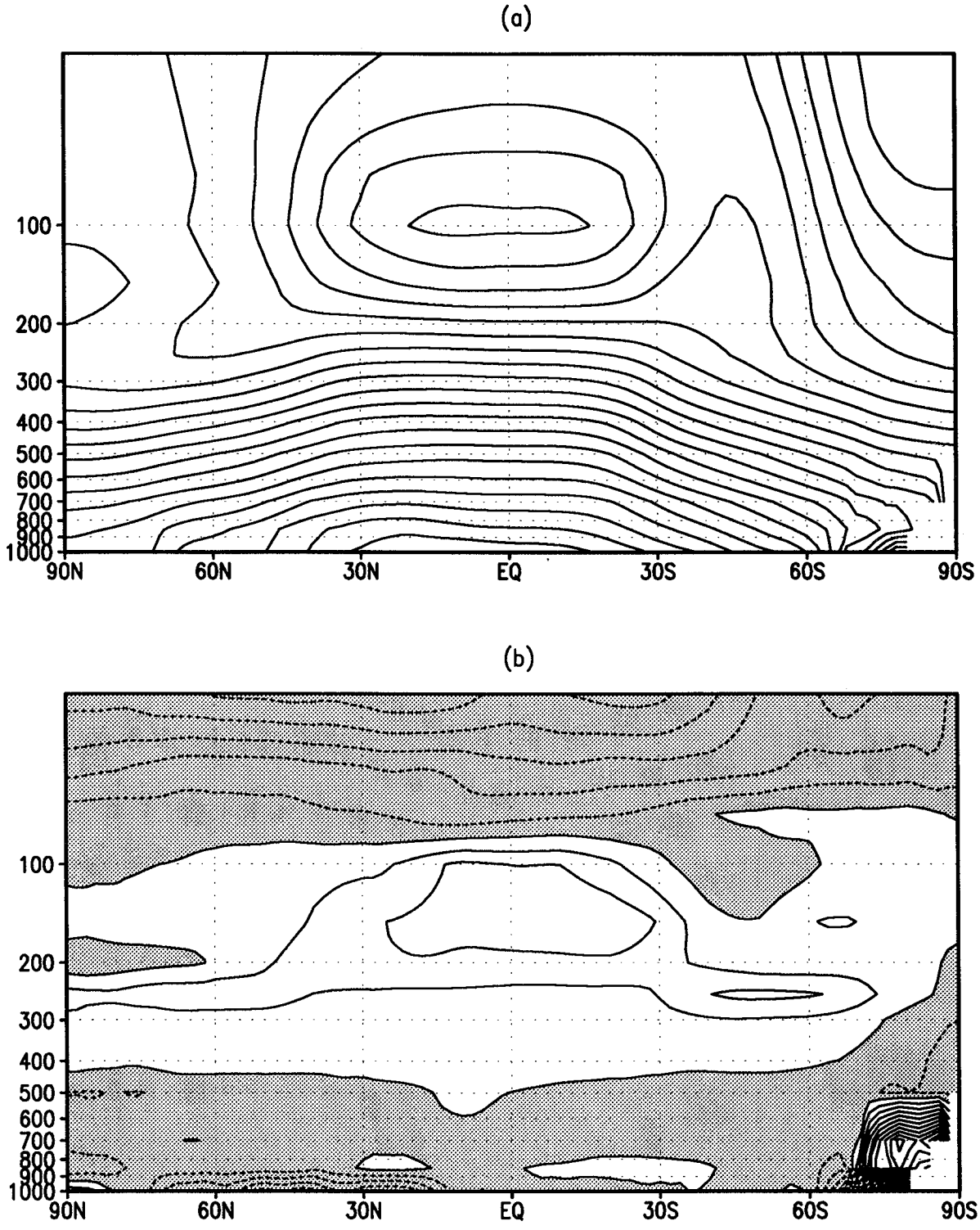


Figure 11: (a) JJA climatology of zonal mean temperature from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 5.0 $^{\circ}C$ in (a) and 1.0 $^{\circ}C$ in (b). Negative regions are shaded. The top level is 30 mb

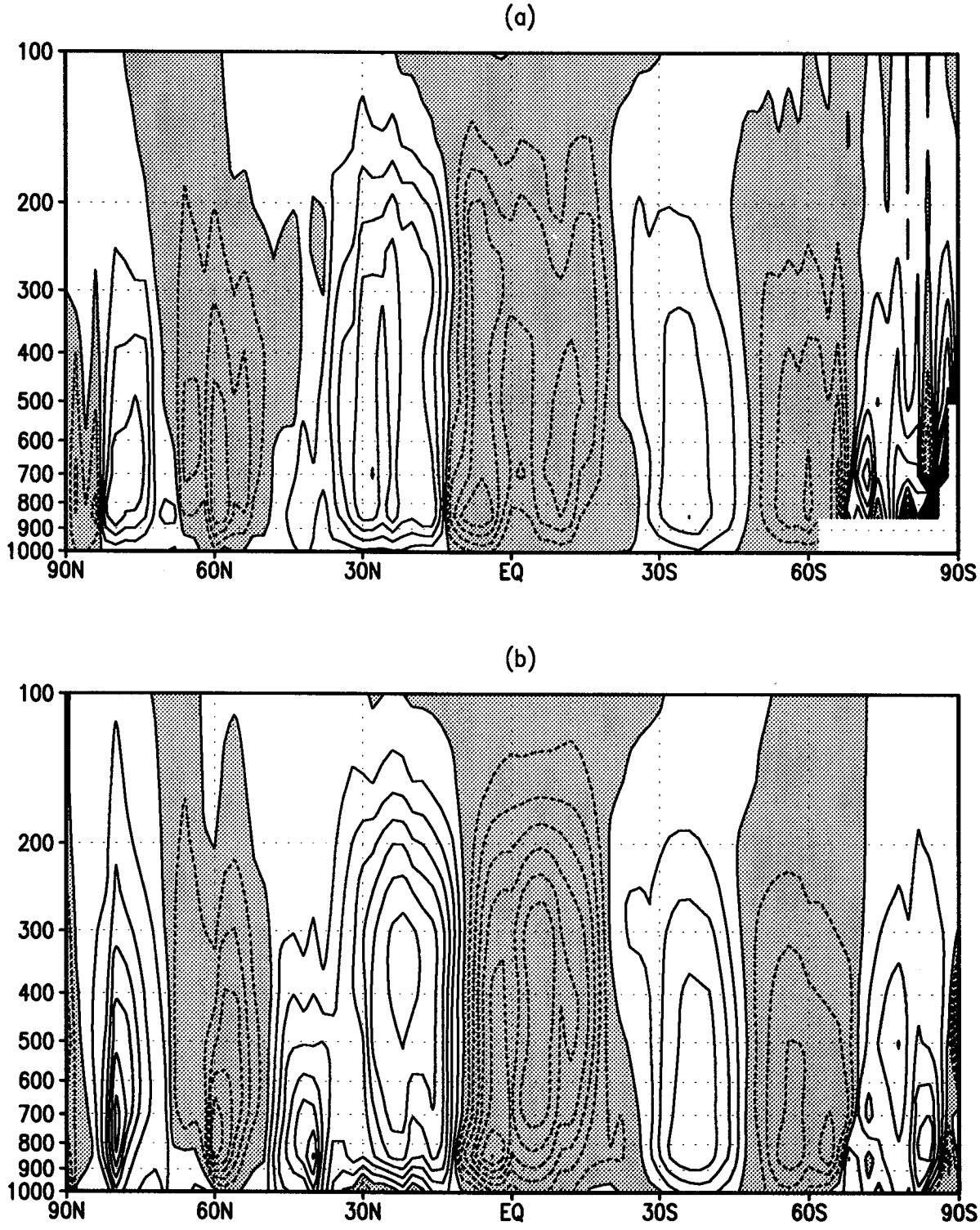


Figure 12: DJF climatology of zonal mean vertical–component of wind from (a) the GEOS–DAS and (b) from the ECMWF analysis. Contour interval is 40.0 nb/sec. Negative regions are shaded.

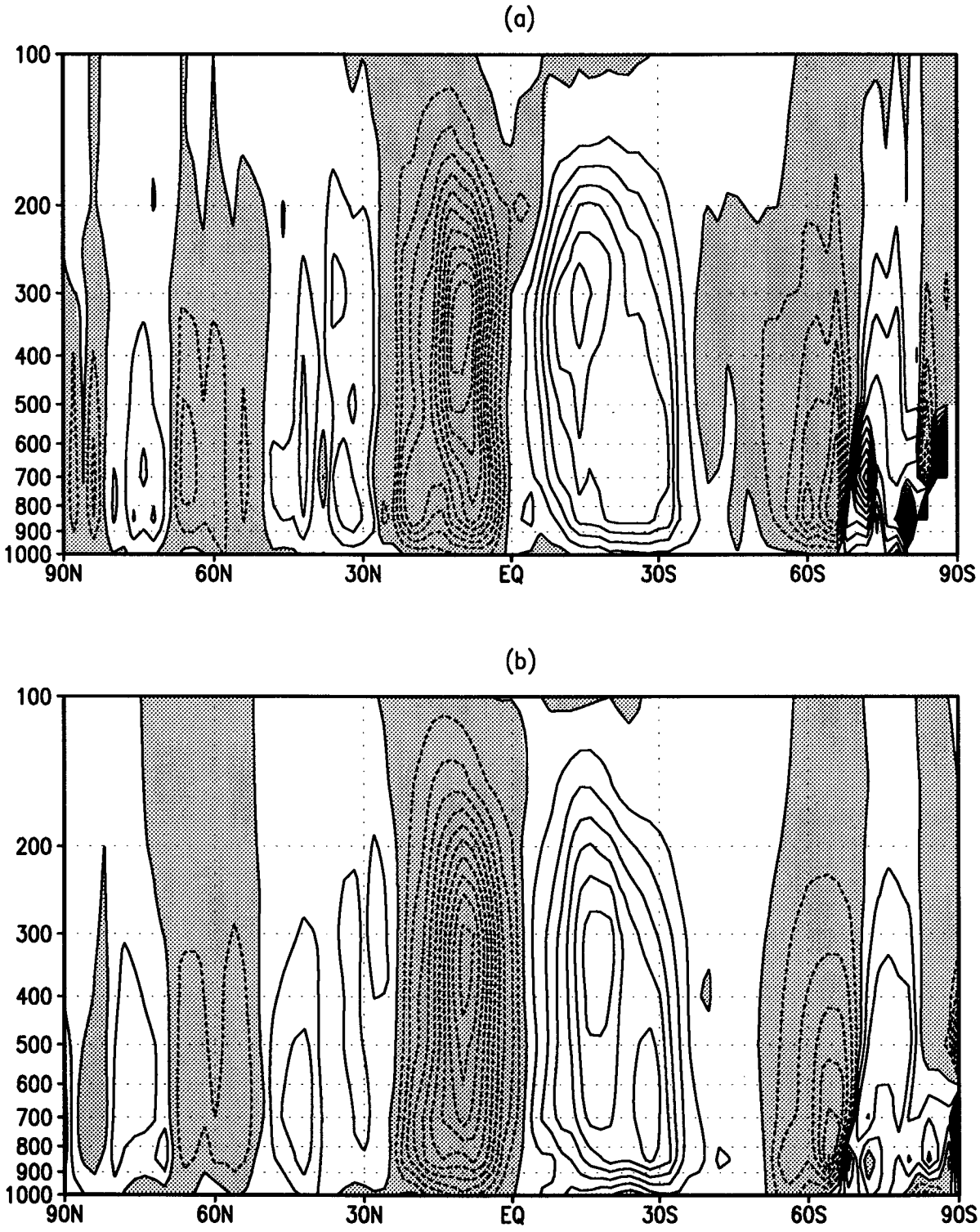


Figure 13: JJA climatology of zonal mean vertical–component of wind from (a) the GEOS–DAS and (b) from the ECMWF analysis. Contour interval is 40.0 nb/sec. Negative regions are shaded. Negative regions are shaded.

4.2.2 Global Prognostic

Figures 14–31 show the global distribution of selected GEOS–DAS quantities along with difference maps with ECMWF analyses, satellite and *in situ* observations.

The general layout for these figures is the same as for the zonal means, with the GEOS results on the top panels and difference fields on the bottom panels. The upper–level GEOS and ECMWF mean zonal winds are quite similar, with differences of less than 2 m/sec in most places. The largest differences occur during JJA (Fig. 16) in the tropics and at high latitudes. At 850 mb the largest difference in the zonal wind occur over the tropics, with the GEOS results showing larger westerlies over the western Pacific during DJF (Fig. 15) and larger easterlies over the eastern Pacific during JJA (Fig. 17). Considerable differences also exist over the tropical land masses.

The meridional wind differences (GEOS−ECMWF) show more complicated patterns than the zonal wind. During DJF the GEOS 200 mb meridional winds (Fig. 18) show a more northerly tendency just north of the equator, while at 850 mb (Fig. 19) the winds have a more southerly tendency in that latitude band; this is consistent with a weaker Hadley cell during this season (see Fig. 8). During JJA the differences at 200 mb (Fig. 20) are primarily at high latitudes and appear to be associated with problems in the ECMWF winds near the poles. At low levels, besides the large differences near the poles, the differences tend to be primarily tied to local topographic features.

Both the stream function and velocity potential fields (Figs. 22–27) show the largest differences (GEOS−ECMWF) over the tropics. Figure 24 shows clearly the signature of the weaker DJF Hadley cell in the GEOS results. The JJA velocity potential differences suggest the GEOS results tend to produce stronger convection over the western tropical Pacific and the Caribbean and stronger subsidence over the eastern tropical oceans. This is consistent with more intense tropical east–west (Walker) circulations in the GEOS results during JJA.

The sea–level pressure differences (Figs. 28–29) largely follow the regions of high orography. This must be primarily due to differences in the algorithm used to reduces the surface pressure to sea–level. The strong signature of the topography in the GEOS results suggests a problem with that algorithm. There are also very large scale (but small amplitude) differences, with the GEOS sea–level pressure being higher than that from ECMWF over much of the Pacific and Indian Oceans during both seasons.

The GEOS–DAS total precipitable water (TPW) over water surfaces is compared with SSM/I values in Figs. 30–31. The GEOS assimilation is drier over most of the oceans, except in middle and high latitudes. The largest differences occur in the tropics and subtropics where magnitudes often exceed $0.5 g/cm^2$. The differences show a clear seasonal shift, with the largest regions of significant dry bias (with respect to the SSM/I values) moving from the boreal subtropics during JJA to the tropics during DJF.

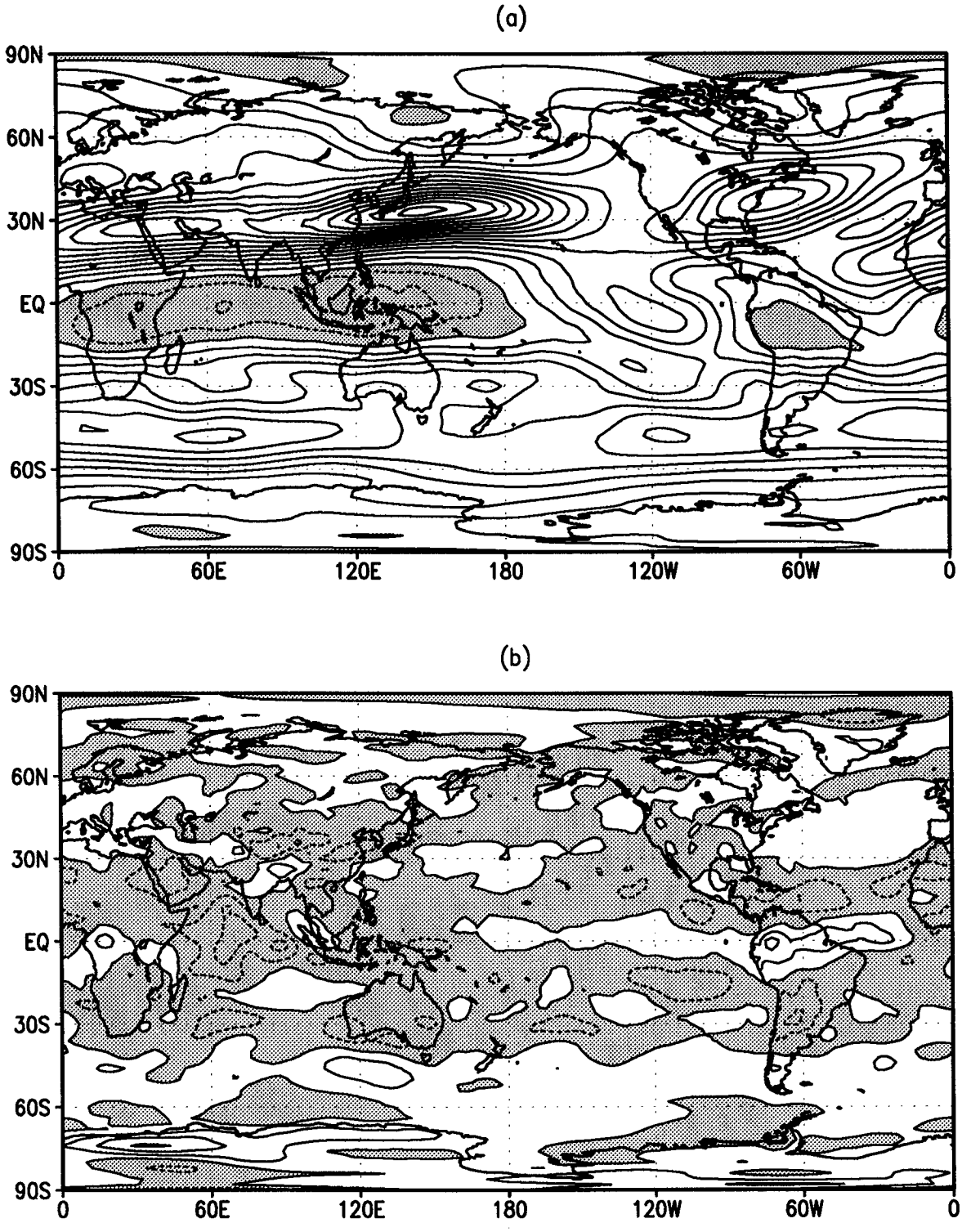


Figure 14: (a) DJF climatology of 200 mb u–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 5.0 m/sec in (a) and 2.0 m/sec in (b). Negative regions are shaded

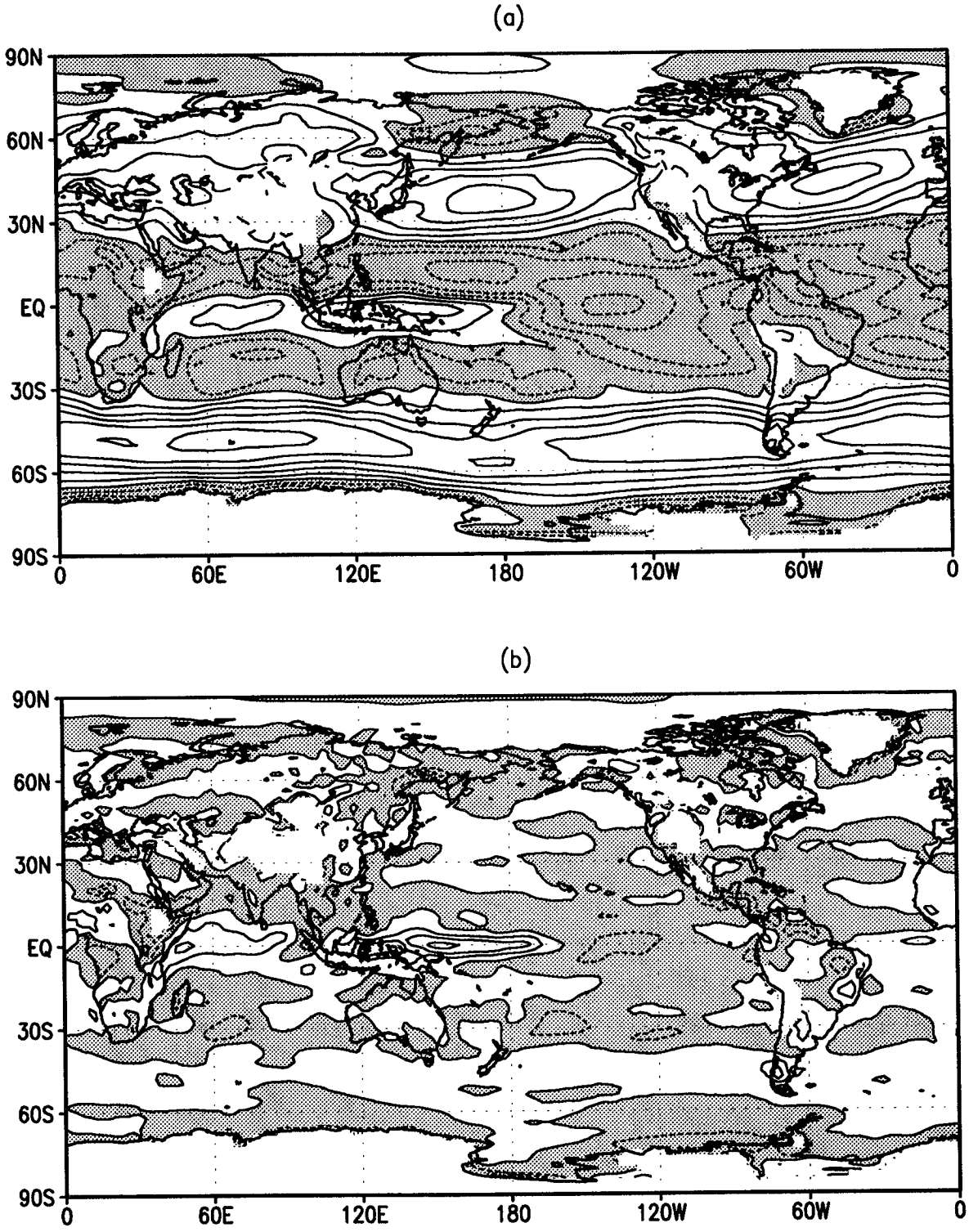


Figure 15: (a) DJF climatology of 850 mb u–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 3.0 m/sec in (a) and 2.0 m/sec in (b). Negative regions are shaded.

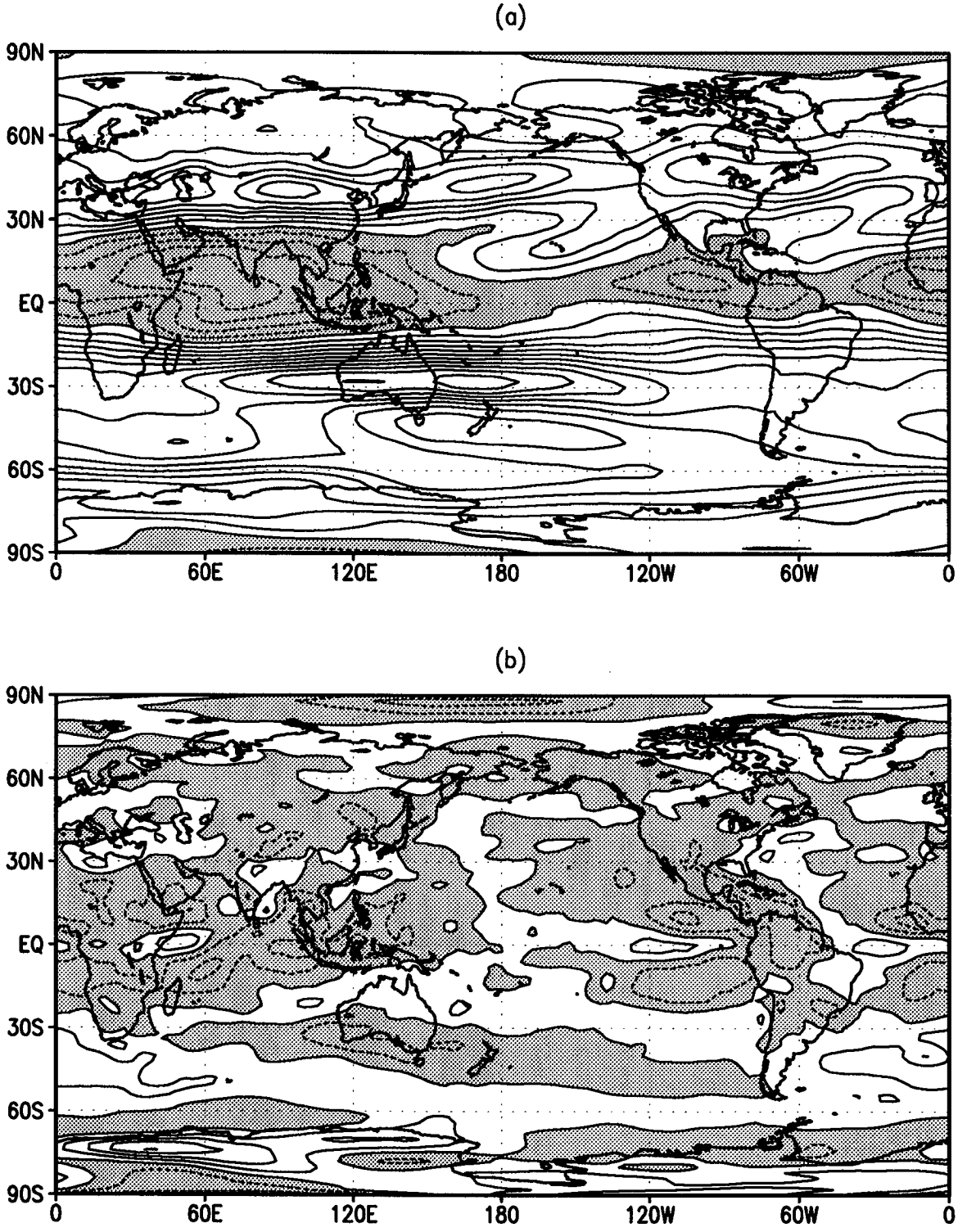


Figure 16: (a) JJA climatology of 200 mb u–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 5.0 m/sec in (a) and 2.0 m/sec in (b). Negative regions are shaded.

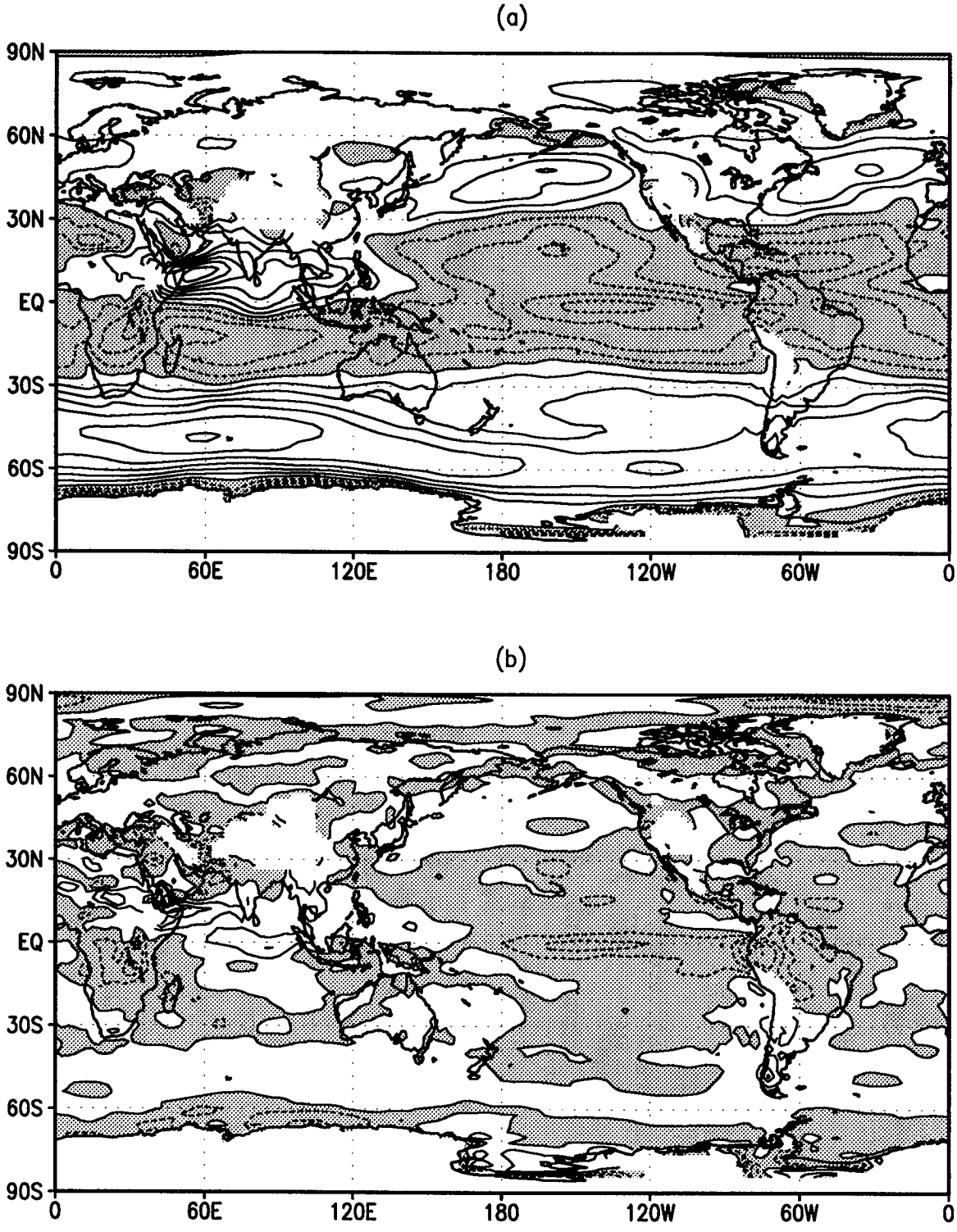


Figure 17: (a) JJA climatology of 850 mb u–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 3.0 m/sec in (a) and 2.0 m/sec in (b). Negative regions are shaded.

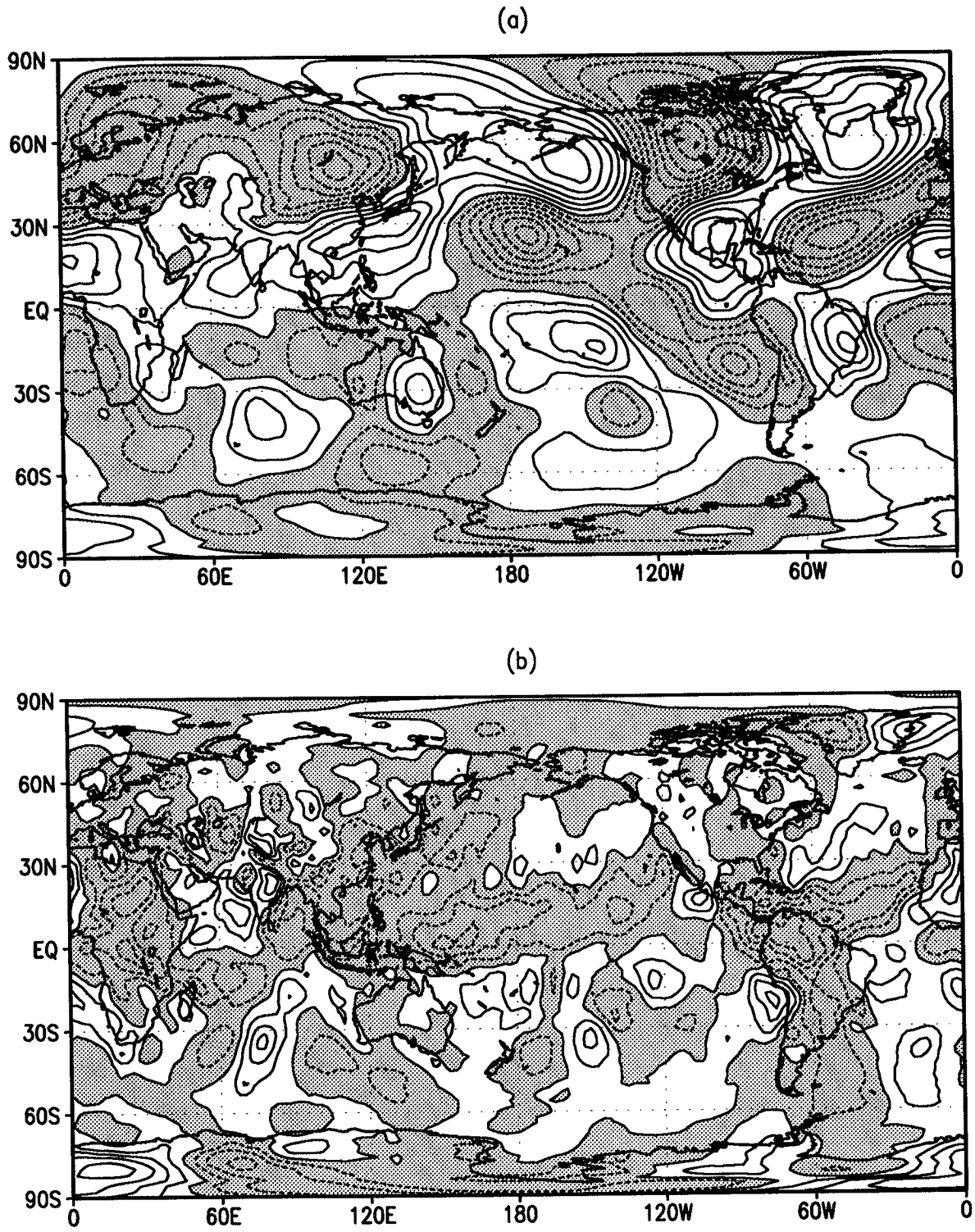


Figure 18: (a) DJF climatology of 200 mb v–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 2.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded.

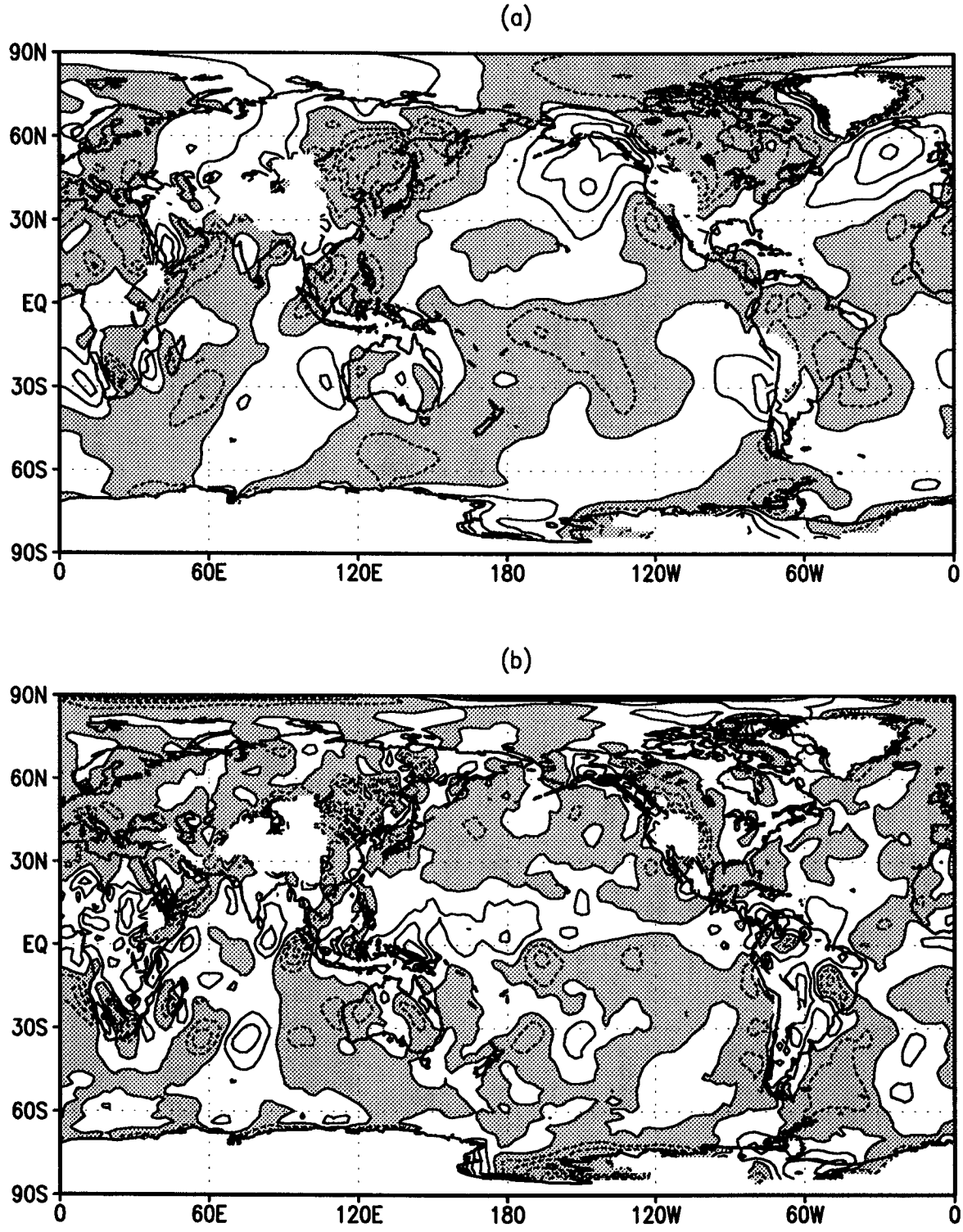


Figure 19: (a) DJF climatology of 850 mb v–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 2.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded.

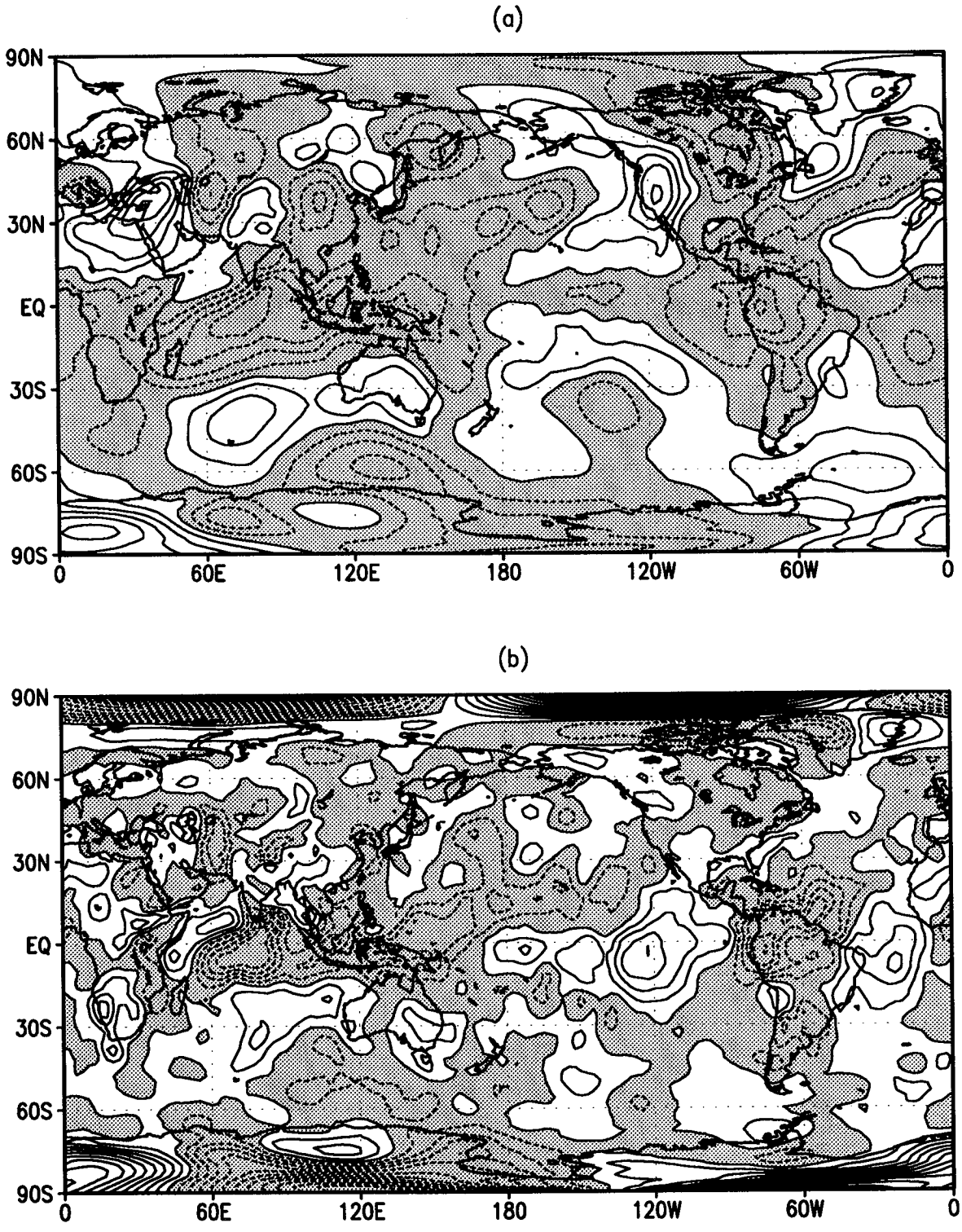


Figure 20: (a) JJA climatology of 200 mb v–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 2.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded.

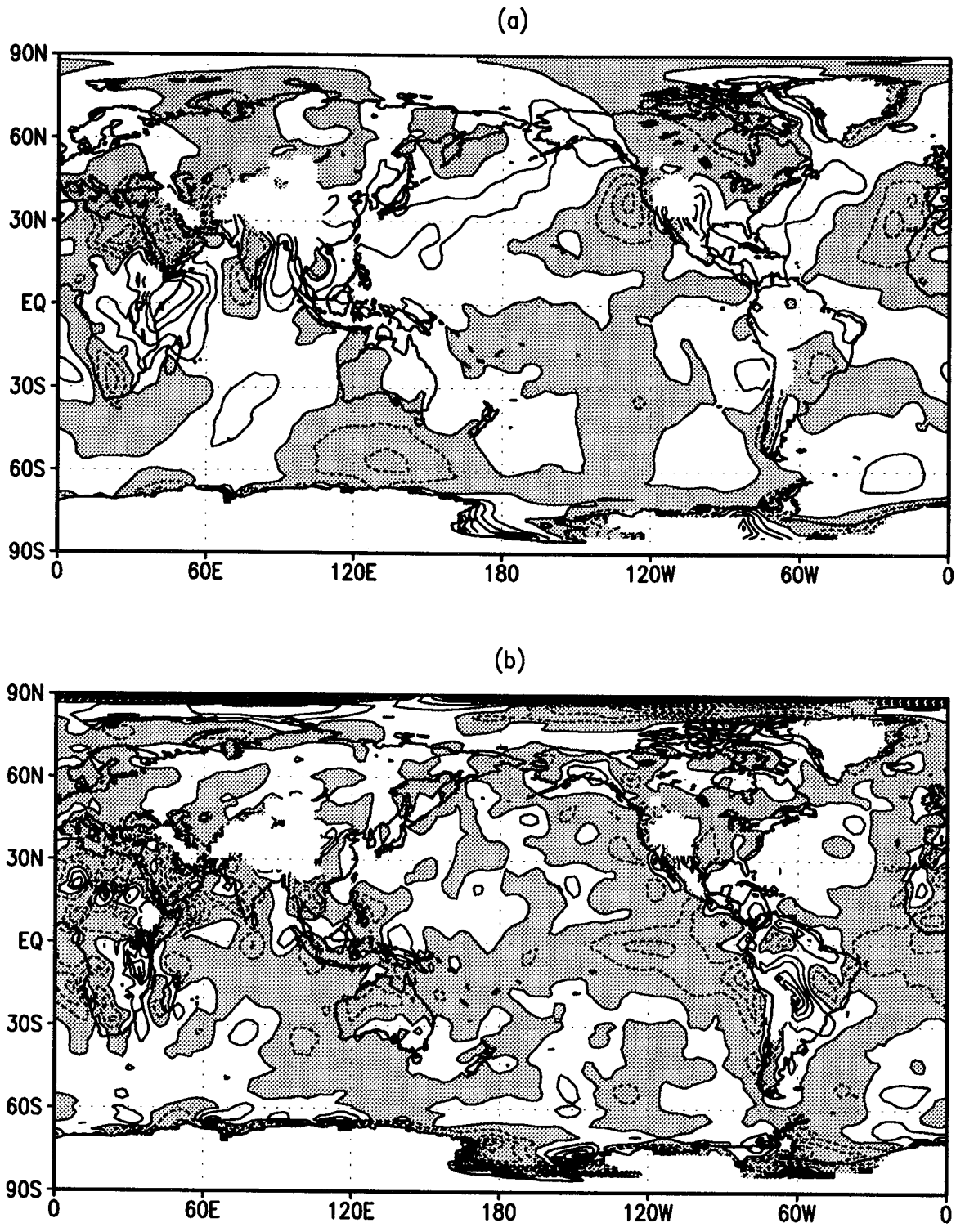


Figure 21: (a) JJA climatology of 850 mb v–wind from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is 2.0 m/sec in (a) and 1.0 m/sec in (b). Negative regions are shaded.

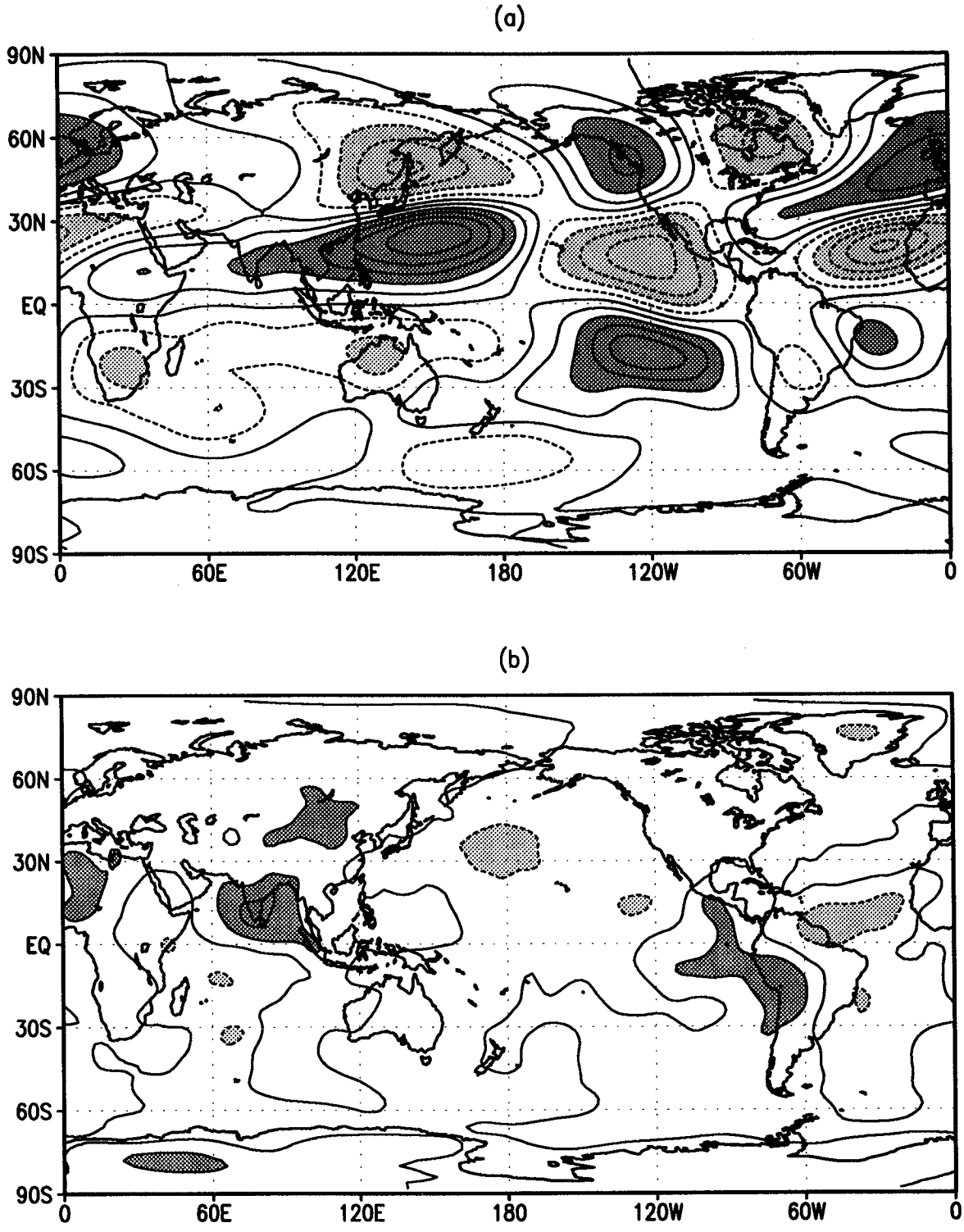


Figure 22: (a) DJF climatology of 200 mb eddy streamfunction from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is $5.0 \times 10^6\ m^2/sec$ in (a) and $2.0 \times 10^6\ m^2/sec$ in (b). The values greater than $10.0 \times 10^6\ m^2/sec$ and less than $-10.0 \times 10^6\ m^2/sec$ are shaded in (a) and the values greater than $2.0 \times 10^6\ m^2/sec$ and less than $-2.0 \times 10^6\ m^2/sec$ are shaded in (b).

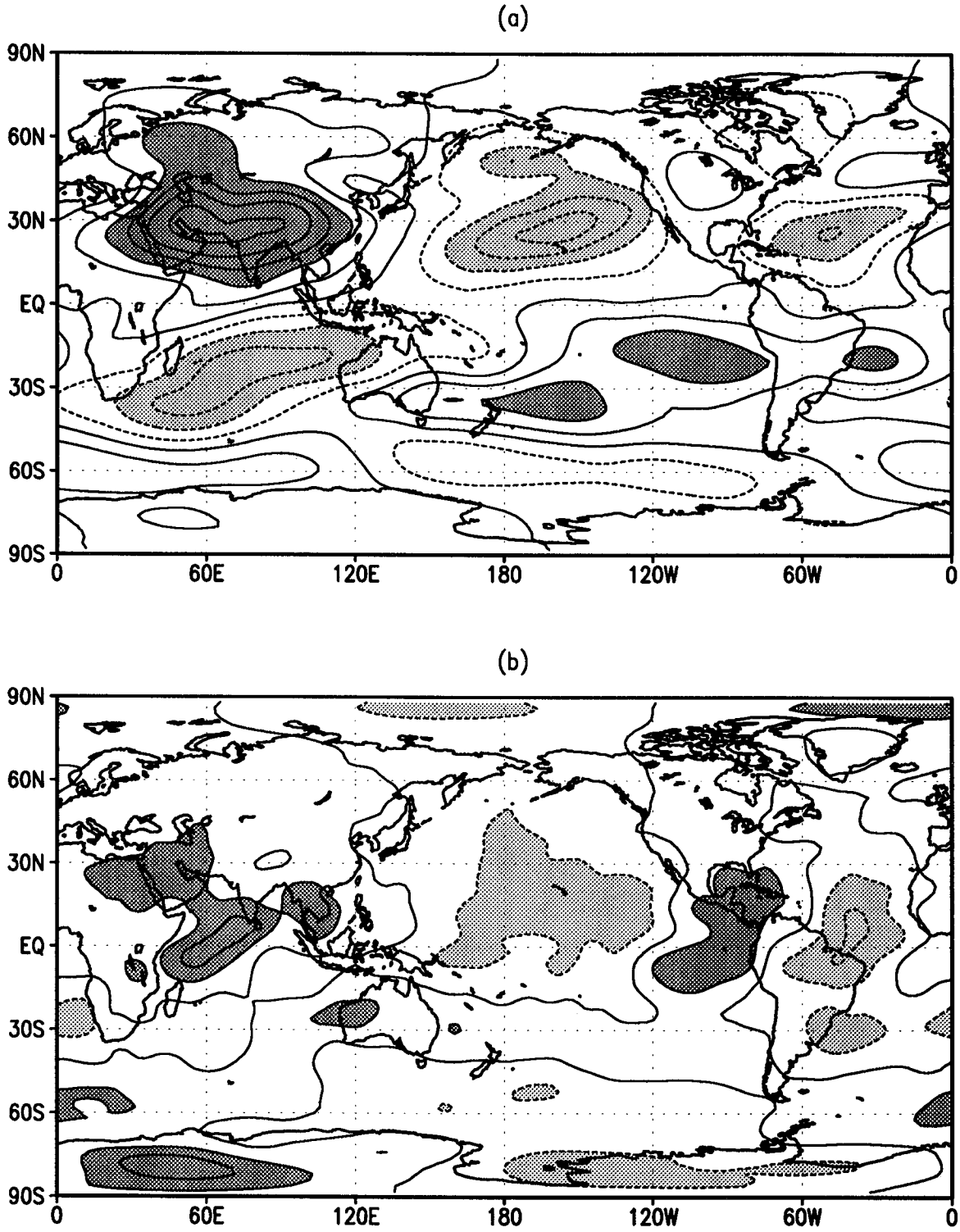


Figure 23: (a) JJA climatology of 200 mb eddy streamfunction from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is $5.0 \times 10^6\ m^2/sec$ in (a) and $2.0 \times 10^6\ m^2/sec$ in (b). The values greater than $10.0 \times 10^6\ m^2/sec$ and less than $-10.0 \times 10^6\ m^2/sec$ are shaded in (a) and the values greater than $2.0 \times 10^6\ m^2/sec$ and less than $-2.0 \times 10^6\ m^2/sec$ are shaded in (b).

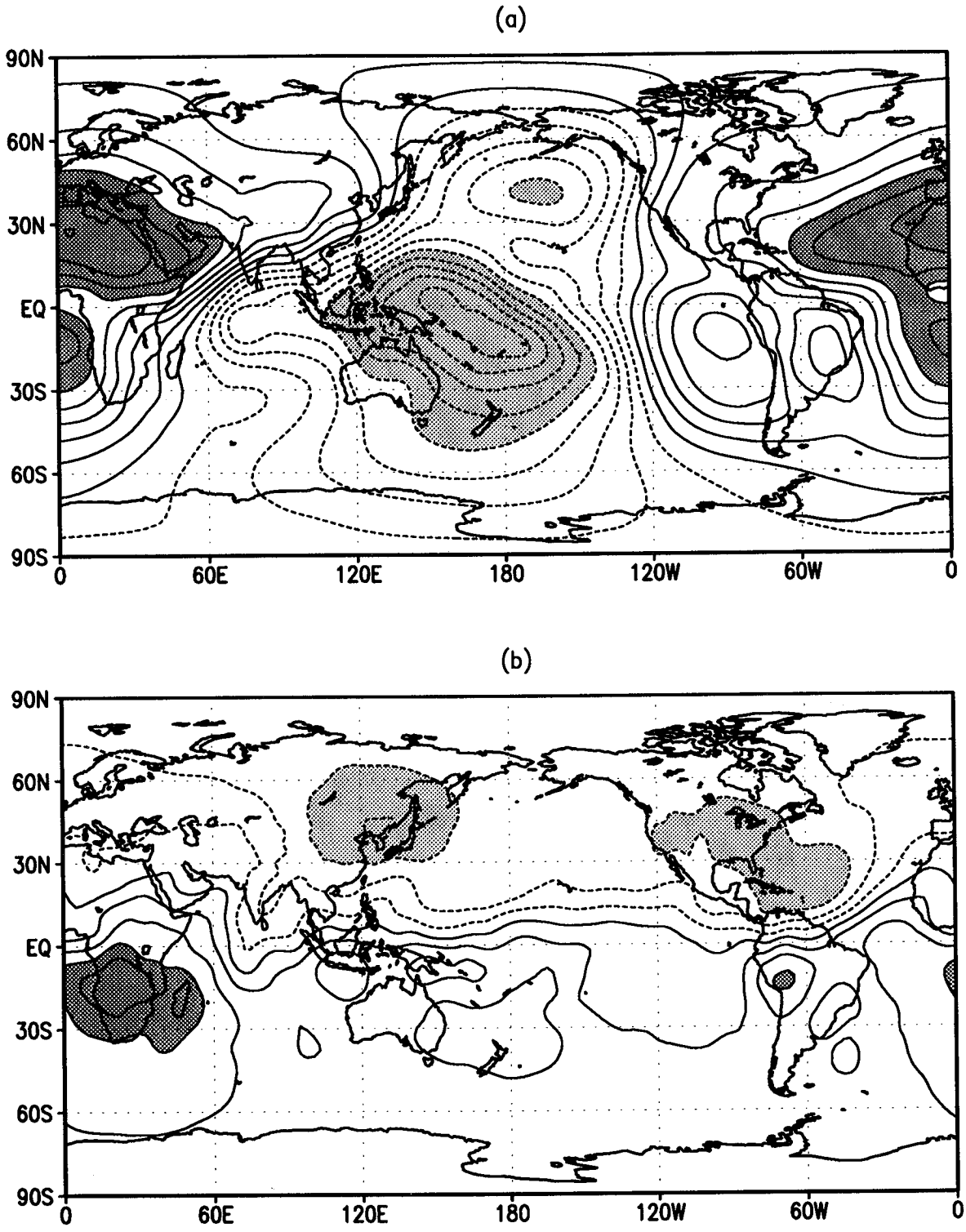


Figure 24: (a) DJF climatology of 200 mb velocity potential from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is $1.0 \times 10^6\ m^2/sec$. The values greater than $5.0 \times 10^6\ m^2/sec$ and less than $-5.0 \times 10^6\ m^2/sec$ are shaded in (a) and the values greater than $3.0 \times 10^6\ m^2/sec$ and less than $-3.0 \times 10^6\ m^2/sec$ are shaded in (b).

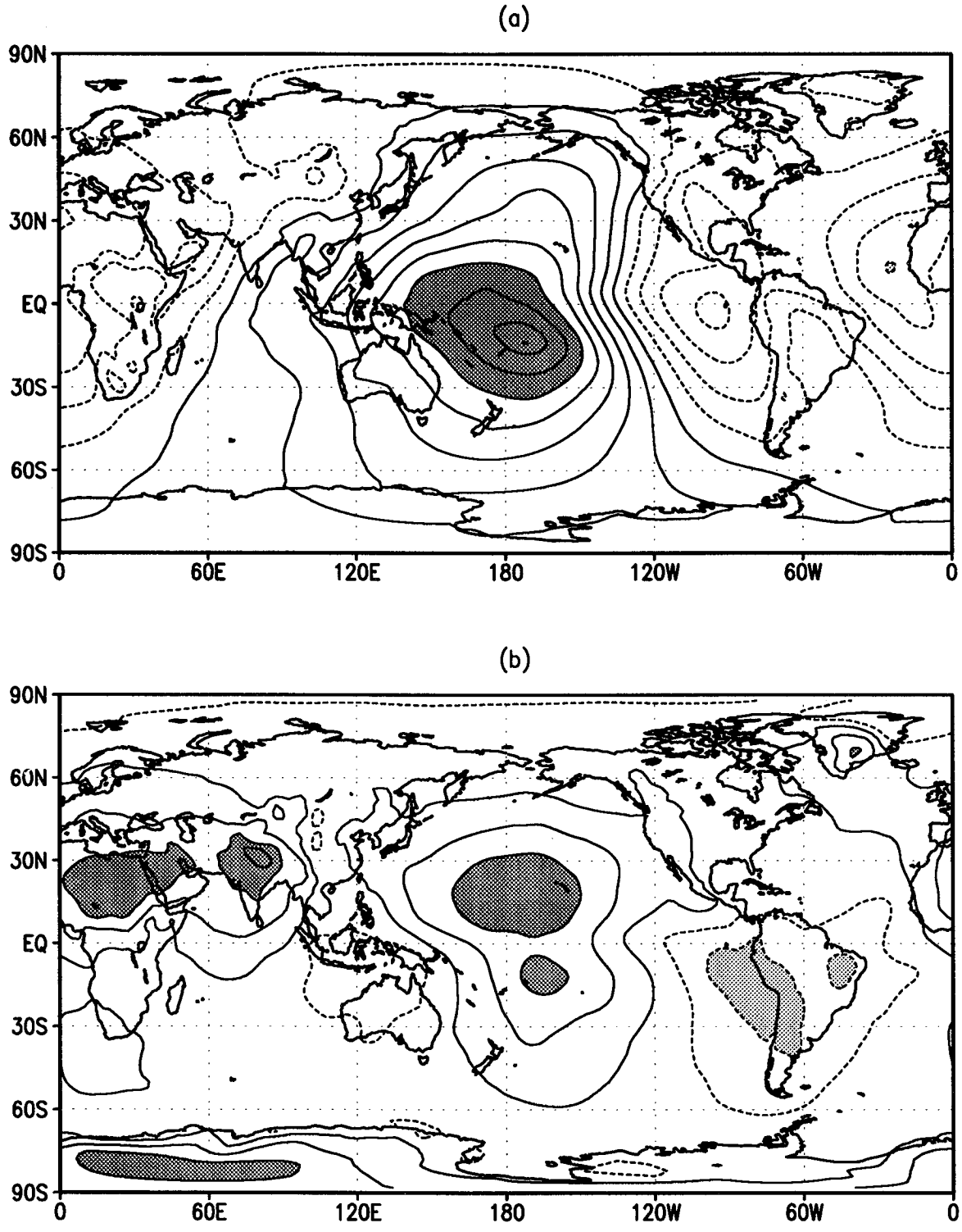


Figure 25: (a) DJF climatology of 850 mb velocity potential from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is $1.0 \times 10^6\ m^2/sec$. The values greater than $5.0 \times 10^6\ m^2/sec$ and less than $-5.0 \times 10^6\ m^2/sec$ are shaded in (a) and the values greater than $2.0 \times 10^6\ m^2/sec$ and less than $-2.0 \times 10^6\ m^2/sec$ are shaded in (b).

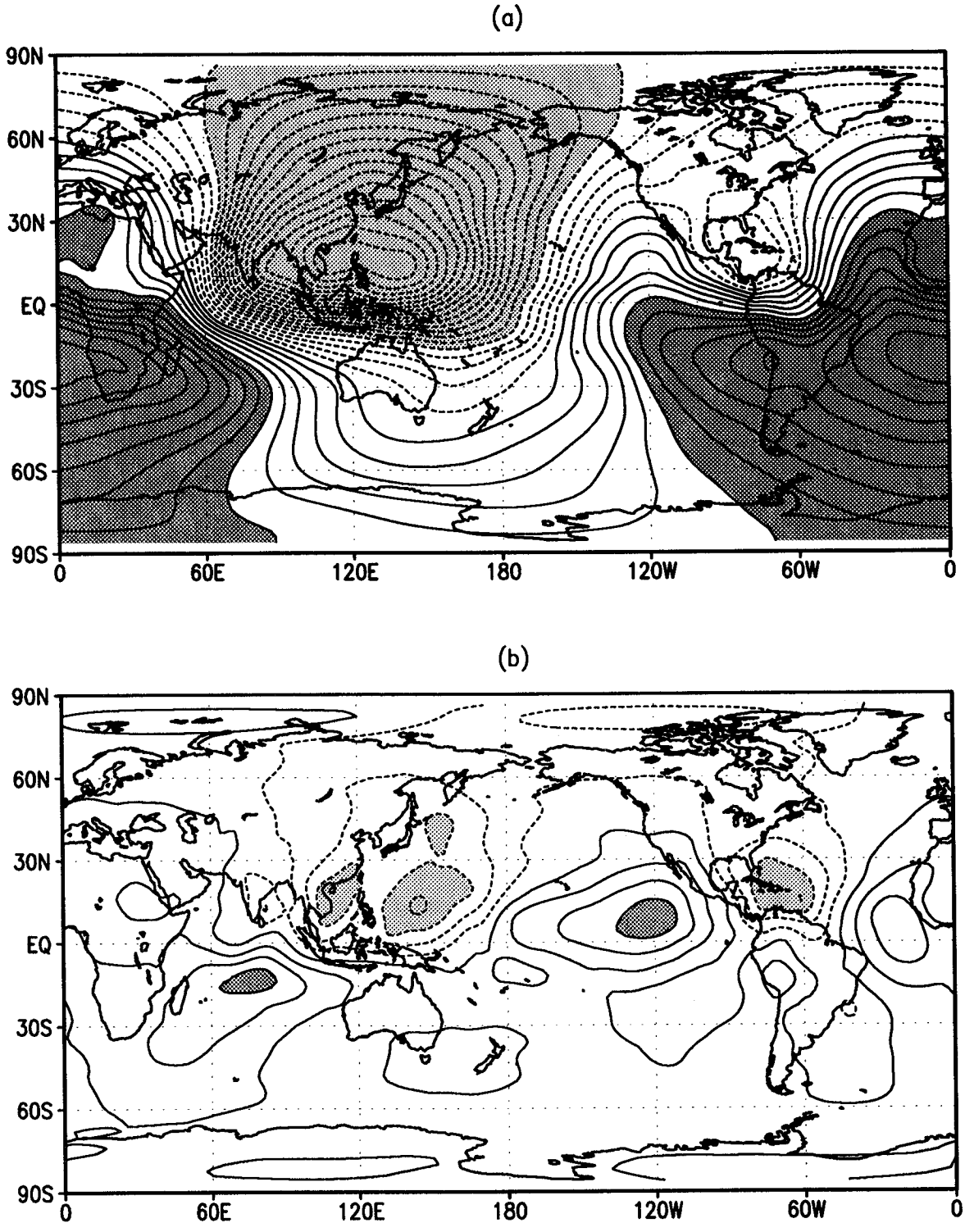


Figure 26: (a) JJA climatology of 200 mb velocity potential from the GEOS–DAS and (b) the difference fields (GEOS–ECMWF). Contour interval is $1.0 \times 10^6\ m^2/sec$. The values greater than $5.0 \times 10^6\ m^2/sec$ and less than $-5.0 \times 10^6\ m^2/sec$ are shaded in (a) and the values greater than $3.0 \times 10^6\ m^2/sec$ and less than $-3.0 \times 10^6\ m^2/sec$ are shaded in (b).

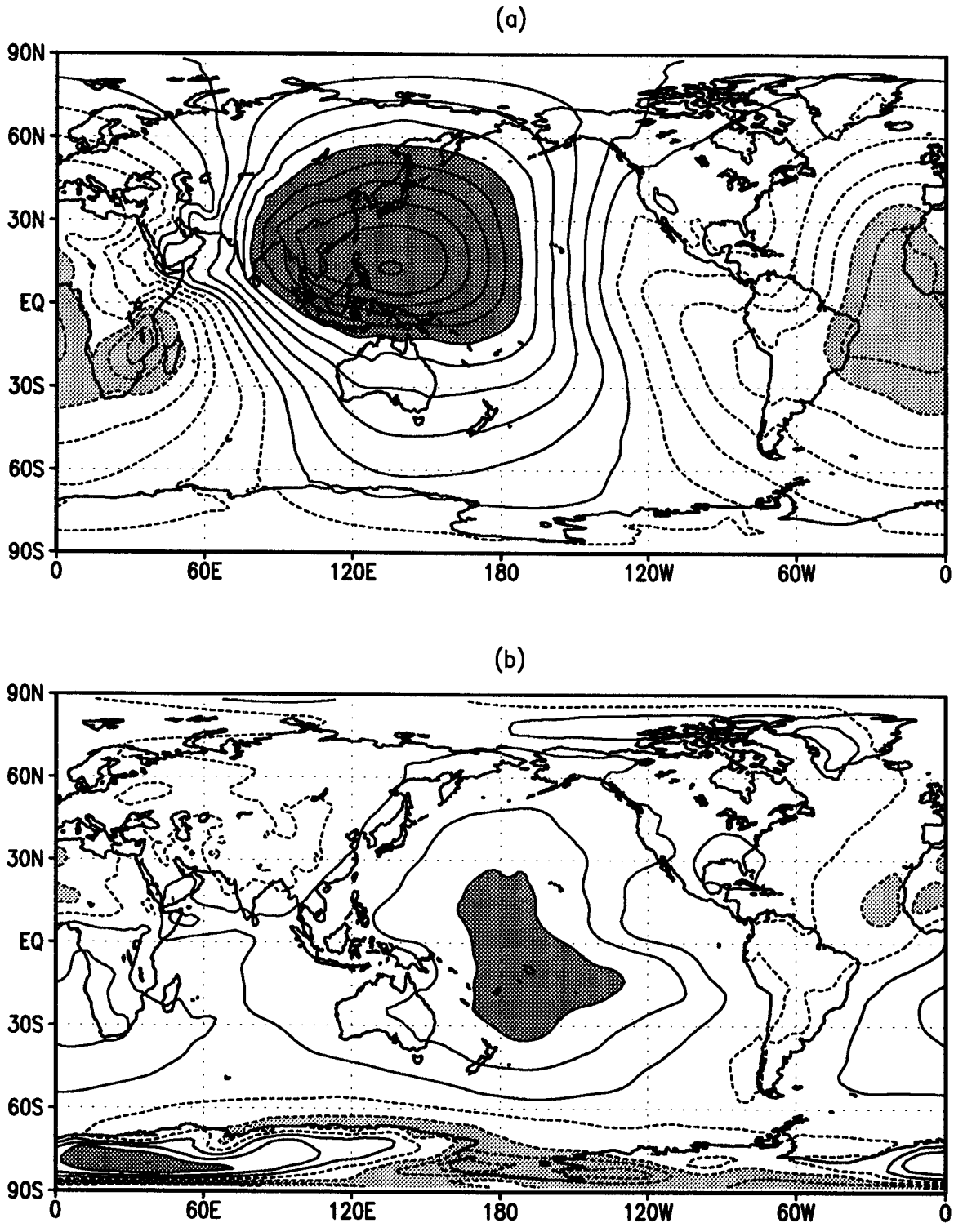


Figure 27: (a) JJA climatology of 850 mb velocity potential from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is $1.0 \times 10^6\ m^2/sec$. The values greater than $5.0 \times 10^6\ m^2/sec$ and less than $-5.0 \times 10^6\ m^2/sec$ are shaded in (a) and the values greater than $2.0 \times 10^6\ m^2/sec$ and less than $-2.0 \times 10^6\ m^2/sec$ are shaded in (b).

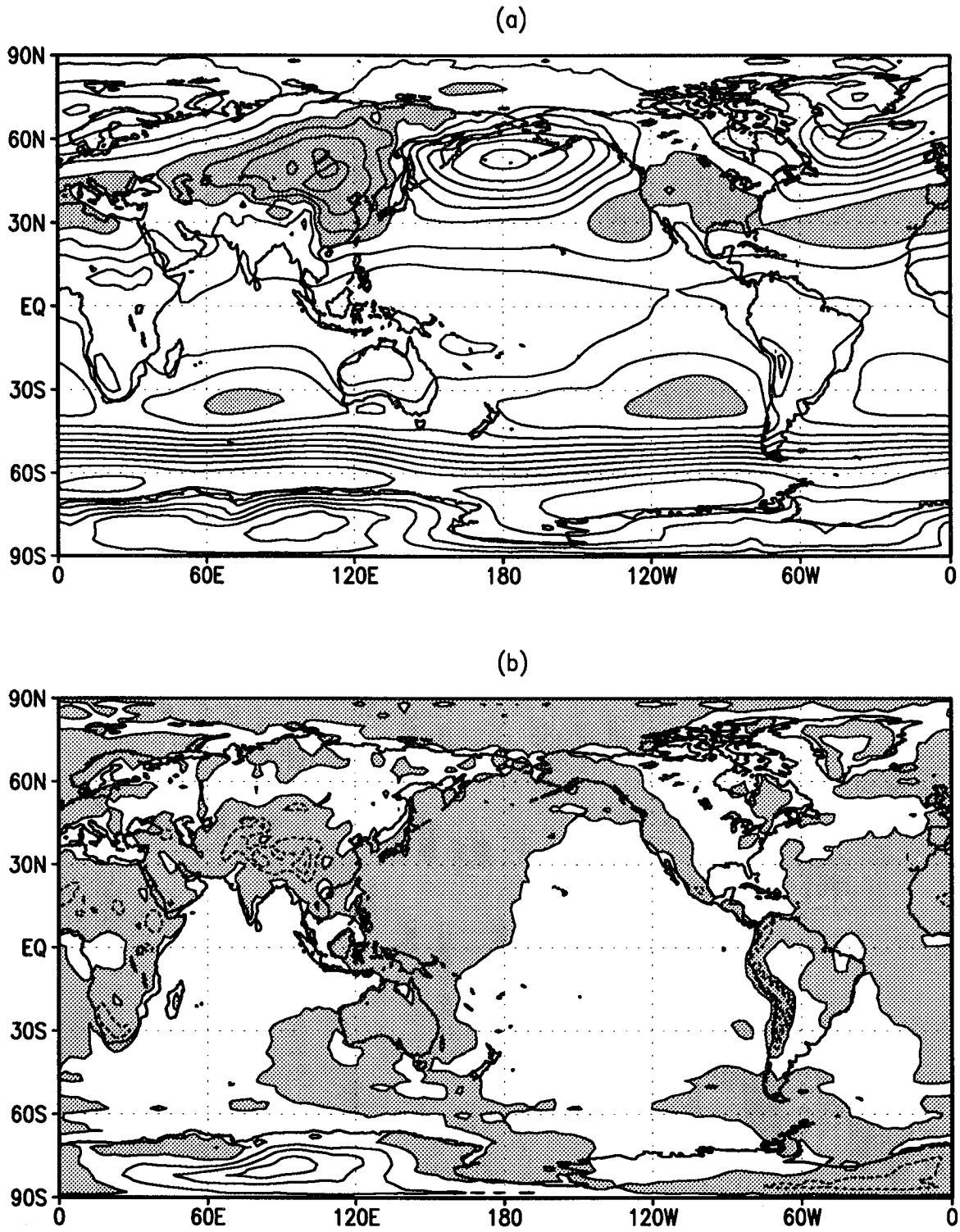


Figure 28: (a) DJF climatology of sea–level pressure from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour intervals are 4.0 *mb*. Values greater than 1020 mb are shaded in (a). Negative regions are shaded in (b).

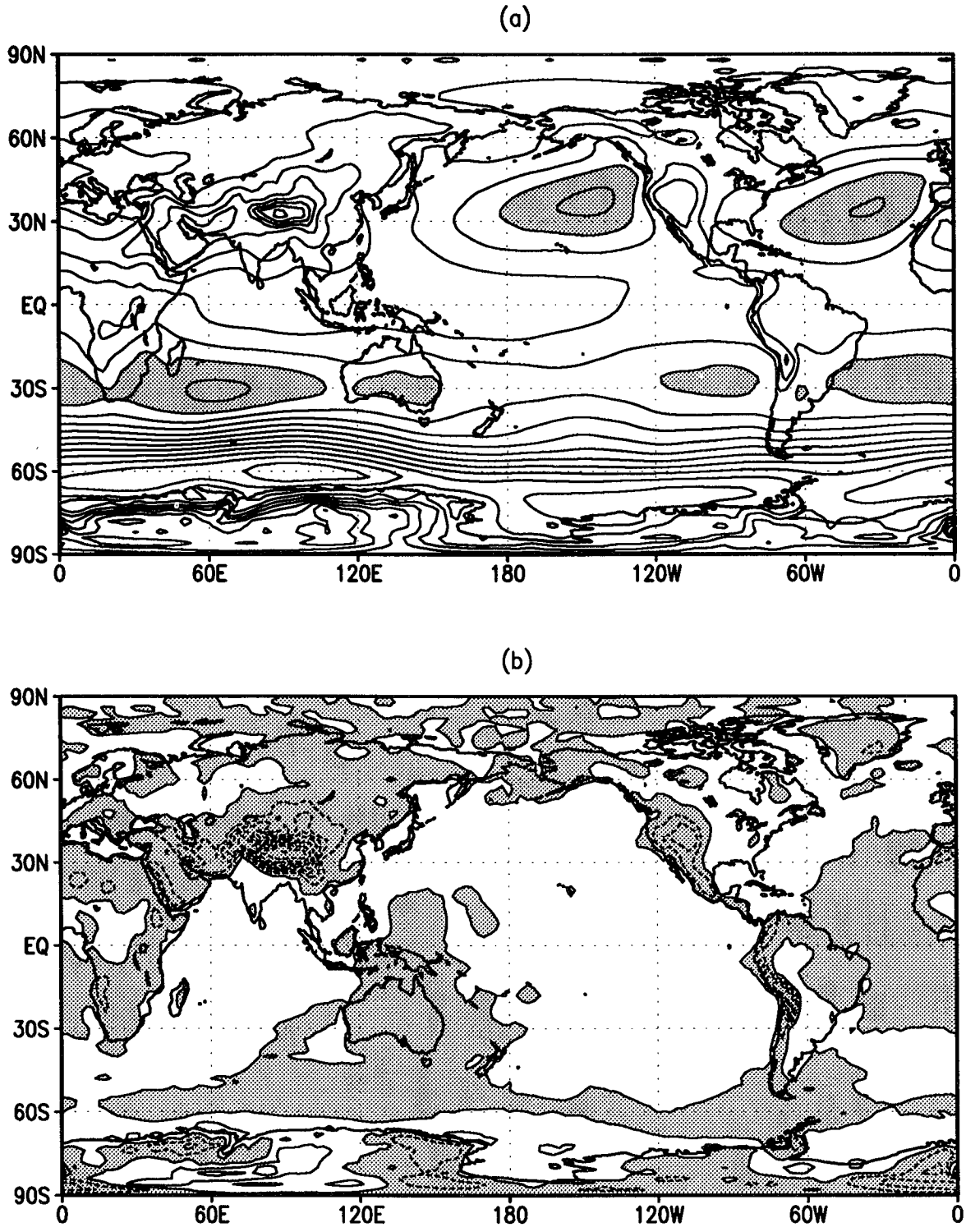


Figure 29: (a) JJA climatology of sea–level pressure from the GEOS–DAS and (b) the difference fields (GEOS−ECMWF). Contour interval is 4.0 *mb*. Values greater than 1020 mb are shaded in (a). Negative regions are shaded in (b).

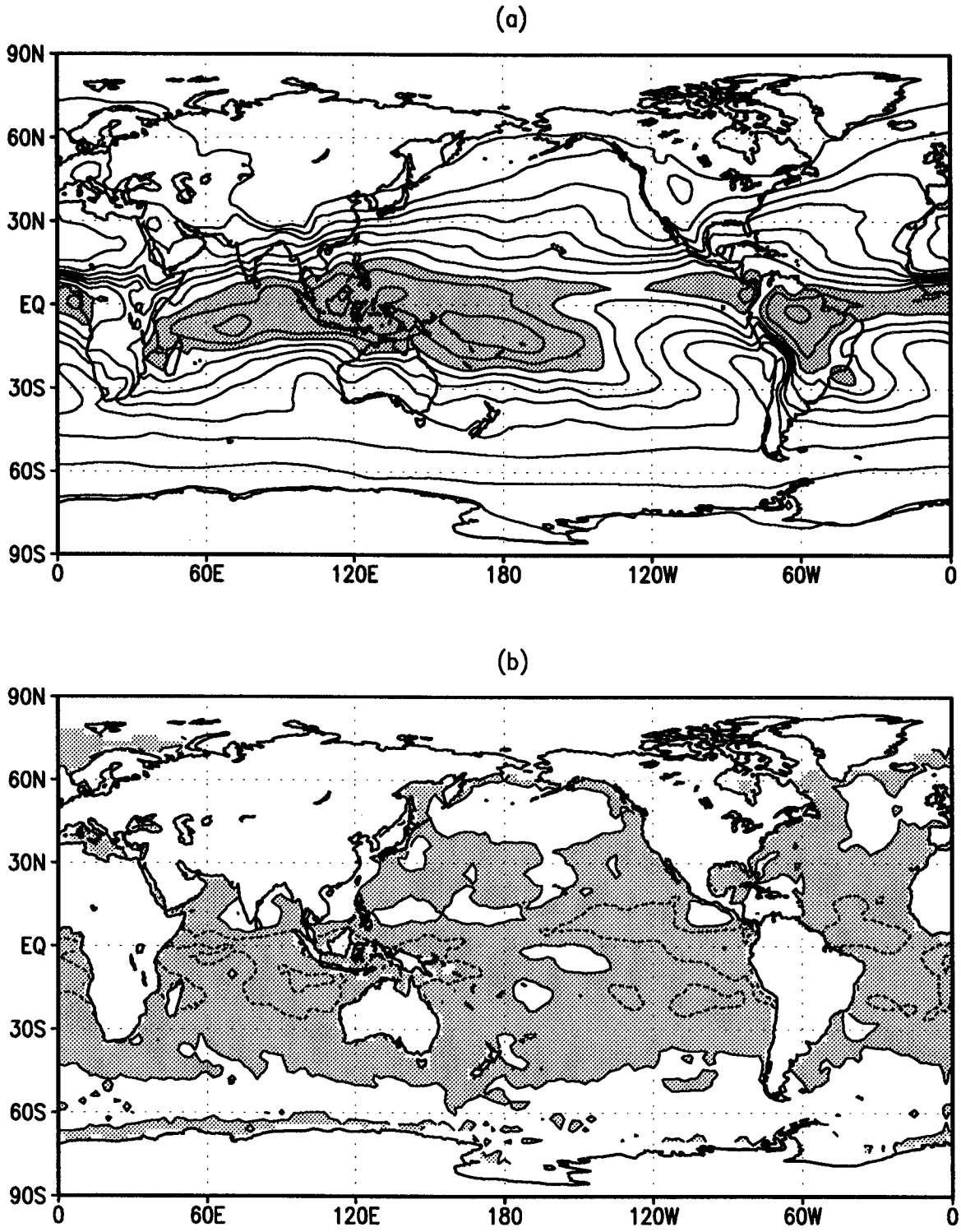


Figure 30: (a) DJF climatology (Dec 1987–Feb 1990) of total precipitable water from the GEOS–DAS and (b) the difference (GEOS–SSM/I) over the oceans. Contour interval is 0.5 gm/cm^2. The values greater than 4.0 gm/cm^2 are shaded in (a) and negative values are shaded in (b).

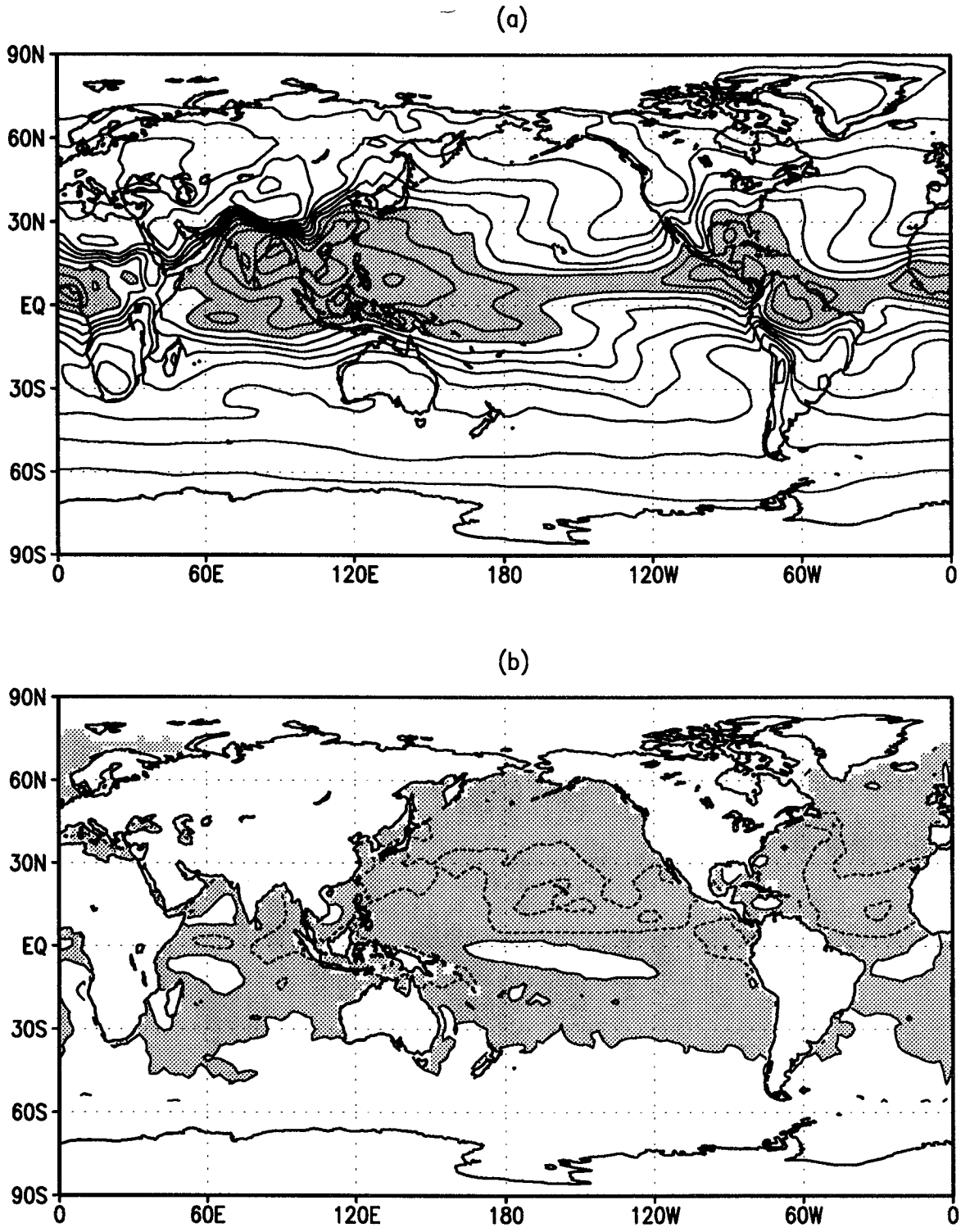


Figure 31: (a) JJA climatology of total precipitable water (Jul 1987–Aug 1989) from the GEOS–DAS and (b) the difference (GEOS–SSM/I) over the oceans. Contour interval is 0.5 gm/cm^2. The values greater than 4.0 gm/cm^2 are shaded in (a) and negative values are shaded in (b).

4.2.3 Surface Energy Balance

This section presents the global distribution of the components of the surface energy balance from the GEOS–DAS and the differences between the GEOS–DAS and COADS results (oceans only). Further comparisons are made with the parallel GEOS GCM simulation for the zonal means. Section 4.1 describes the COADS data and the parallel model run. The net surface energy balance can be expressed as

$$F_{sw\downarrow} - F_{sw\uparrow} - (F_{lw\uparrow} - F_{lw\downarrow}) - SH - LH = Q_{net} \tag{5}$$

where F_{sw} terms are the upward and downward shortwave fluxes, F_{lw} terms are the longwave fluxes, SH is the sensible heat flux, LH is the latent heat flux, and Q_{net} is the net surface heat flux. Since the sea–surface temperature (SST) is specified from observed monthly mean values, the net surface heating over the oceans serves mainly as a bulk diagnostic indicator of the interaction at the ocean–atmosphere interface. Note that, since the GEOS-1 DAS does not perform a surface temperature analysis, there is no explicit forcing term in (5) coming from the analysis (see section 2).

The latent heat flux is assumed to depend chiefly on the surface wind magnitude, the air–sea moisture gradient, and the surface layer stability. Because of the similarity in technique used to estimate evaporation by the GEOS GCM and by the COADS algorithm (see section 4.1), any differences between the GEOS–DAS and COADS results reflects differences in the near–surface gradients of humidity or winds and in surface roughness. Both seasons (Figs. 32–33) show maxima over the subtropical oceans in the regions of the subtropical highs. There is a strong seasonal shift associated with large evaporation rates over the summer continents and boreal winter east coasts of the two major continents. The GEOS–DAS results are generally close to the COADS results but tend to underestimate the magnitude, especially over the Kuroshio and Gulf currents during winter, and in the Northern Hemisphere (NH) over the east tropical Pacific during summer. Molod et al. (1995) provide a further discussion of these results and comparisons with the parallel GCM simulation (see also the zonal mean results below).

The flux of sensible heat at the surface is an important mechanism for energy transfer from the surface to the atmosphere, although over the oceans somewhat less so than the latent heat flux. The most important determinants of the sensible heat flux over the oceans are the air–sea temperature gradient and the ocean surface roughness. Oceanic sensible heat flux is generally from the surface to the atmosphere, but can be downward over sea ice or cold ocean currents. The global distribution of the sensible heat flux (Figs. 34–35), shows that, over the oceans, the regions of maximum flux occur over the Gulf Stream and Kuroshio current during winter, when the temperature gradients are strongest. Over land, the largest fluxes occur over the summer desert regions. The difference plot between the DAS and COADS, indicates an underestimation of the Gulf Stream sensible heat flux. As

was the case for the latent heat flux, the similarity in the parameterizations used to calculate sensible heat flux in the GEOS–DAS and by the COADS algorithm implies that differences in resultant fluxes must be due to differences in the air–sea temperature gradient and slight differences in the determination of surface roughness.

The major energy source term for the earth's surface, providing energy for sensible and latent heat fluxes as well as the surface heating and terrestrial emissions, is the net shortwave radiation at the surface. The major factors influencing the distribution of net surface shortwave flux are the incoming shortwave at the top of the atmosphere, the surface albedo, and the absorption and reflection by clouds. The net shortwave and difference with the COADS results are shown in Figs. 36 and 37 for the two extreme seasons. The largest values occur over the summer hemisphere subtropics, corresponding to the incident shortwave on the atmosphere, with a local minimum over the intertropical convergence zone (ITCZ) corresponding to reflection by clouds, and a final minimum over the dark winter pole. The seasonal differences in the tropics are strongly tied to the movement of the ITCZ. Compared with COADS, the GEOS–DAS results tend to overestimate the net surface shortwave in the extratropics during both seasons, with a greater disparity in the summer. There is an underprediction of net surface shortwave in the tropics, especially during boreal summer over the western Pacific and the Caribbean Sea. These differences are related to an underestimation of cloud cover in the extratropics, and an overestimation of the cloud cover in the tropics by the DAS (see Molod et al. 1995).

The longwave radiation at the surface is the difference between the upward surface emission and the downward emission by clouds and the moist atmosphere. The net surface longwave radiation (upward minus downward flux) is shown in Figs. 38 and 39. The net longwave is a smaller component of the surface heat budget than the net surface shortwave and exhibits a smaller latitudinal and seasonal variability, particularly over the oceans. The largest seasonal variations occur over the extratropical storm track regions of both hemispheres, the polar regions, and over the desert land areas. The smaller values during summer over the storm tracks suggest that the influence of clouds dominates any differences due to the seasonal variations in sea–surface temperatures. The GEOS–DAS overpredicts (compared with COADS) the net surface longwave during both seasons, especially over the boreal winter storm tracks. The SST values used by COADS and GEOS–DAS are essentially the same, so the discrepancy in surface longwave must be in an underestimation of the downward component, either due to clouds or due to water vapor emissions. In the extratropics, Molod et al. (1995) found that the TOA longwave cloud radiative forcing was underestimated by the DAS, which implies a low estimate of cloud–related downward emissions as well. In the tropics, the low downward longwave emissions are probably related to the dry bias (Figs. 30–31).

The net surface heat flux, Q_{net}, represents the residual heating available to raise the ground temperature over land or the local sea surface temperature. Figures 40 and 41 show a positive maximum in the summer hemisphere throughout the extratropics and subtropics,

corresponding to the maximum solar irradiance and the relative minimum of evaporative flux in the extratropics. The positive values extend to the equator over the eastern oceans, corresponding to local minima in evaporative and longwave flux. The boreal winter fluxes are dominated by the strong negative fluxes over the Kuroshio and Gulf streams. The main difference with the COADS results occurs in the summer hemisphere where the GEOS–DAS tends to overestimate the net heat flux into the oceans, corresponding to the overestimate of net surface radiative flux. Large differences also occur over the tropical and subtropical Indian and Pacific Oceans and the Caribbean Sea during JJA; these are associated with the differences in the net longwave flux mentioned earlier (see Fig. 39).

The following zonal mean plots are computed over oceans only to allow comparisons with the COADS data. The figures include the results from the parallel GEOS GCM simulation with the same boundary conditions used in the GEOS–DAS (see section 4.1).

The zonal mean sensible heat flux (Fig. 42) shows a maximum in the winter hemisphere extratropics throughout the year, corresponding to the largest air–sea temperature gradients. There exists a fairly good correspondence between GCM simulation, assimilation and COADS, with a tendency for the GCM to overpredict the flux in the extratropics. The sparcity of COADS data in the Southern Ocean minimizes the significance of the differences in that region. The zonal mean latent heat flux clearly shows the maxima over the subtropical oceans, where precipitation is low and the air above the surface is relatively dry due to the subsidence associated with the subtropical highs. The winter hemispheres tend to have a greater latent heat flux, presumably associated with the higher wintertime surface winds. The COADS and GEOS–DAS results agree quite well, though there is a tendency for the DAS fluxes to be weaker. The GCM results show considerably stronger fluxes; this appears to be primarily due to the stronger simulated surface winds, which are corrected in the GEOS–DAS (see Molod et al. 1995).

The zonal mean net shortwave and longwave radiation at the surface are shown in Fig. 43. The GEOS–DAS and GEOS–GCM fluxes are quite similar, except in the NH extratropics during JJA, where the GEOS–DAS has significantly larger fluxes, apparently associated with reduced cloudiness in the DAS. In this latitude band the GCM simulation is actually closer to the COADS results. The large differences between the DAS and COADS in the summer hemisphere net surface short wave flux noted earlier is quite evident in the zonal averages. Note the substantial poleward shift in the maxima compared with COADS. In all three estimates, the zonal mean net surface longwave flux shows little latitudinal or seasonal variation, but the COADS results show consistently lower values.

The zonal mean net surface radiation (Fig. 44) is dominated by the shortwave component and, therefore, exhibits the same latitudinal and seasonal behavior, as seen in the shortwave flux (Fig. 43). The DAS and GCM estimates of the net surface radiation are closer to the COADS data than the individual components due to the compensating overprediction errors in both longwave and shortwave. The zonal mean net heat flux (Fig. 44), shows a

positive maximum in the summer hemisphere throughout the extratropics. The zonal mean flux has a slight local maximum near the equator, corresponding to the local minimum in the zonal mean evaporative flux. The implied ocean fluxes, therefore, are from summer to winter hemisphere, with a slightly greater hemispheric disparity during the southern summer; however, the known sensitivity of the implied ocean fluxes to small changes in the zonal averages requires a more thorough analysis of the quality of these results.

In the tropics the DAS and GCM agree relatively well with the COADS results. In the summer hemisphere both the GCM and DAS tend to overestimate the net heat flux into the oceans, corresponding to the overestimate of net surface radiative flux. The implication is that the parameterization of the net surface radiation, probably in the simulation of cloud–radiative interaction, is biased even when given a more accurate atmospheric temperature and moisture profile by the DAS. In the NH winter the GCM shows a stronger negative surface heat flux than either the DAS or COADS data, due in large part to the overestimation of the latent heat flux.

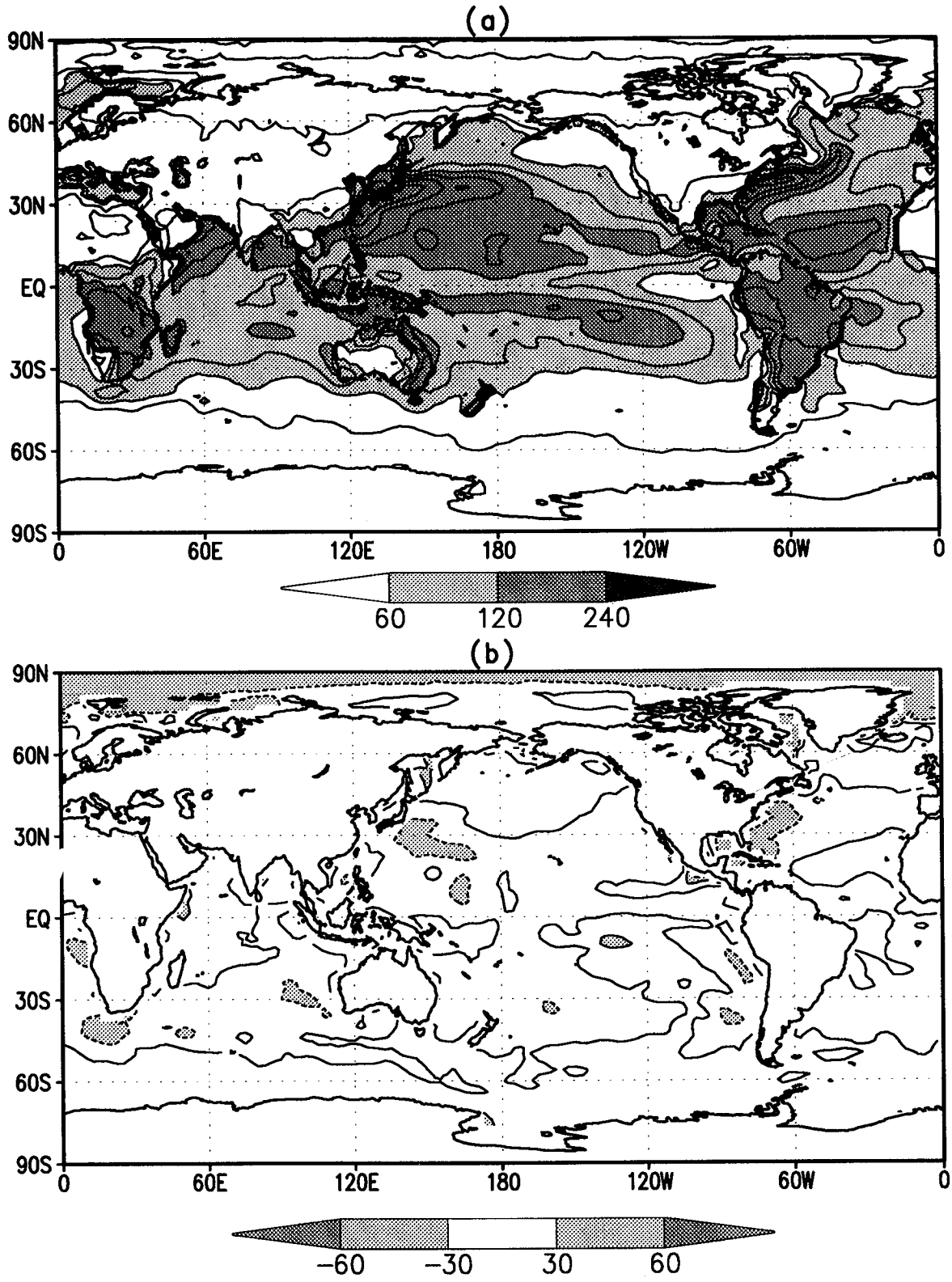


Figure 32: The latent heat flux for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Contour interval is 30 W/m^2.

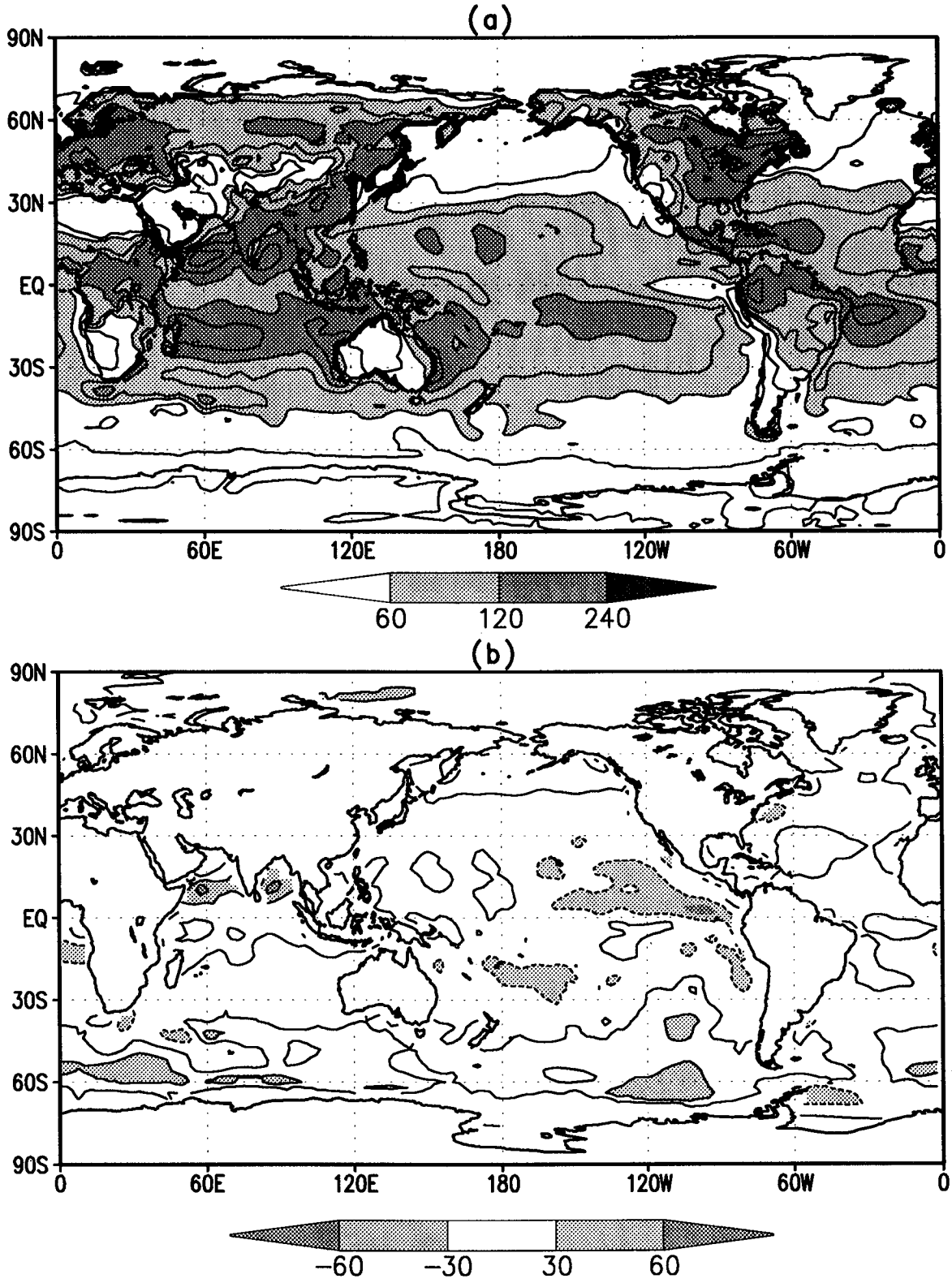


Figure 33: The latent heat flux for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Contour interval is 30 W/m^2.

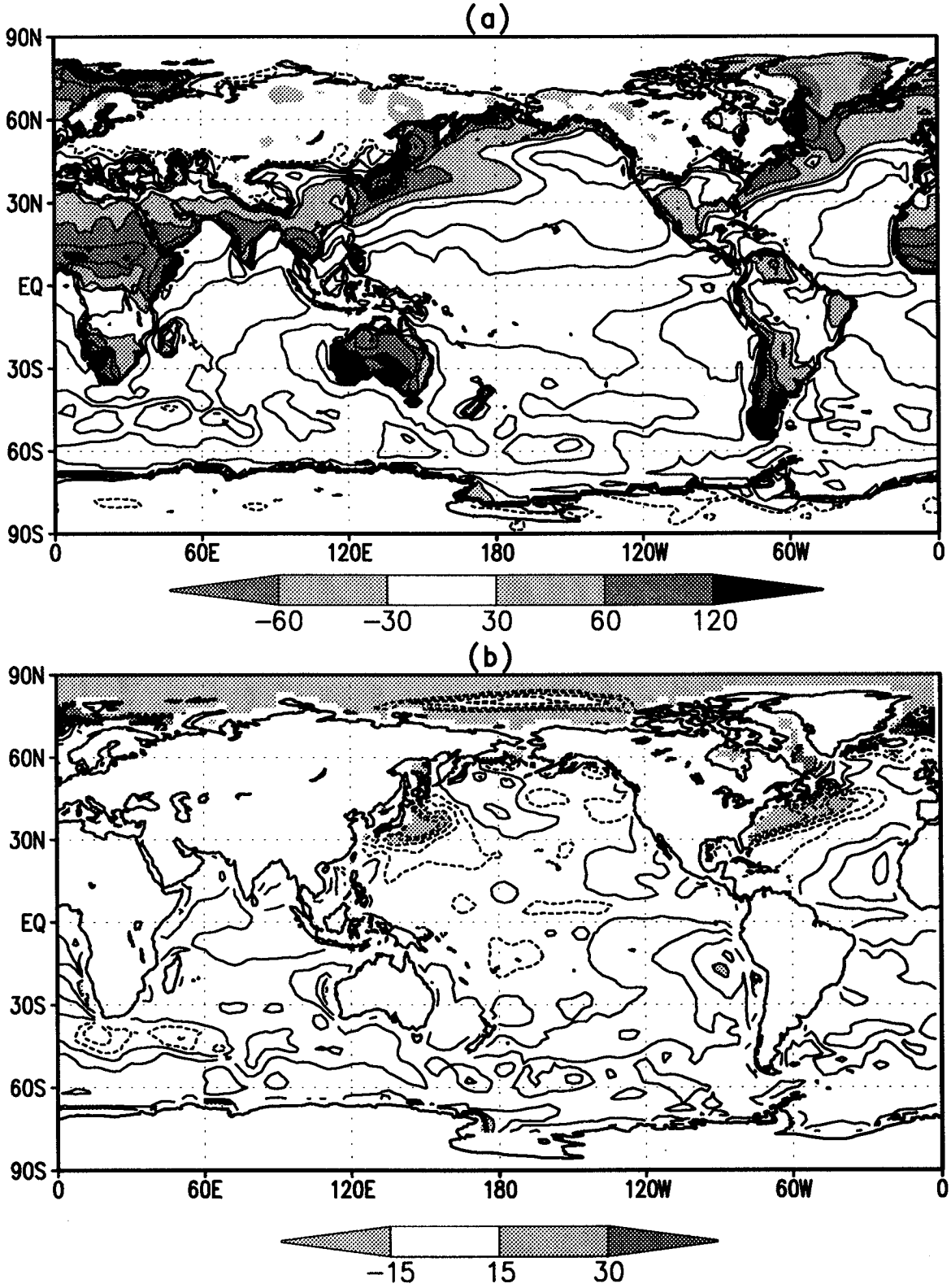


Figure 34: The sensible heat flux for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour intervals are -5 0 5 10 20 30 60 90 120 150 180 in (a) and every 5 in (b).

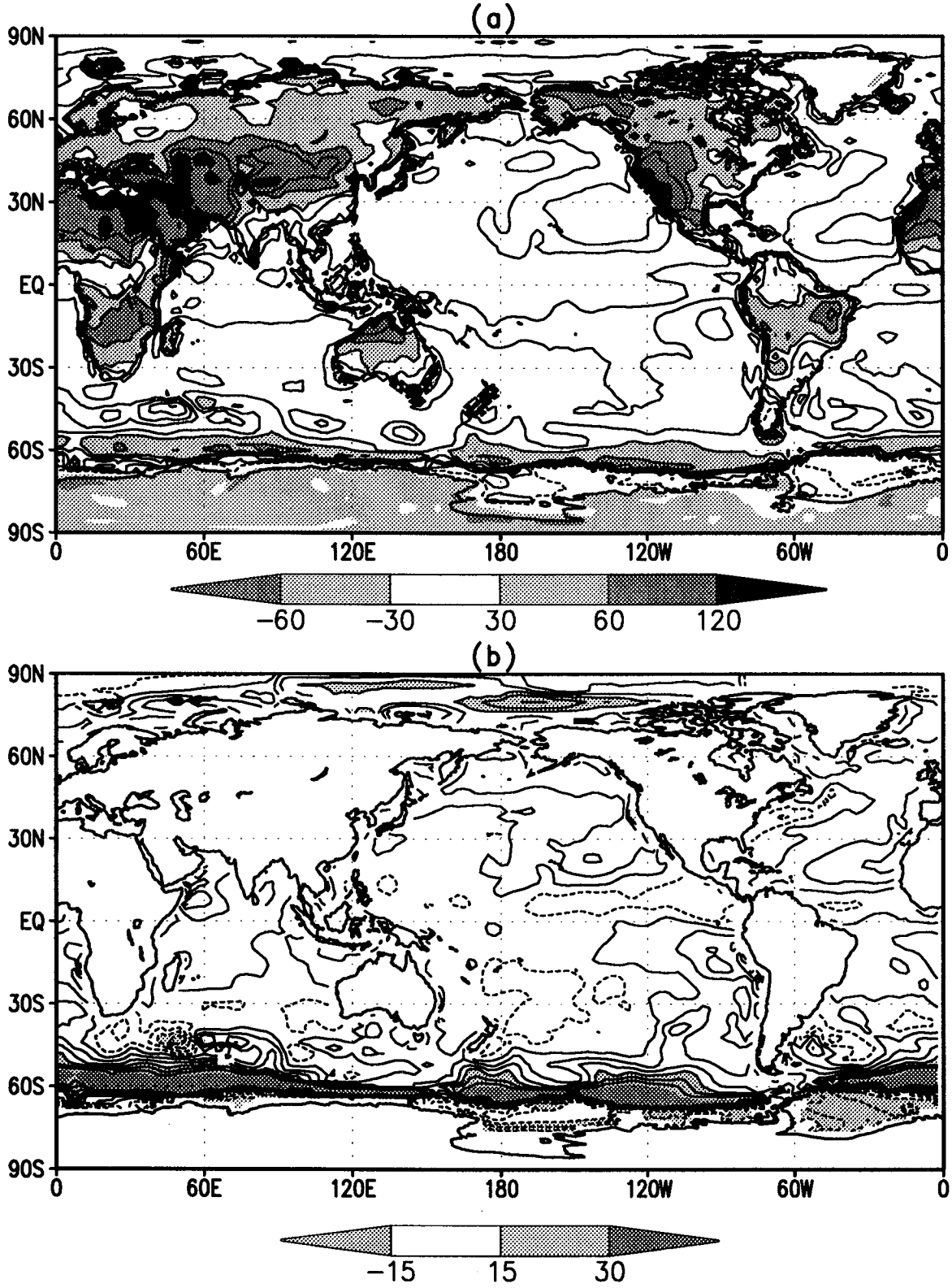


Figure 35: The sensible heat flux for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour intervals are -5 0 5 10 20 30 60 90 120 150 180 in (a) and every 5 in (b).

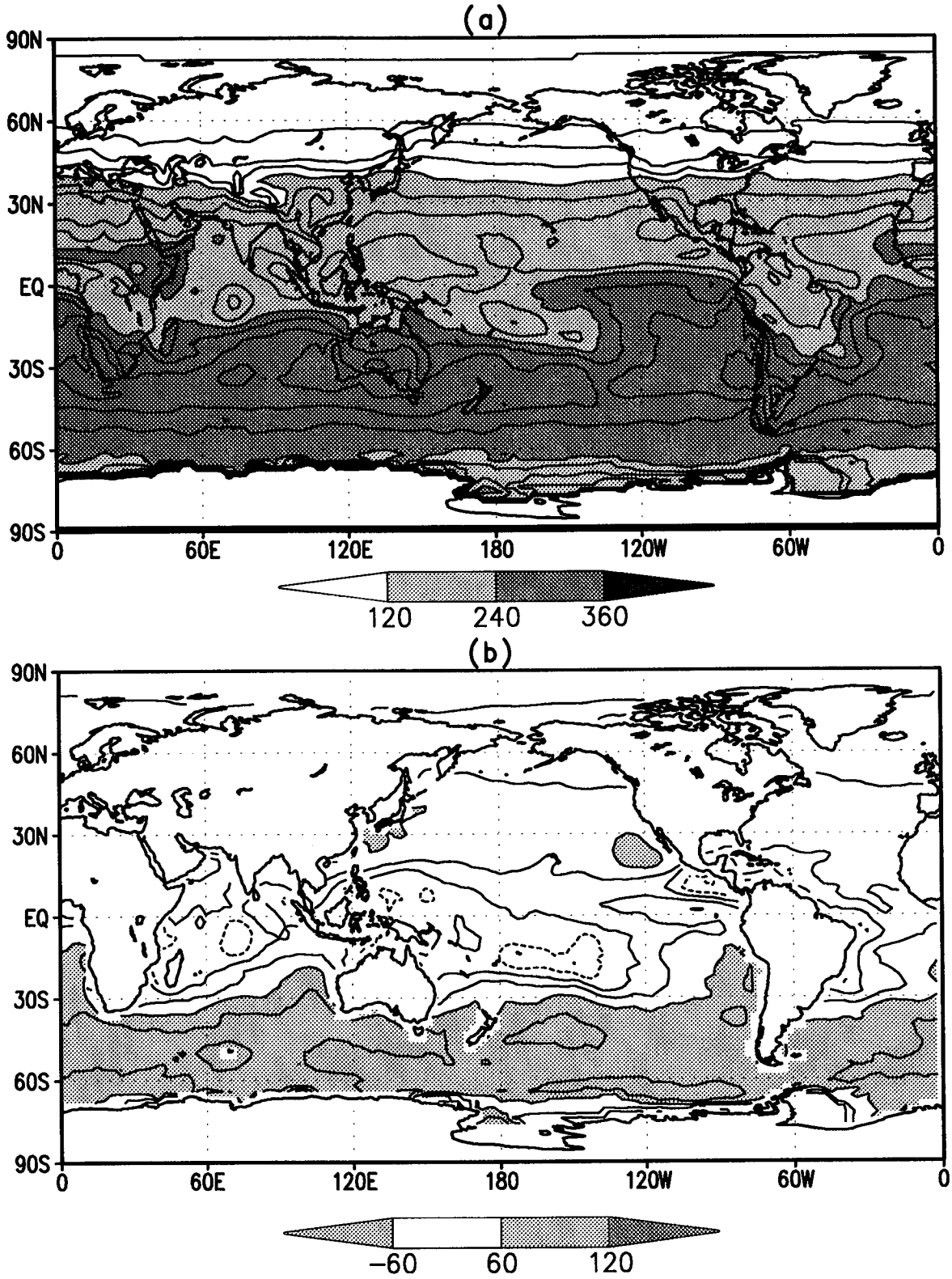


Figure 36: The net surface shortwave radiation for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Contour interval is 30 W/m^2.

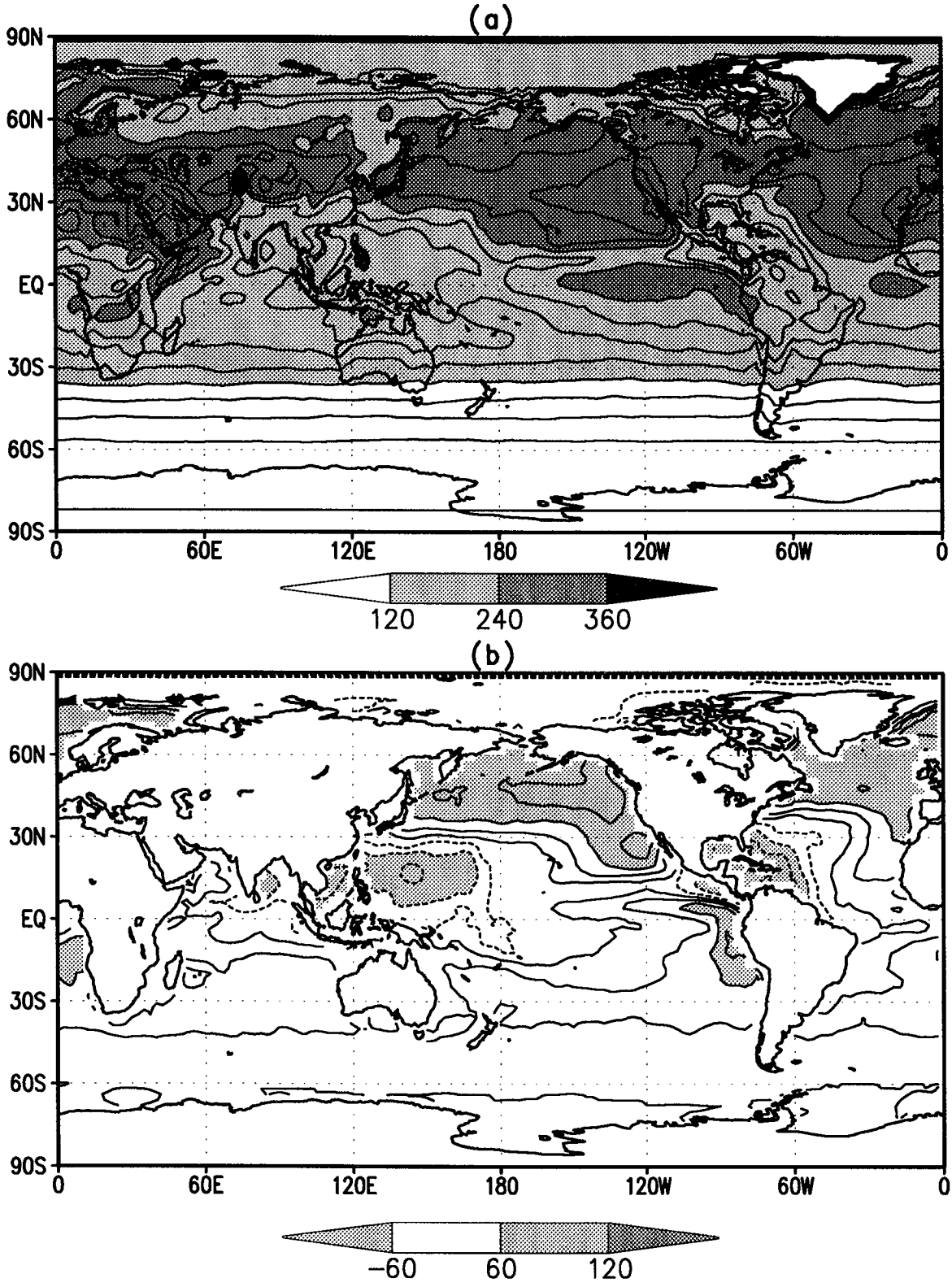


Figure 37: The net surface shortwave radiation for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Contour interval is 30 W/m^2.

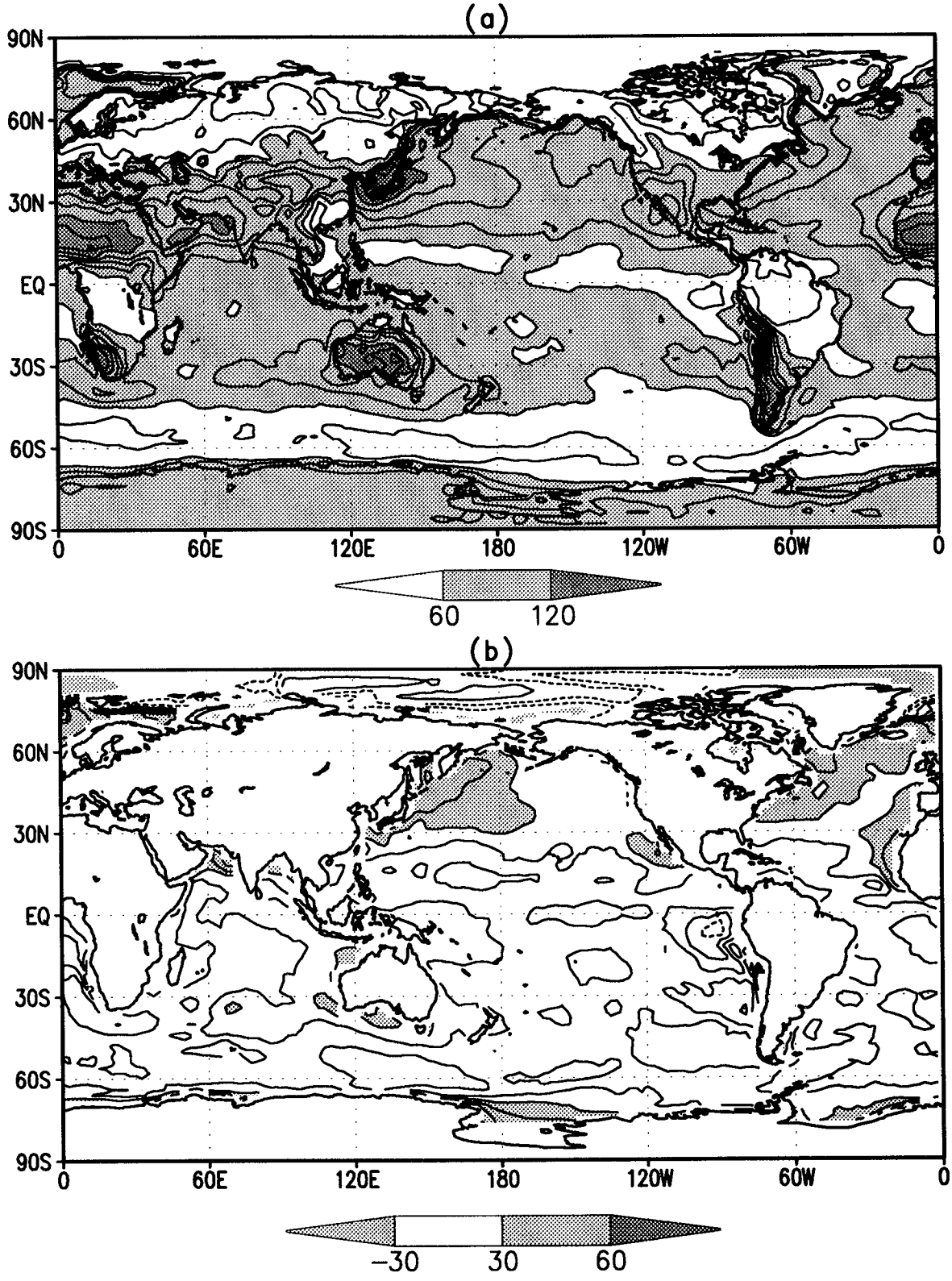


Figure 38: The net surface longwave radiation for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour interval is 15 in (a) and 10 in (b).

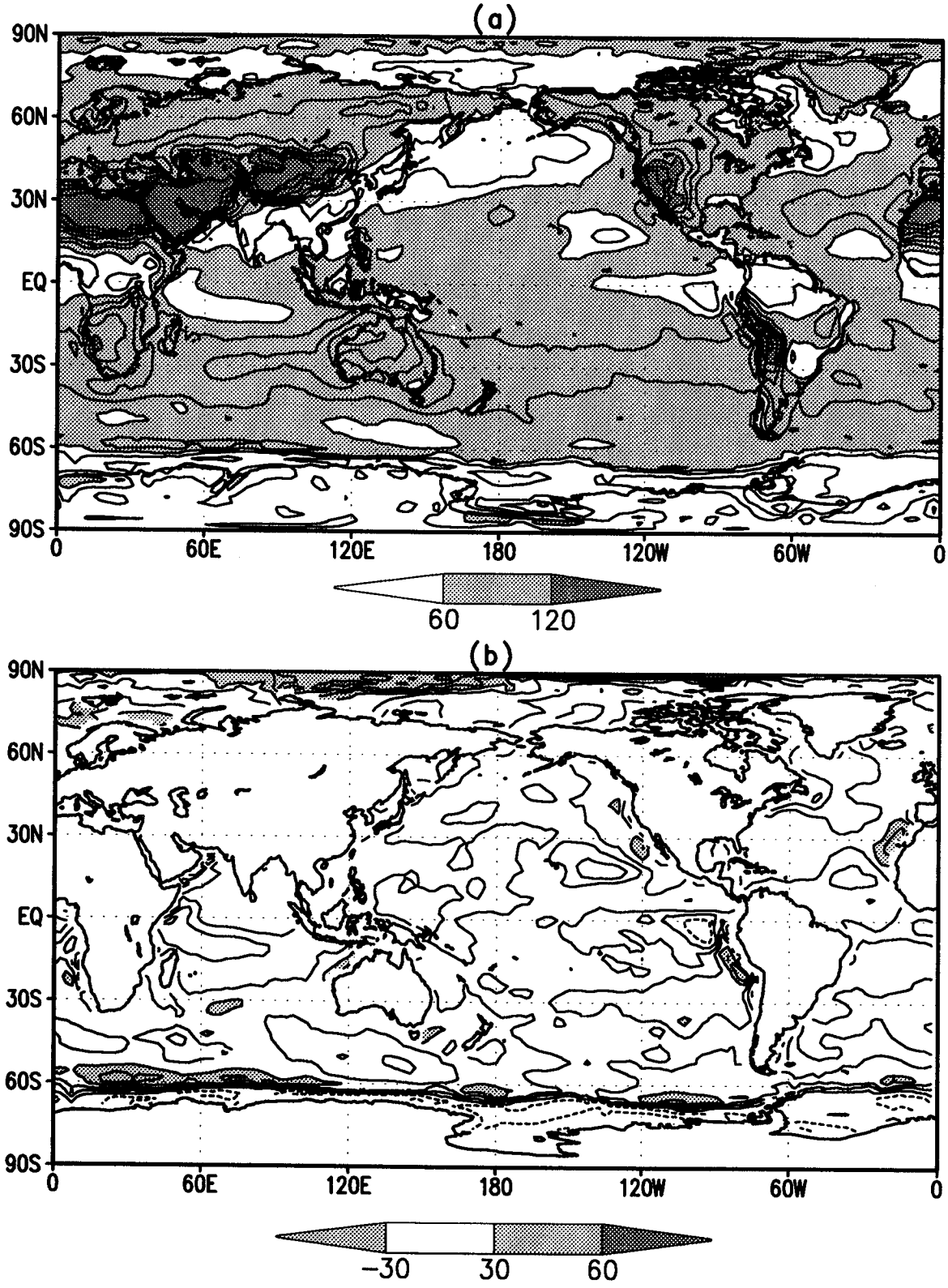


Figure 39: The net surface longwave radiation for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour interval is 15 in (a) and 10 in (b).

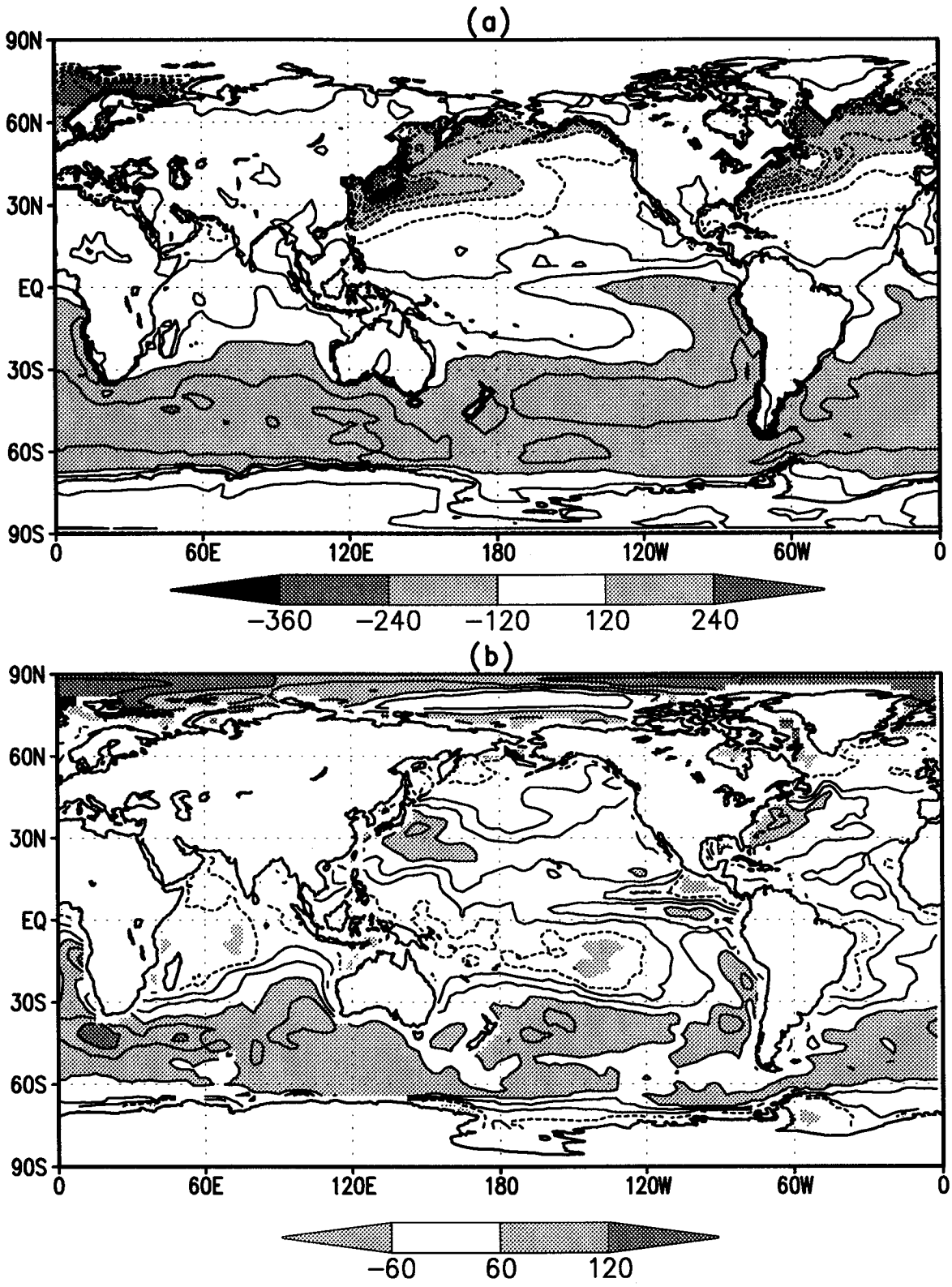


Figure 40: The net surface heat flux for DJF for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour interval is 60 in (a) and 30 in (b).

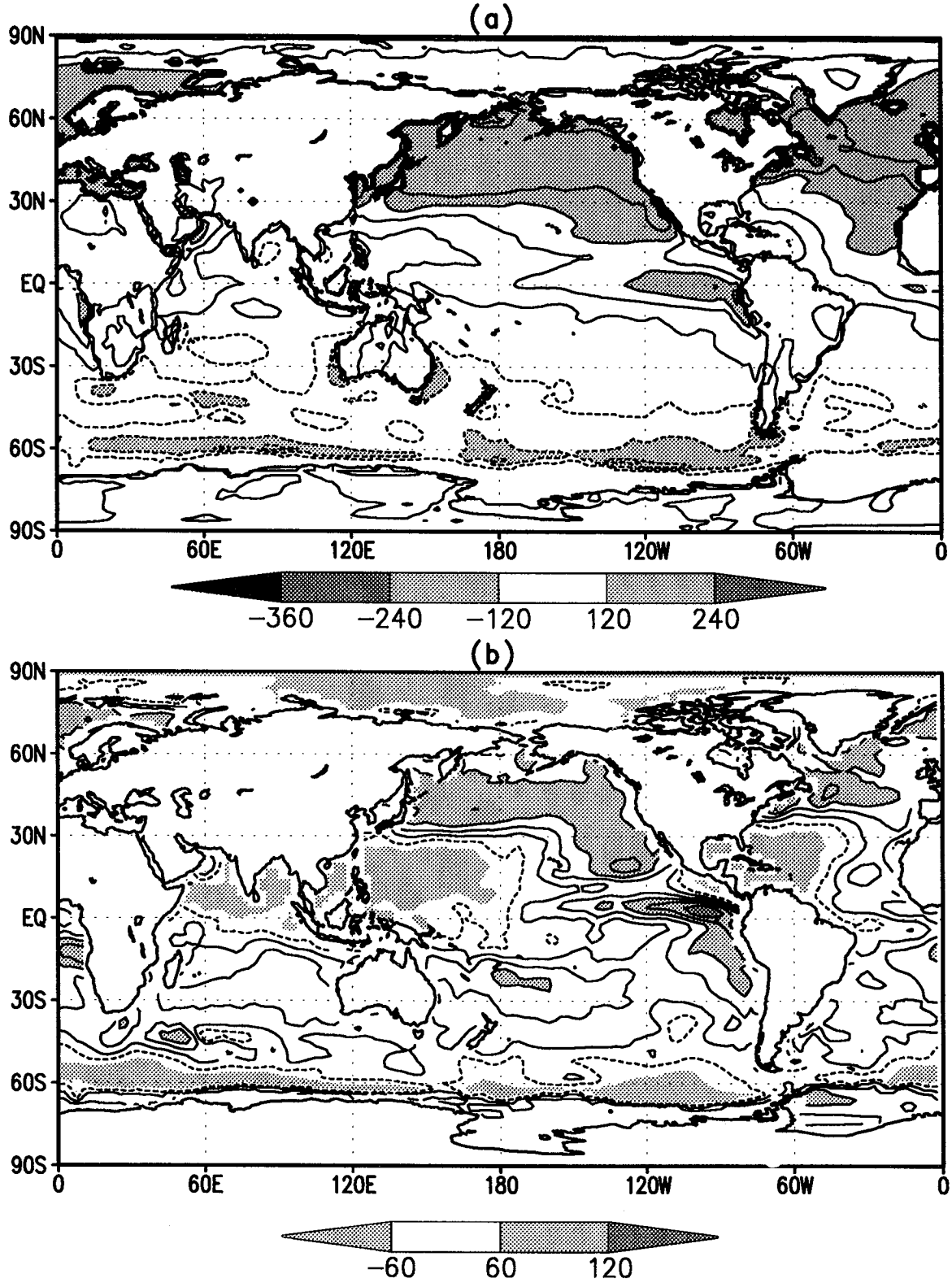


Figure 41: The net surface heat flux for JJA for a) GEOS–DAS, and b) GEOS–DAS minus COADS. Units:W/m^2. Contour interval is 60 in (a) and 30 in (b).

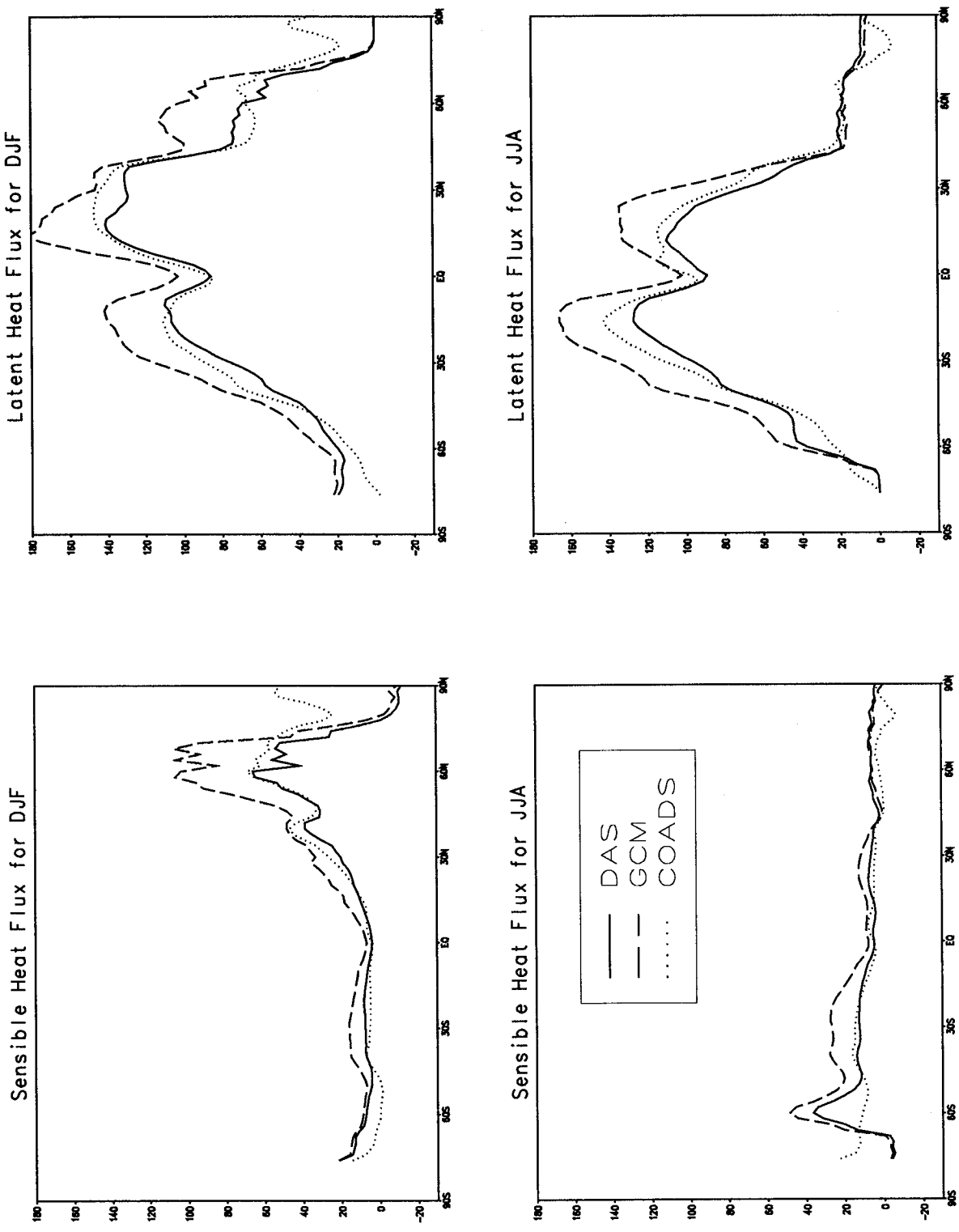


Figure 42: The zonal mean (oceans only) sensible and latent heat flux for the GEOS–DAS, GEOS–GCM, and COADS. Units:W/m^2.

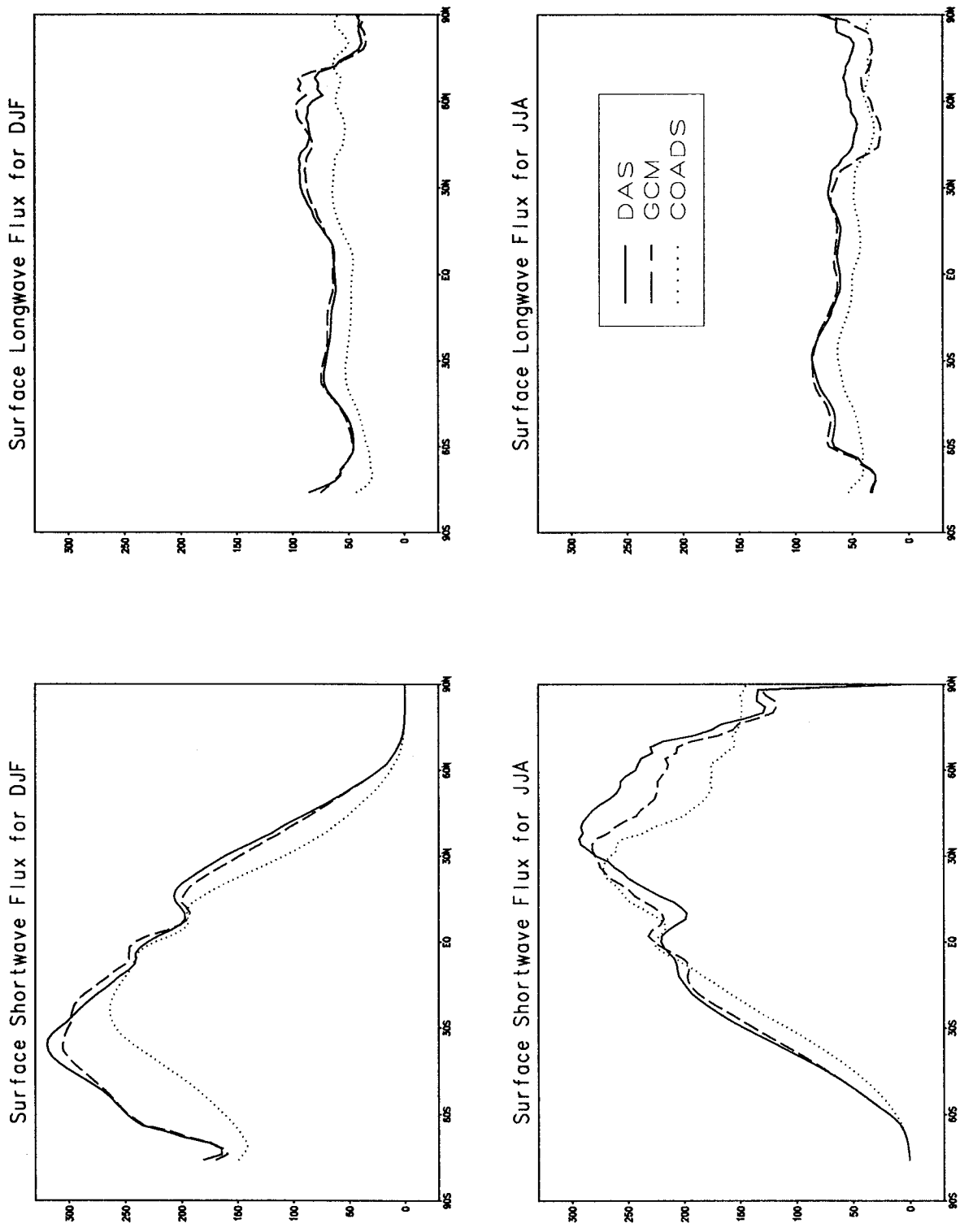


Figure 43: The zonal mean (oceans only) surface shortwave and longwave flux for the GEOS–DAS, GEOS–GCM, and COADS. Units:W/m^2.

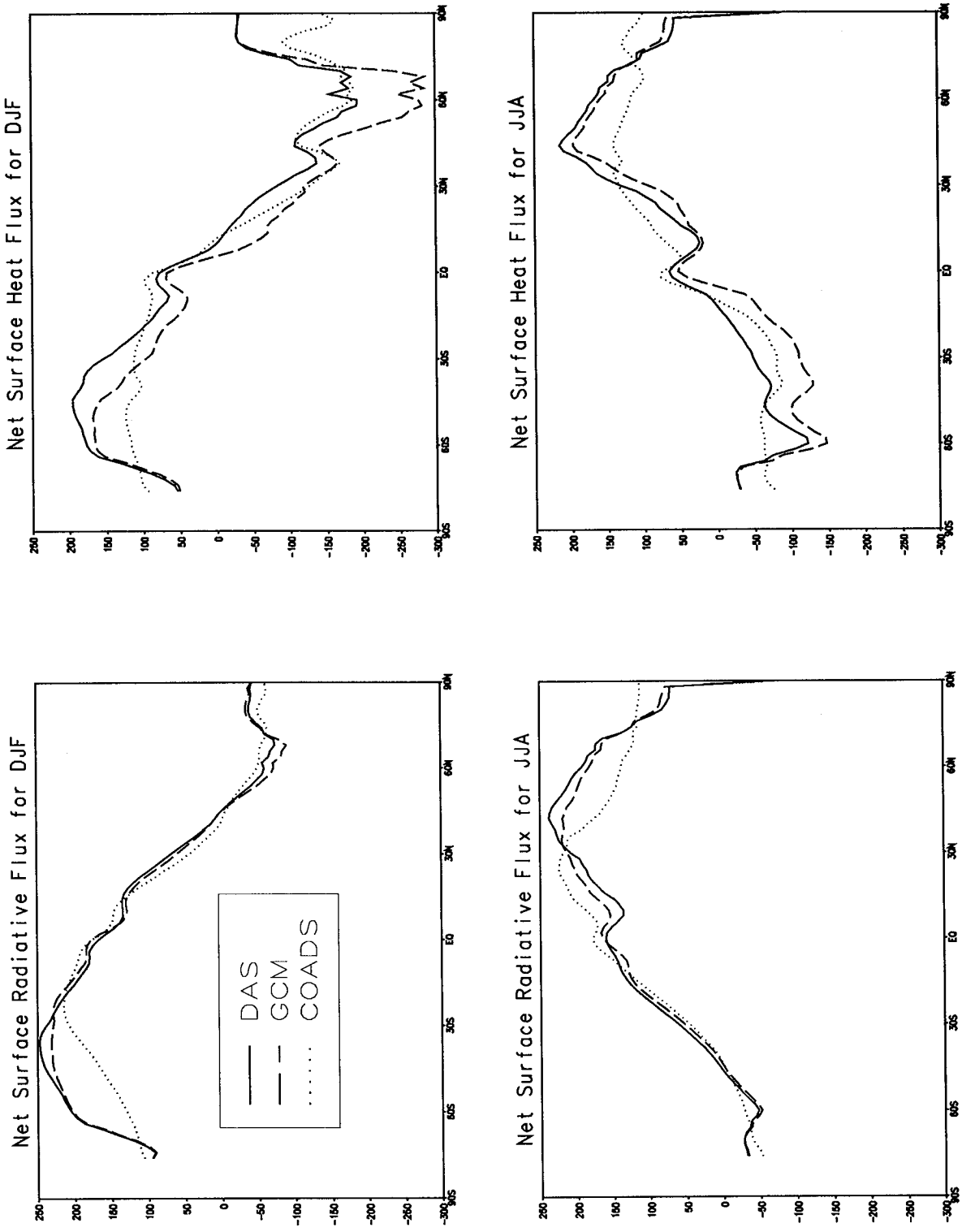


Figure 44: The zonal mean (oceans only) net surface radiative flux and net surface heat flux for the GEOS–DAS, GEOS–GCM, and COADS. Units:W/m^2.

4.2.4 Precipitation and the U.S. Moisture Budget

Global Precipitation

The global precipitation fields and the corresponding difference fields (GEOS–station observations) are shown in Figs. 45 and 46, respectively. The precipitation patterns show a clear seasonal shift, with the heaviest tropical precipitation moving northward with the sun. The heavy summer monsoon rainfall over India is also evident. Other features of note are the precipitation associated with the boreal winter middle latitude storm tracks and the boreal summer increase in rainfall over land. The largest differences with station observations occur over the tropics in the regions of heavy rainfall. The JJA GEOS precipitation appears to be too heavy over Southeast Asia, northern South America, and much of the NH middle and high latitude land masses. The DJF GEOS precipitation appears to be somewhat too weak over parts of the United States and Europe. The differences in the tropics (Africa and South America) suggest the ITCZ in the GEOS–DAS may be shifted too far north during DJF.

The seasonal cycle of precipitation from the GEOS–DAS and station observations averaged over selected regions is given in Figs. 48–49. The regions consist of the land areas encompassed by the rectangles in Fig. 47. The seasonal cycle is computed from monthly means averaged over the five years (Mar 1985–Feb 1990).

The station observations show that the general behavior of the seasonal cycle varies substantially between the eight regions. For example, there is a pronounced annual cycle over India associated with the monsoons, while there is very little seasonal variation in the area–averaged precipitation over the Aleutian region. The general behavior in each of the regions is reproduced in the GEOS–DAS; however, the GEOS values are generally larger than the station observations, particularly during boreal summer. There is also a tendency for GEOS to underestimate the winter time precipitation over North America (including the Aleution region) and Europe. A more detailed look at the precipitation over the continental United States is provided in the next section. The deviations from the seasonal cycle are shown in section 4.4.3.

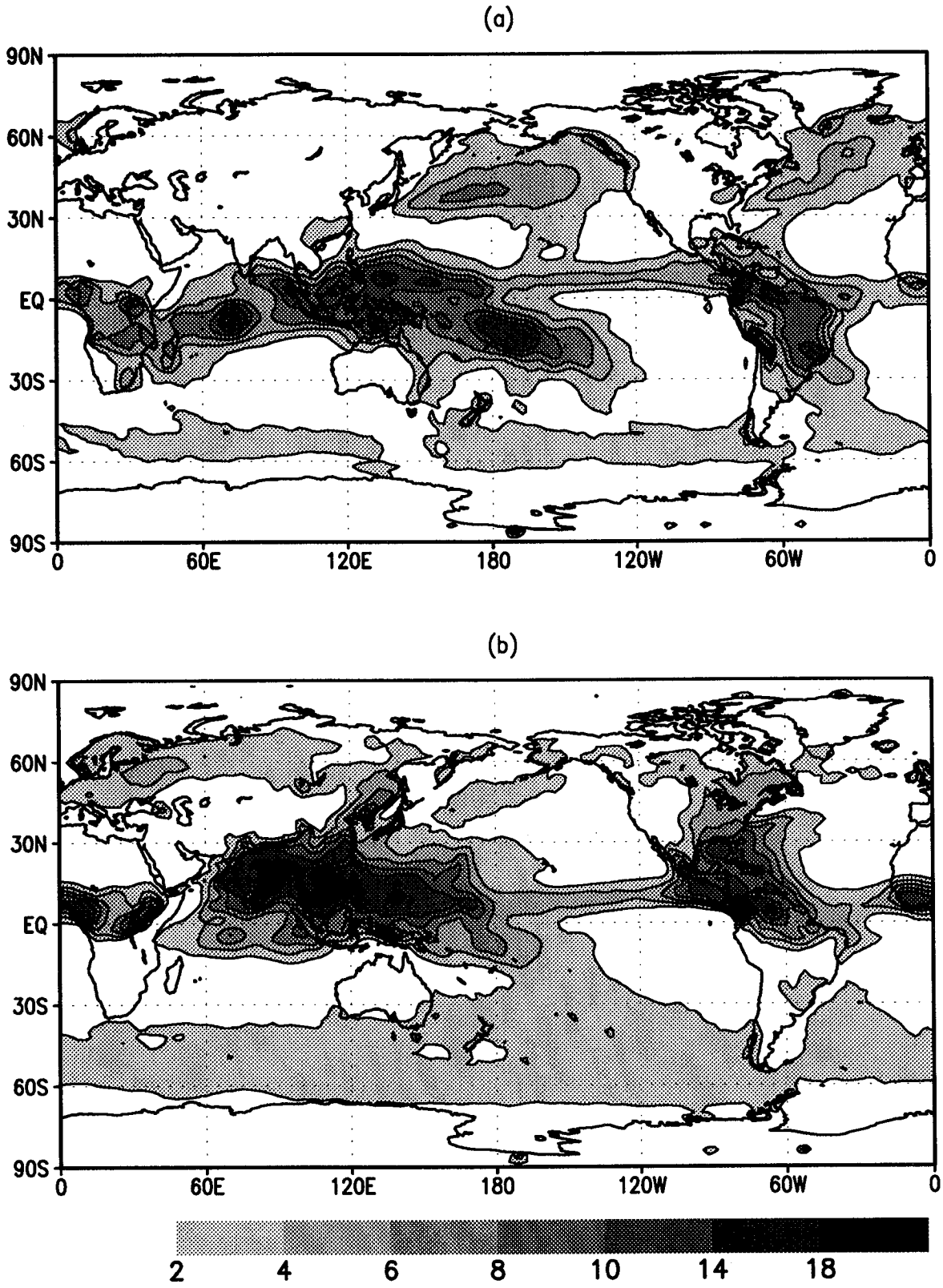


Figure 45: The GEOS–DAS precipitation climatologies for (a) DJF and (b) JJA. Units: mm/day.

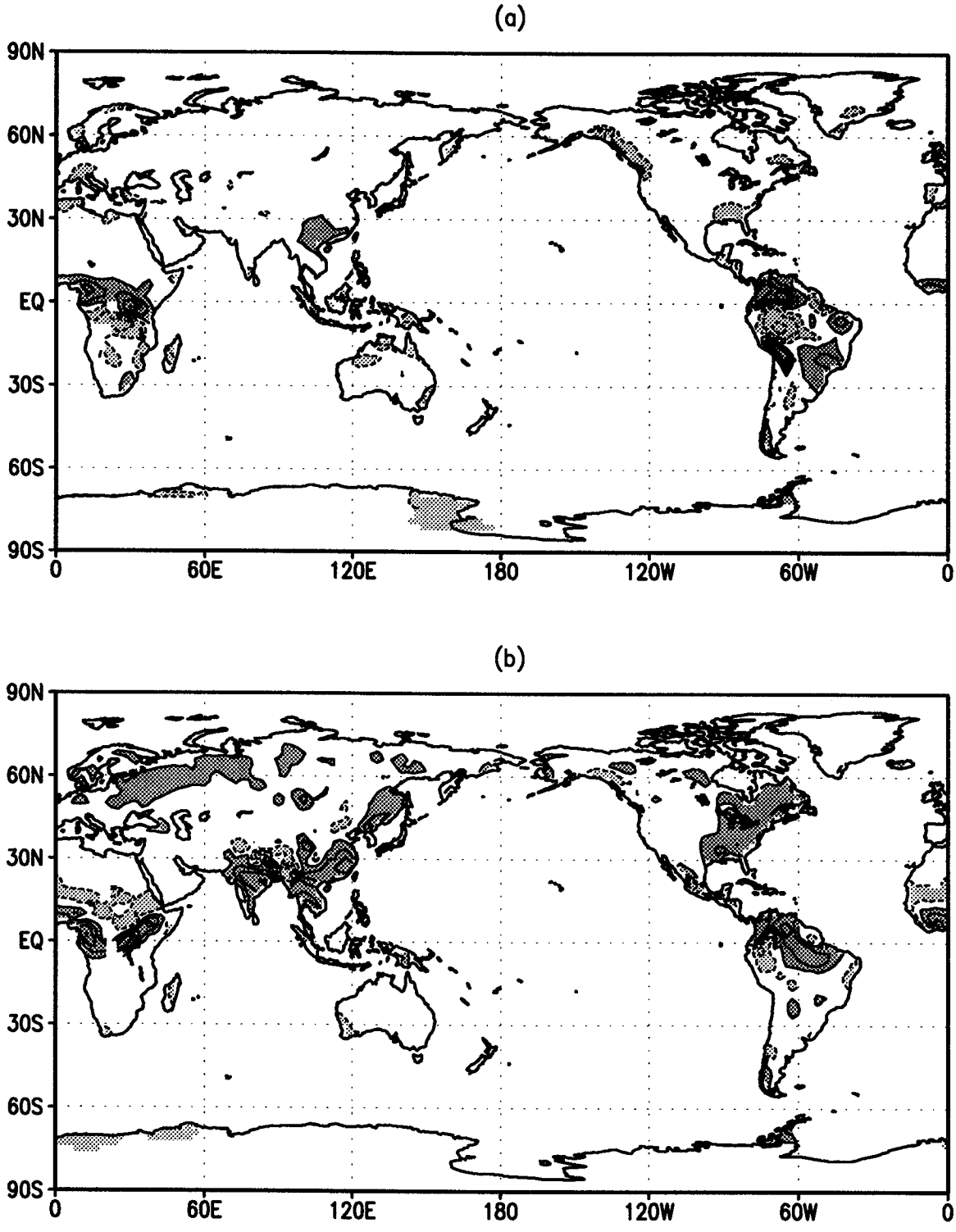


Figure 46: The climatological mean difference (GEOS–DAS minus station observations) of precipitation for (a) DJF and (b) JJA. The contour interval is 2 mm/day. Negative values are dashed. The values greater than 1 mm/day and less than -1 mm/day are shaded.

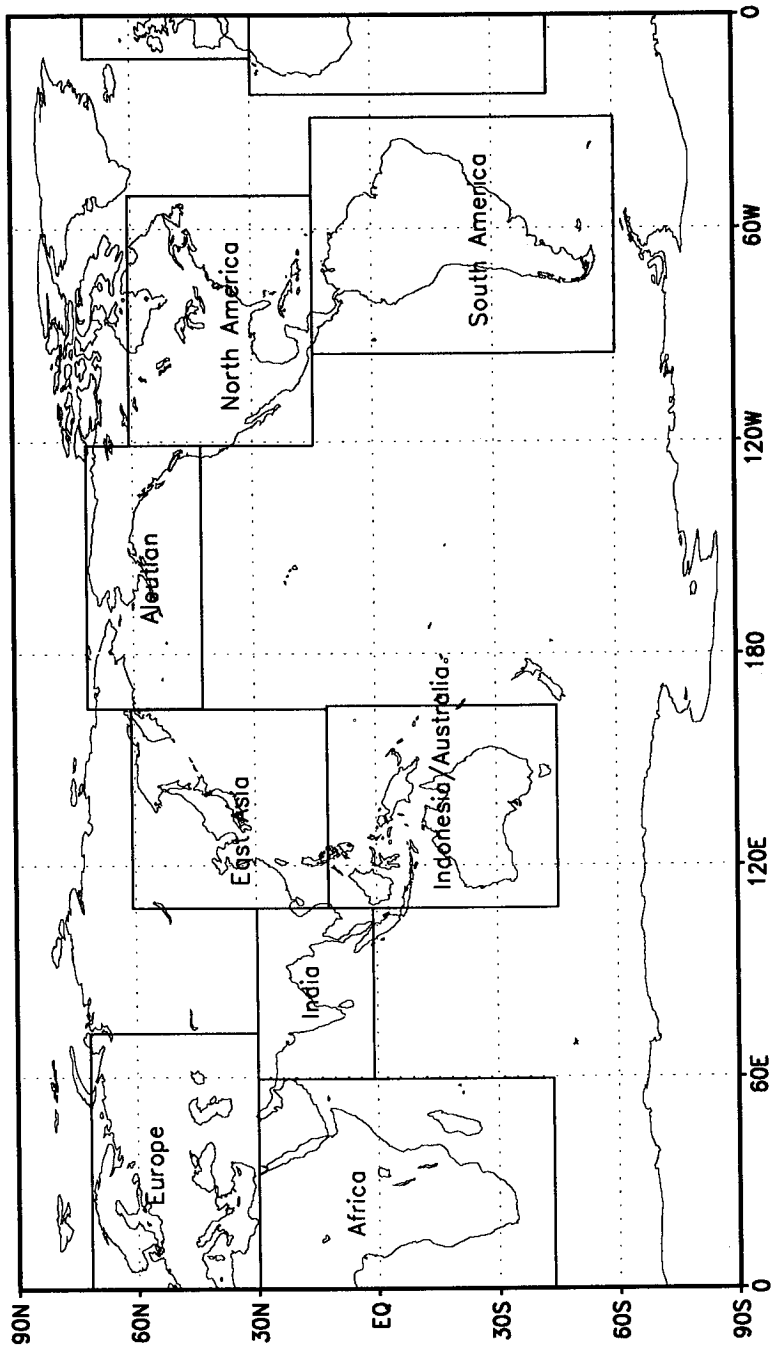


Figure 47: The boxes contain the land areas over which the averages are done for the precipitation line plots.

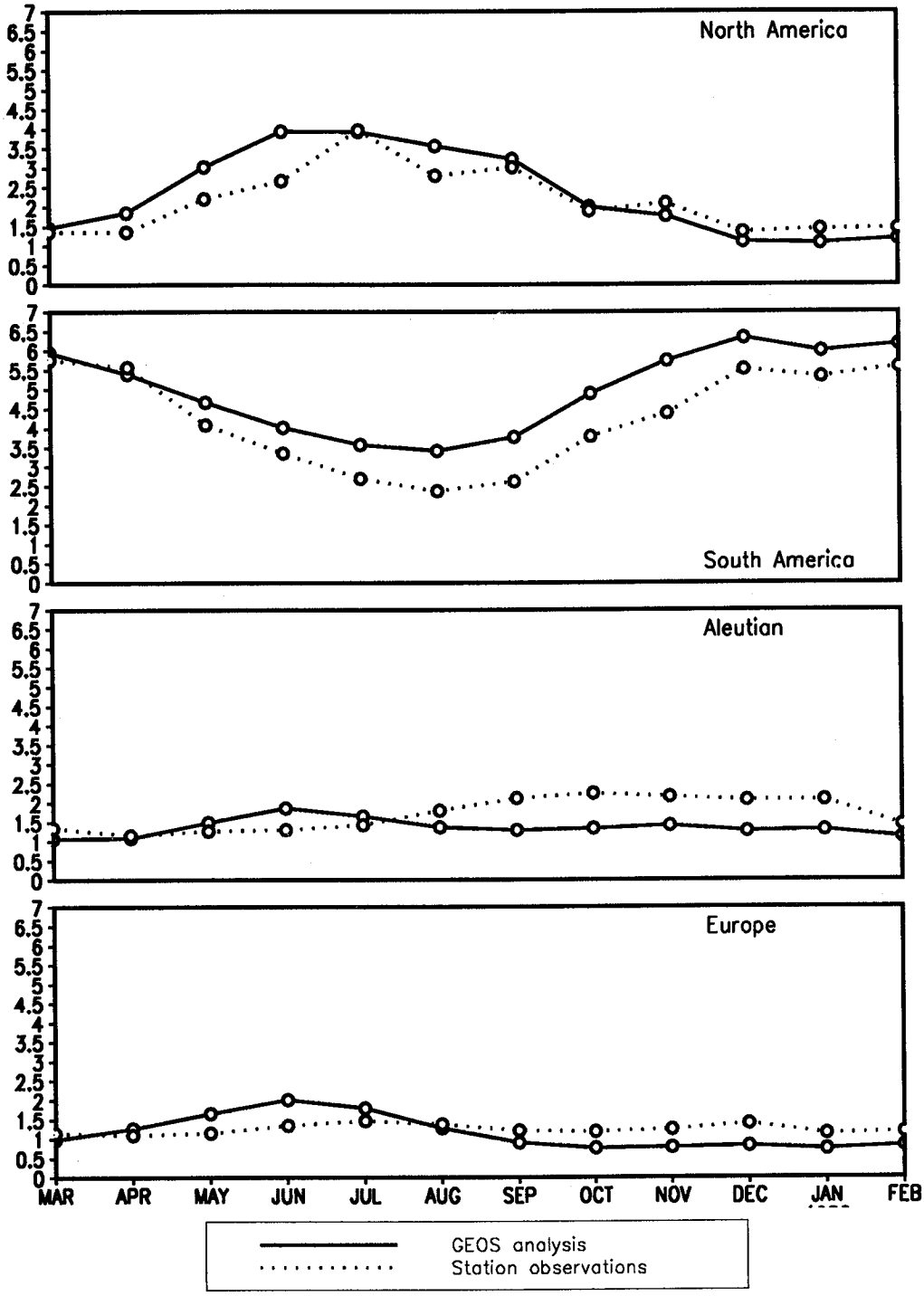


Figure 48: The seasonal cycle of the regional mean precipitation over land from GEOS and the station data. Units: mm/day.

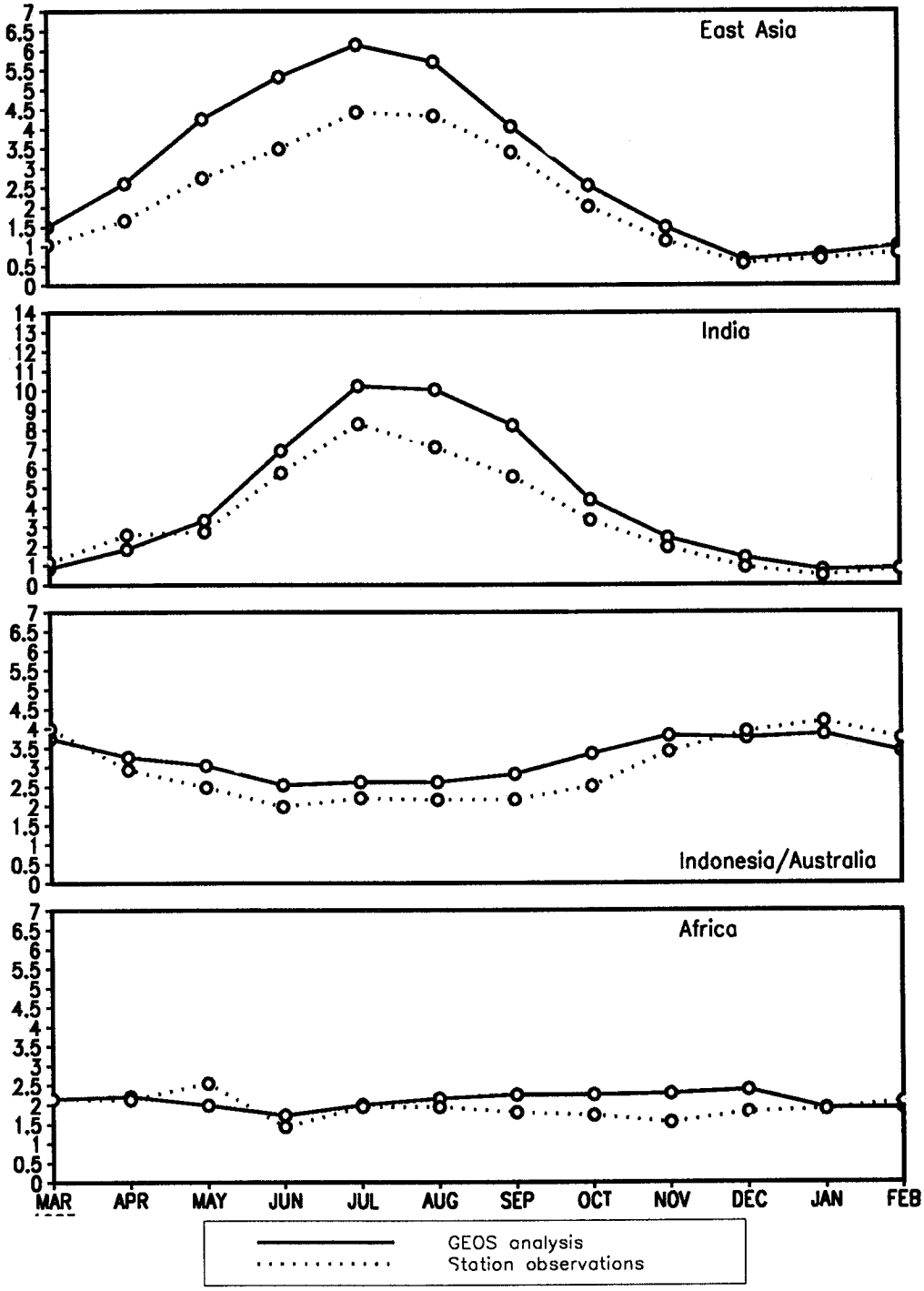


Figure 49: The seasonal cycle of the regional mean precipitation over land from GEOS and the station data. Units: *mm/day*.

U.S. Moisture Budget

The vertically integrated moisture budget in the GEOS–DAS may be written as

$$\langle \frac{\partial q\pi}{\partial t} \rangle = -\langle \nabla \cdot q\pi \vec{V} \rangle - P + E + \langle \Delta q\pi \rangle + \langle F_q \rangle, \tag{6}$$

where π is the surface pressure minus the constant prescribed pressure at the top of the model atmosphere (10 mb), P is the precipitation, E the evaporation, and $\Delta q\pi$ the mass–weighted analysis increment. The last term includes the effects of the Shapiro filter and any filling of moisture to remove negative specific humidities (see section 3.2). The angle brackets denote a vertical integral. The filter and filling terms were not saved in the assimilation; however, off–line tests showed that the residual is dominated by the Shapiro filter. In the results shown here, the filter term is estimated as a residual of the other terms. As shown below, the filter contributes significantly only to the local budget: in the area average this term is very small. The moisture increments will contribute to the time mean budgets if the first guess is biased.

Figures 50–58 show the components of the vertically–integrated moisture budget over the continental United States from the GEOS–DAS averaged over two month intervals. Comparisons are made with the observed precipitation from the U.S. climate division data set described in section 4.1.

The GEOS–DAS substantially overestimates the summertime precipitation over the eastern United States, while it tends to underestimate the winter precipitation over the southeast. The precipitation is closest to the observations during the transition seasons. Both the analysis increments and the filter contribute significantly to the local budgets throughout the year. The analysis increments are generally negative (water is being removed, suggesting a wet bias in the GCM first guess) especially over the western United States, except during the summer when much of the eastern part of the continent has postive analysis increments (water is being added, suggesting the first guess is too dry). Some caution must be used in interpreting the analysis increments, especially during the summer months when spurious feedbacks may occur with the convection (see Molod et al. 1995).

The area averages of each of the terms in the vertically–integrated moisture budget are shown in Fig. 56. The area averages are computed over all the continental land points between latitudes 30°N and 50°N. The overestimate of the summer precipitation and the underestimate of the winter precipitation is clearly evident. The analysis increments—unlike the filter, which does not contribute to the area averaged budget—make a substantial contribution. The mean contribution of the analysis increment is comparable in magnitude to the convergence term and tends to act in the opposite sense. The apparent connection between the analysis increments and the convergence is perplexing since a diagnostic study of the analysis increments suggests that (during late spring) the first guess bias is primarily

associated with errors in the precipitation and evaporation fields (Schubert and Chang, 1995).

The moisture fluxes entering and leaving the region 30°N– 50°N and 80°W–120°W are shown in Figs. 57–58. Note this region includes some ocean points not included in Fig. 56. The results show that most of the moisture enters the continent from the west and south and leaves across the east coast below 500 mb. During the spring and summer months the Great Plains low level jet (LLJ) contributes substantially to the moisture transport, accounting for about 30% of the moisture entering the continent (see also Helfand and Schubert 1995).

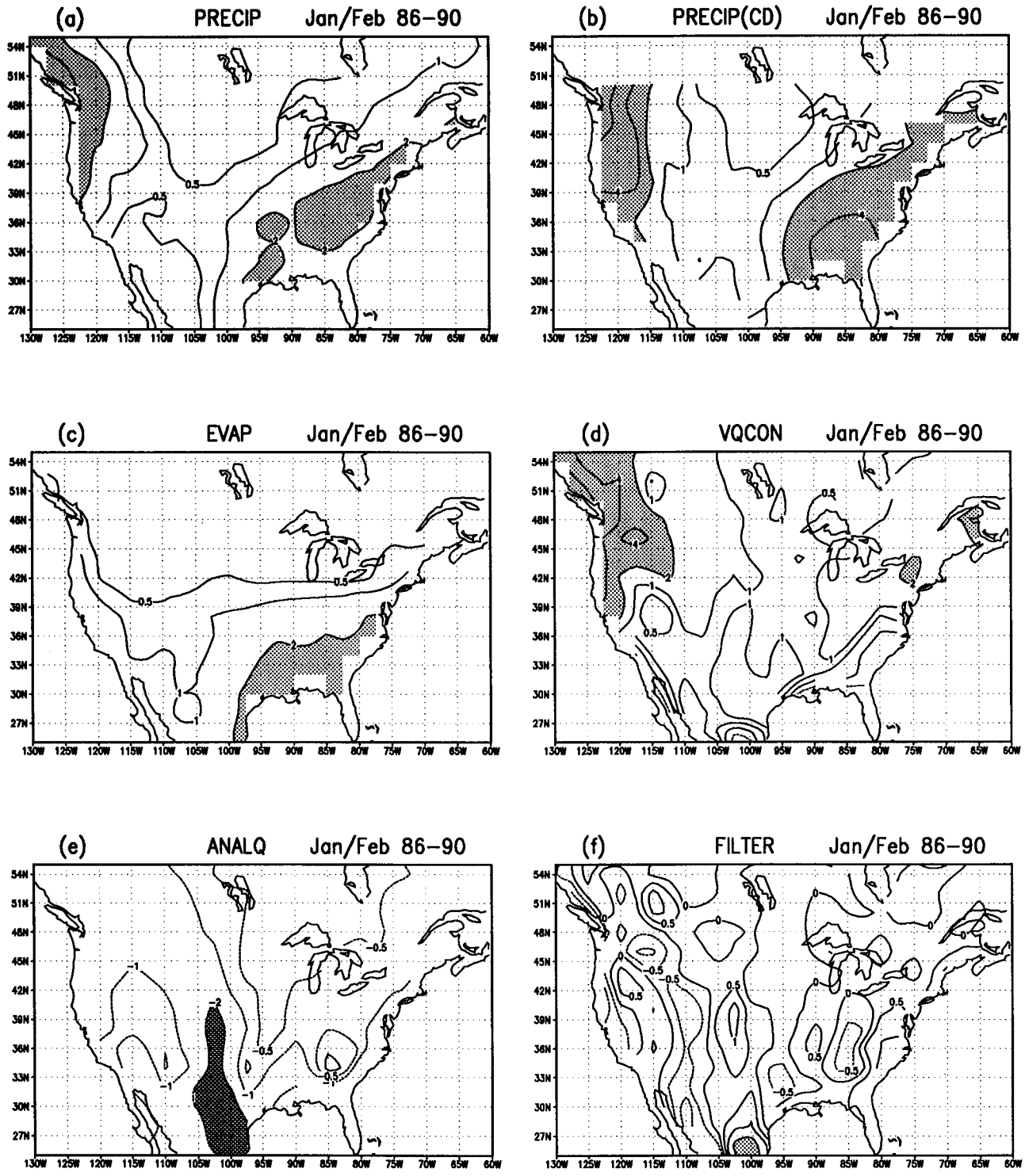


Figure 50: Components of the vertically–integrated moisture budget from the GEOS–1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a–c–d–e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

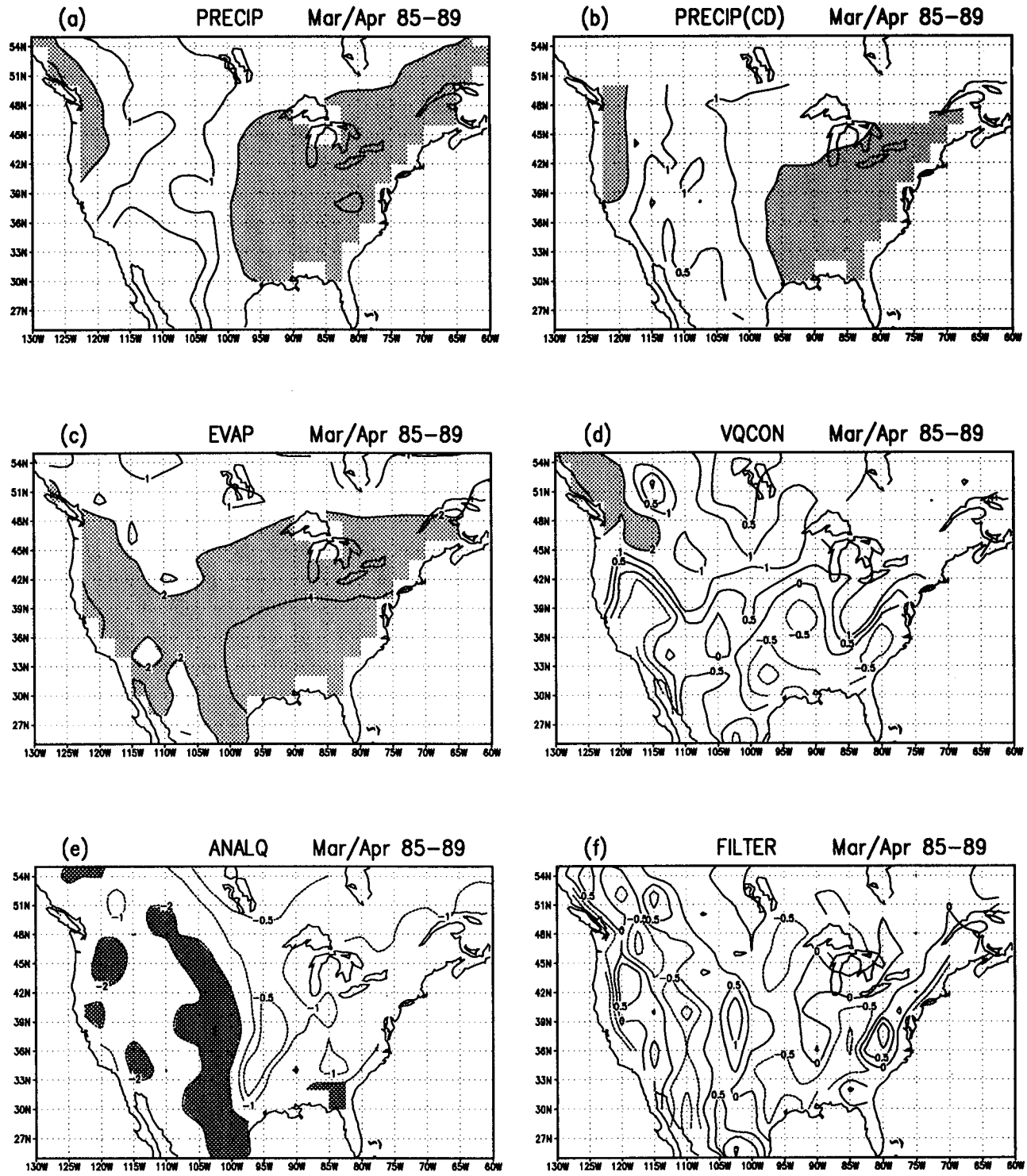


Figure 51: Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a–c–d–e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded.

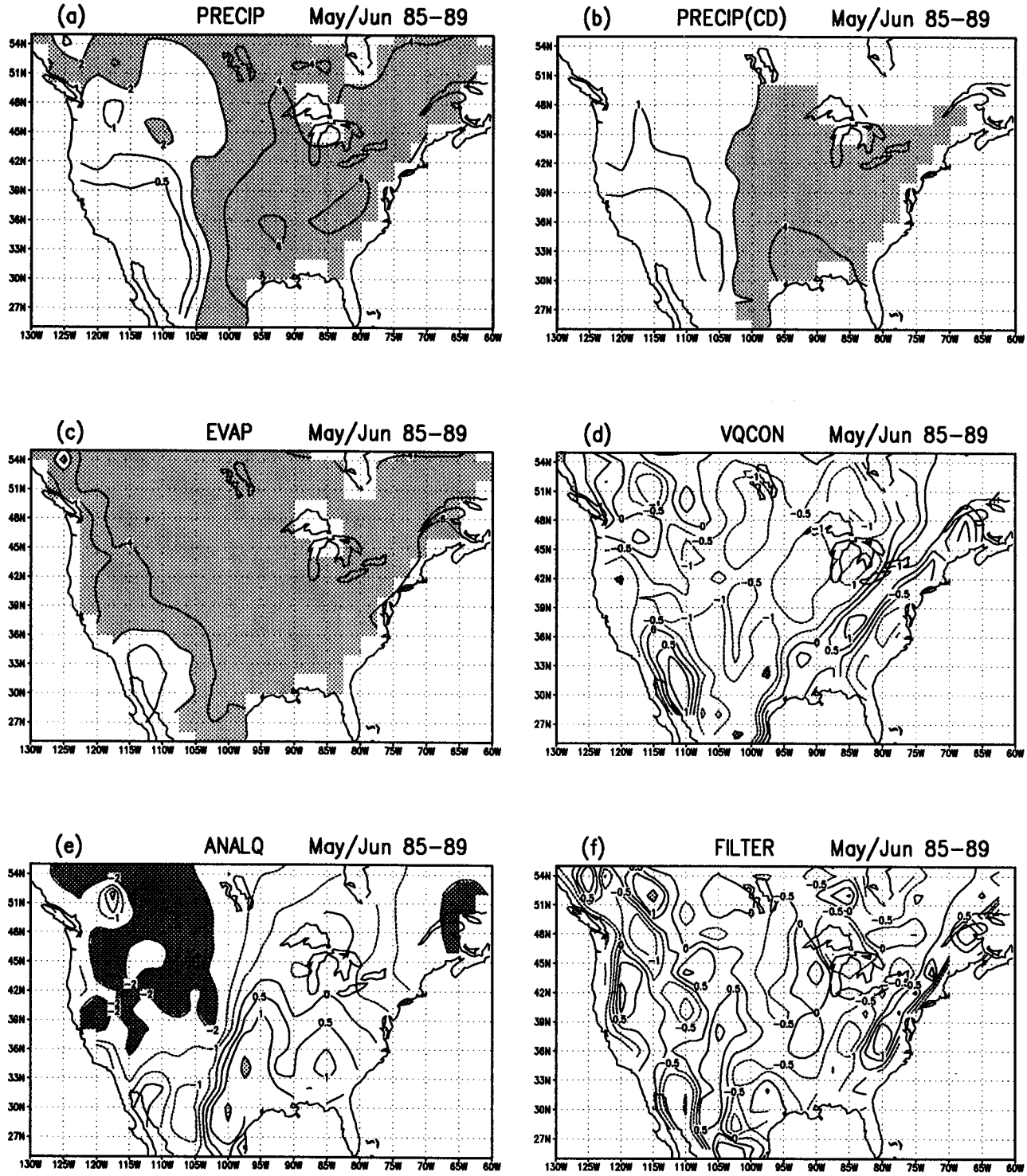


Figure 52: Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a–c–d–e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded.

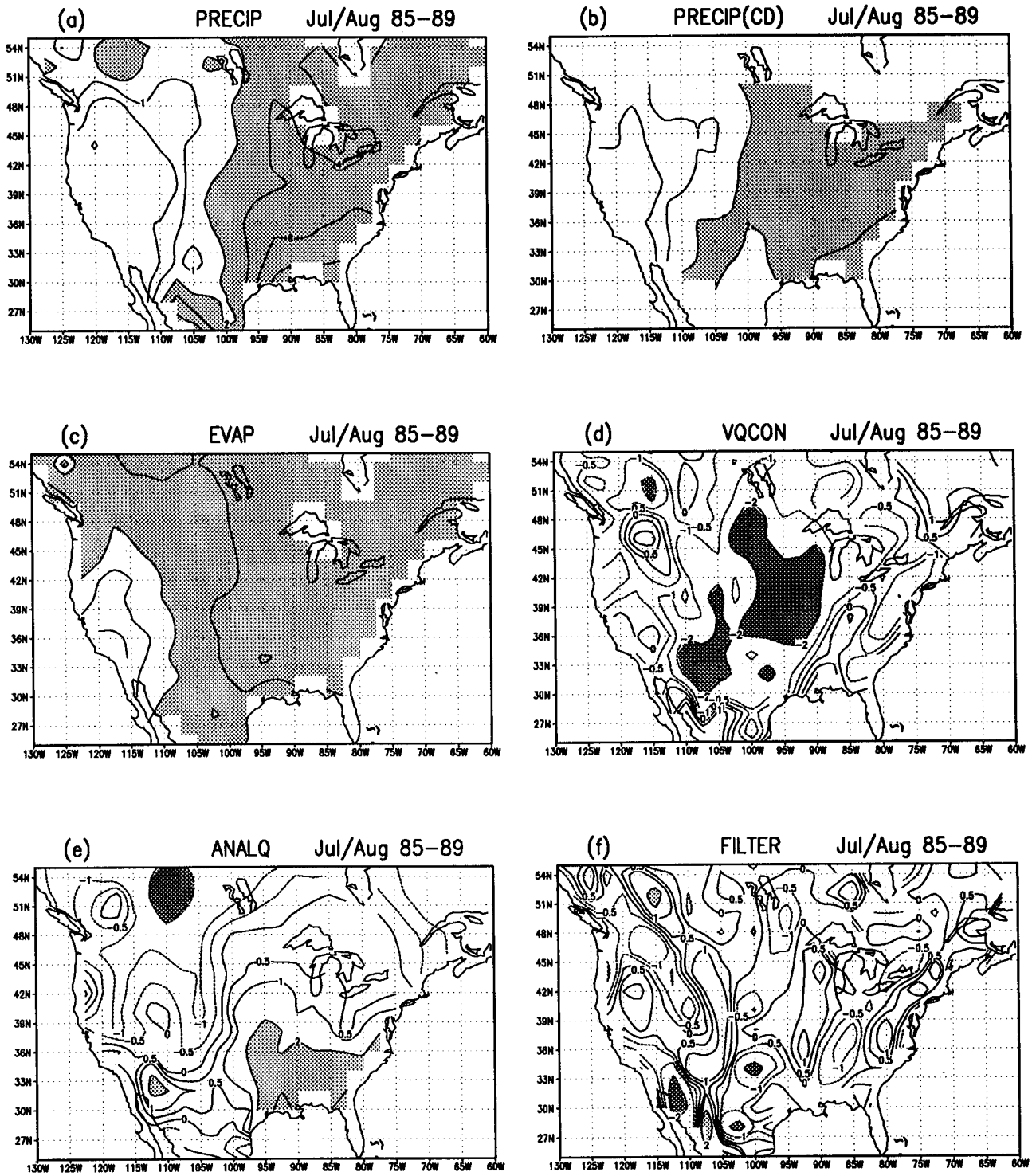


Figure 53: Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a–c–d–e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded.

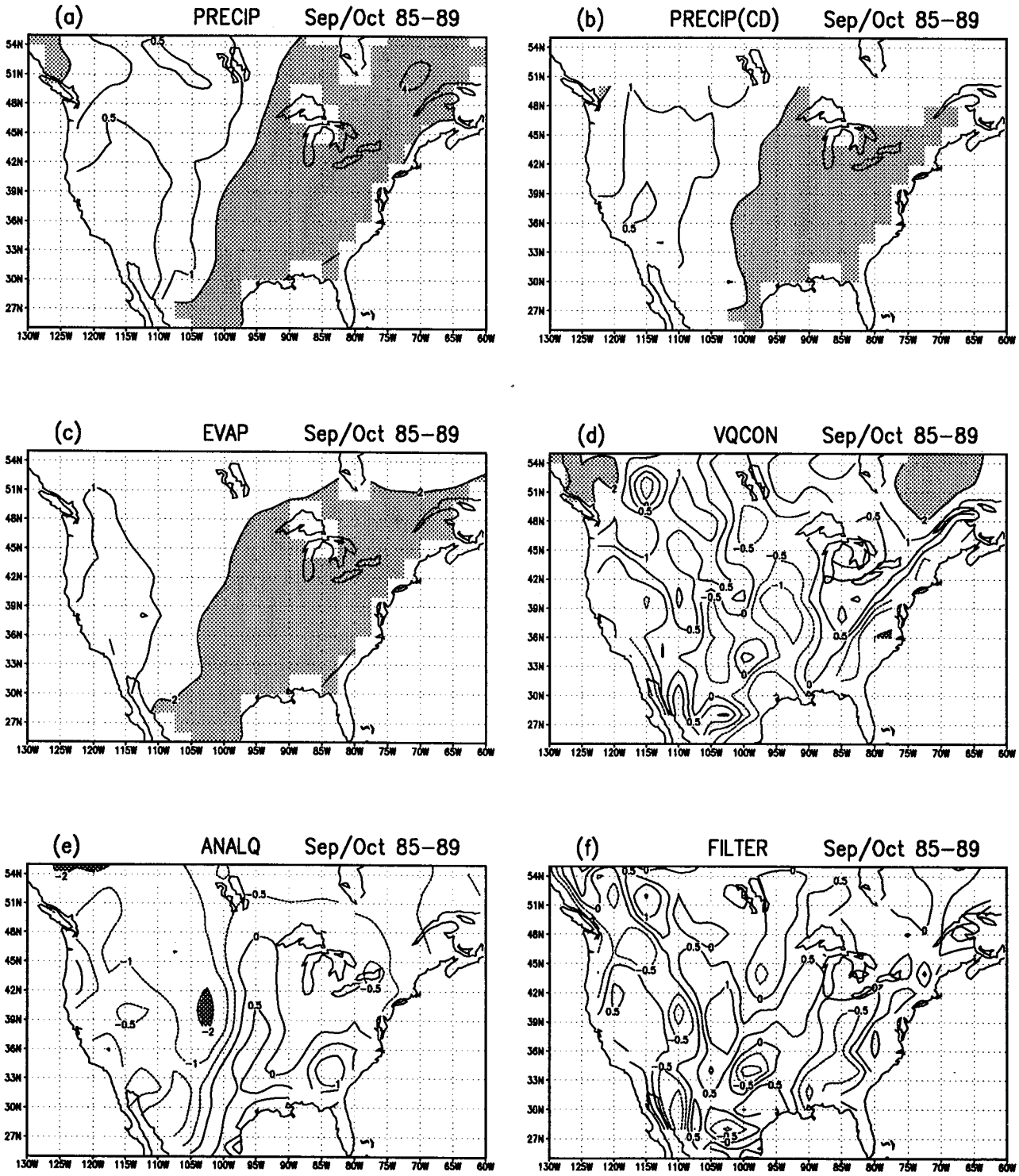


Figure 54: Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a–c–d–e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded.

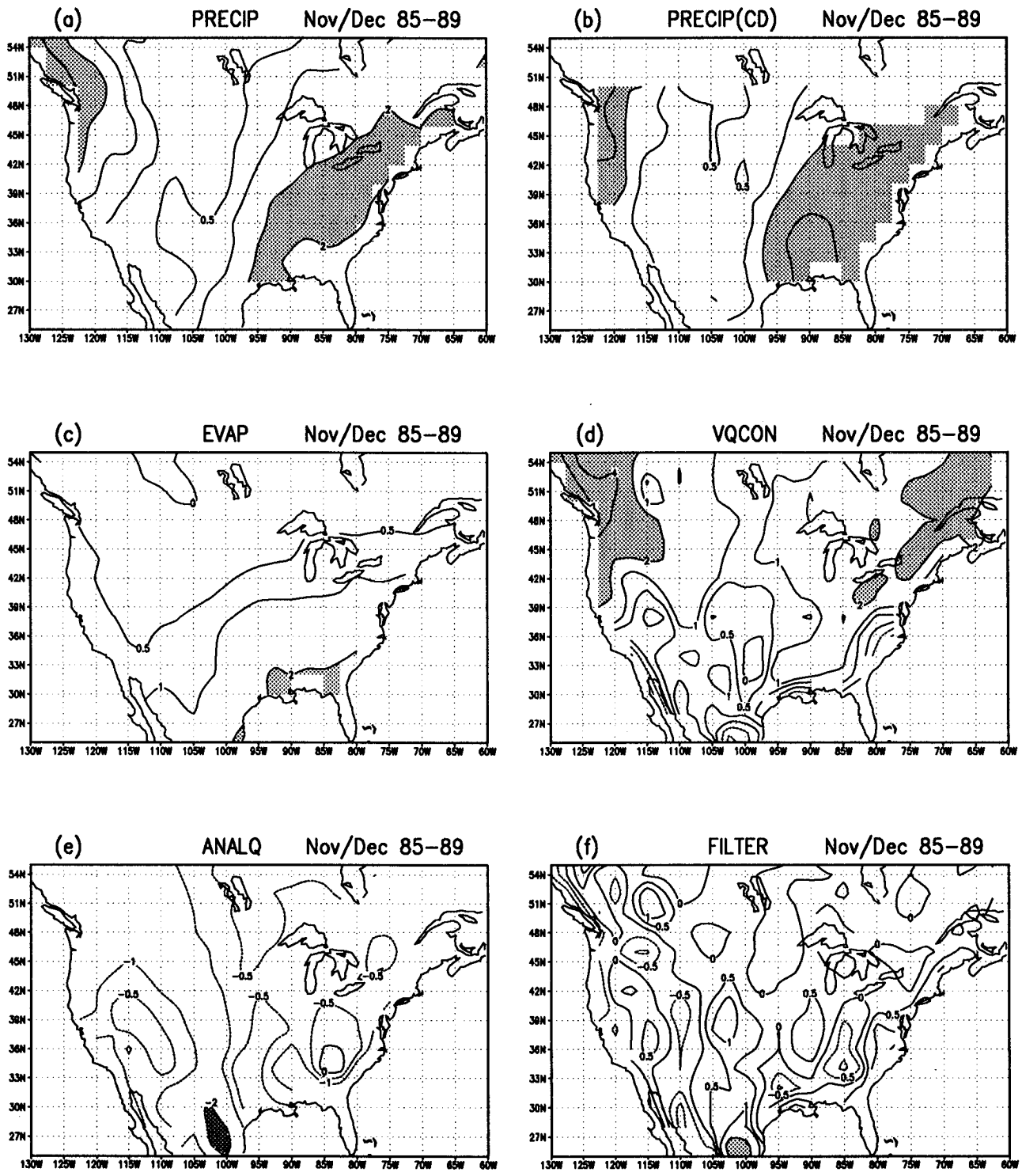


Figure 55: Components of the vertically-integrated moisture budget from the GEOS-1 DAS. a) precipitation, b) precipitation verification from climate division observations, c) evaporation, d) moisture convergence, e) the analysis increments,and f) the Shapiro filter (a–c–d–e). Contour intervals are −10 −8 −6, −4, −2, −1, −0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 *mm/day*. Absolute values greater than 2 are shaded.

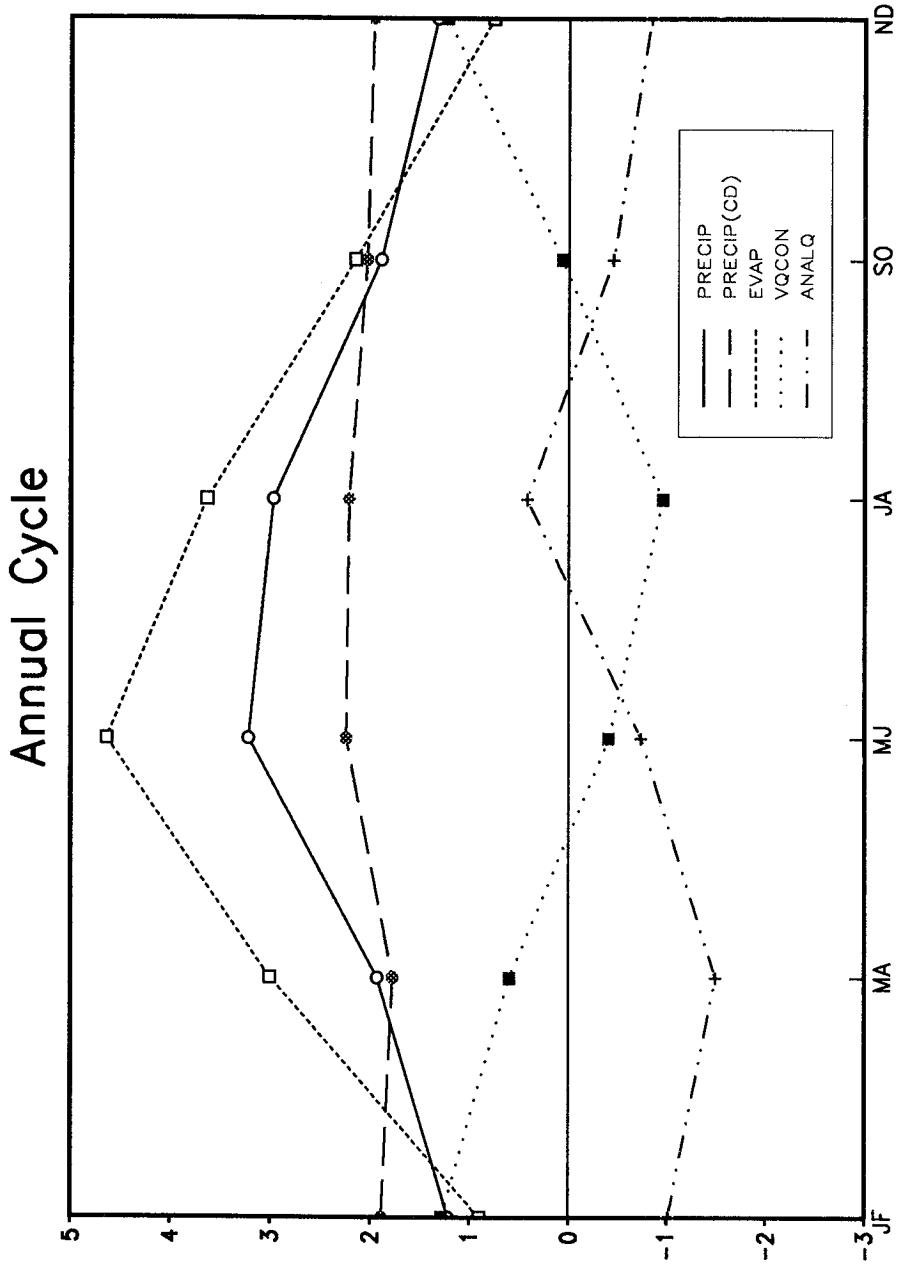


Figure 56: Annual cycle of vertically-integrated moisture budget terms averaged over the continental United States (30°N–50°N). The CD refers to the climate division precipitation. Units are mm/day.

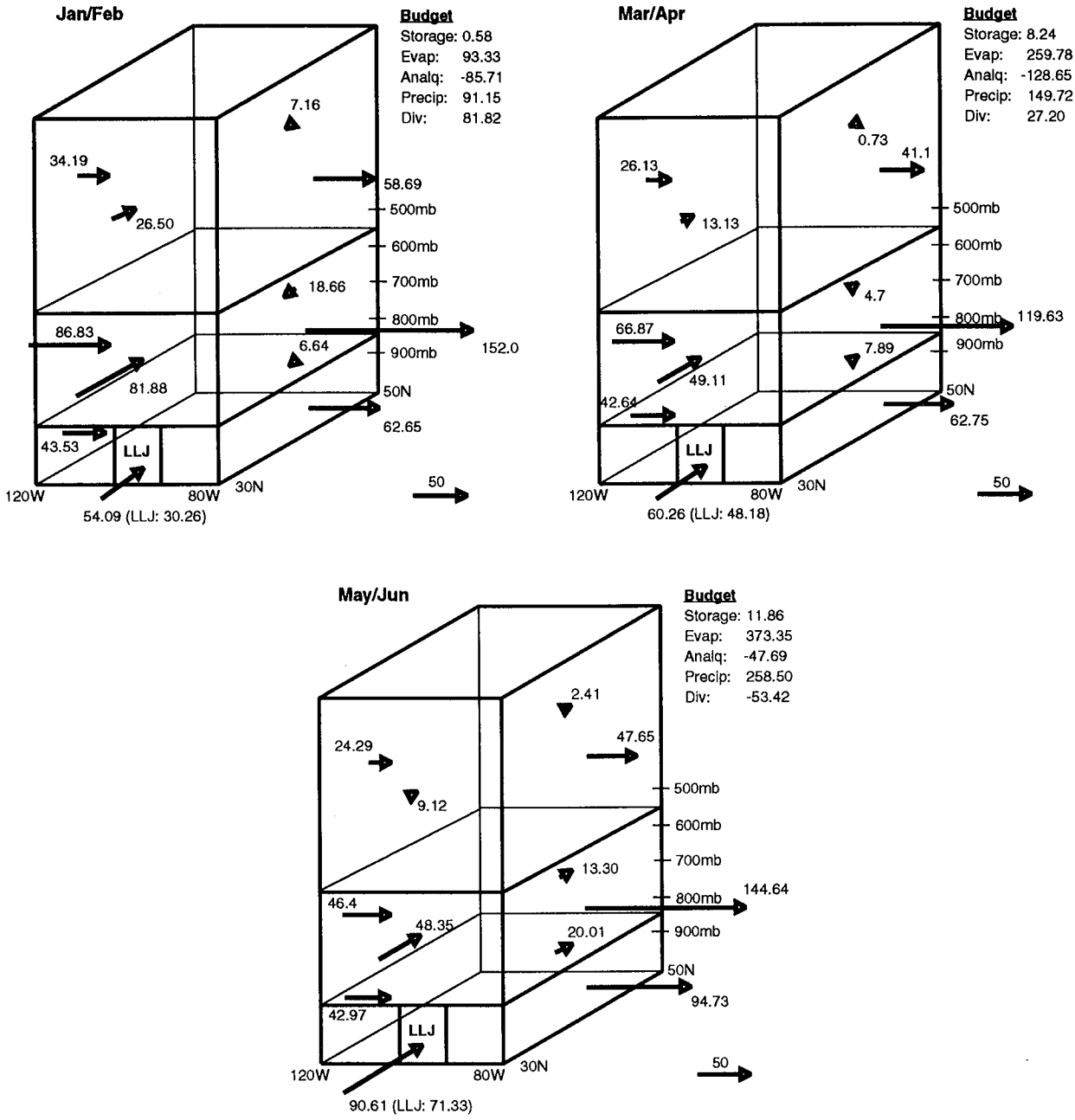


Figure 57: Moisture transport into the continental United States for January through June. LLJ (Low Level Jet) refers to the region between 92.5°W and 102.5°W, and below about 850 mb. Units are 10^6 kg/sec. To obtain units of $cm/month$ divide by 29×10^6.

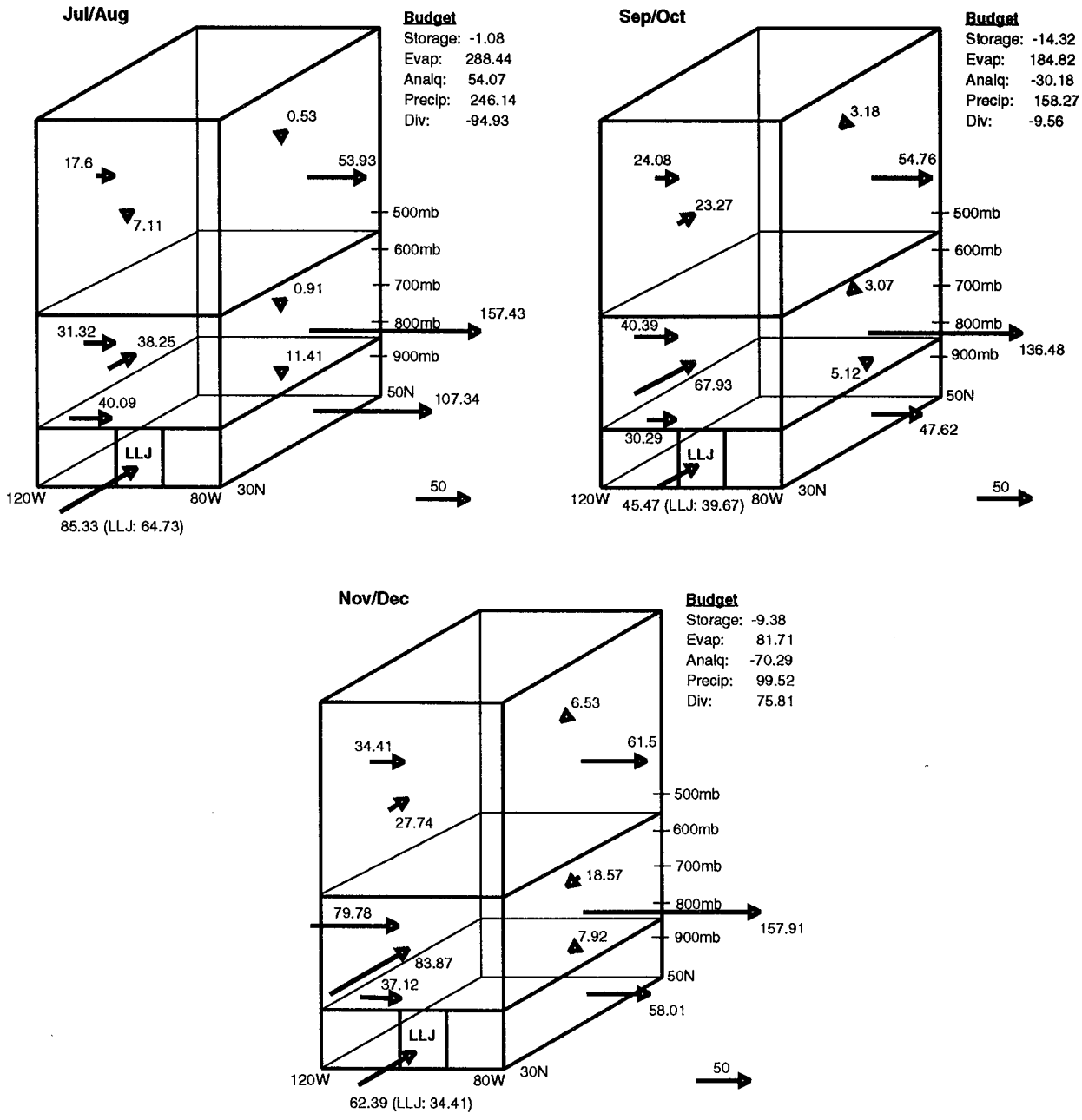


Figure 58: Moisture transport into the continental United States for July through December. LLJ (Low Level Jet) refers to the region between 92.5°W and 102.5°W, and below about 850 mb. Units are 10^6 kg/sec. To obtain units of $cm/month$ divide by 29×10^6.

C-2

4.3 Intraseasonal Variability

In this section we examine the intraseasonal variations in the GEOS-1 assimilated data. We first examine the root-mean-square differences between GEOS-1 and ECMWF analyses. These fields are based on daily averaged anomalies (the respective seasonal cycles have been removed from both for each year separately) and thus reflect the non-systematic differences (on time scales shorter than one year).

Further comparisons are made by filtering the data to isolate various time scales of interest. The Madden-Julian oscillation is compared in terms of the velocity potential at the equator filtered to retain time scales between 20 and 70 days. Shorter time-scale phenomena are examined with filters that retain time scales 10 to 20 days and 2 to 10 days.

4.3.1 Global RMS Differences

Figures 59–61 show the global distribution of the non-systematic root mean square (rms) differences in the GEOS and ECMWF wind and height fields. The rms differences were computed as follows: The fields were first averaged to produce daily mean values. Next, the seasonal cycle (computed separately for each analysis) was removed to produce anomalies. The seasonal cycle was computed separately for each year by performing a least squares fit to the data of the first two harmonics (annual and semi-annual). The squared differences of the anomalies were then averaged for each season and, finally, a square root was taken.

During both seasons the largest rms differences ($> 6m/s$) in the 200 mb zonal wind (Fig. 59) are concentrated over the equatorial oceans. Values are also relatively large at high latitudes during JJA. The large values at the equator may be partly the result of a problem with the GEOS-1 analysis scheme (see Appendix C.4 in Pfaendtner et al. 1995), which can produce unrealistic divergence when pairs of wind observations straddle the equator. The 200 mb meridional wind rms differences (Fig. 60) show latitudinally somewhat broader regions of large values ($> 5m/s$) over the wintertime subtropics and the high latitude austral winter. In the polar regions, the large rms differences during JJA appear to be the result of unrealistic ECMWF winds.

The 500 mb rms non-systematic height differences (Fig. 61) highlight the large uncertainties in the Southern Hemisphere (SH) middle and high latitudes. In the NH middle latitudes the rms is largest over the oceans (> 20) and smallest (< 20) over the land areas. That the smallest differences occur over the NH land areas is consistent with the distribution of rawinsonde stations (Fig. 97). While satellite retrievals (Fig. 100) provide temperature information on a more global basis, these are less reliable than the station observations. Further errors enter into the height field analyis as a result of errors in the first guess surface pressure field, which is needed to convert the retrieved thickness values to geopotential heights. These factors may help to explain the larger SH rms differences especially at high

latitudes. Over Antarctica the steep terrain and associated interpolation problems may also contribute to the large differences.

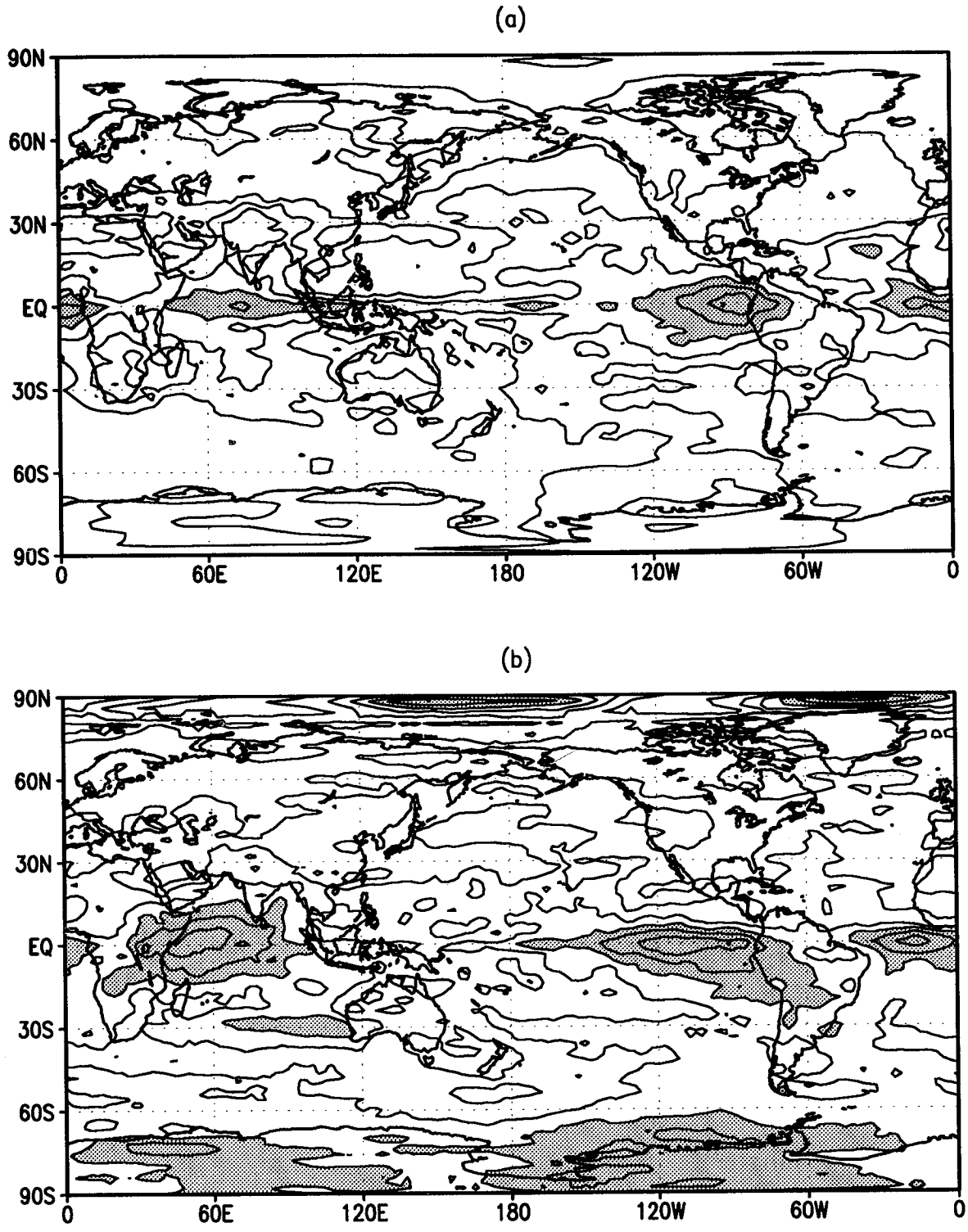


Figure 59: (a) DJF and (b) JJA 200 mb zonal wind rms (non–systematic) difference between the GEOS–DAS and the ECMWF analysis based on daily time–series. Contour intervals are 1 m/sec. Values greater than 6 m/sec are shaded.

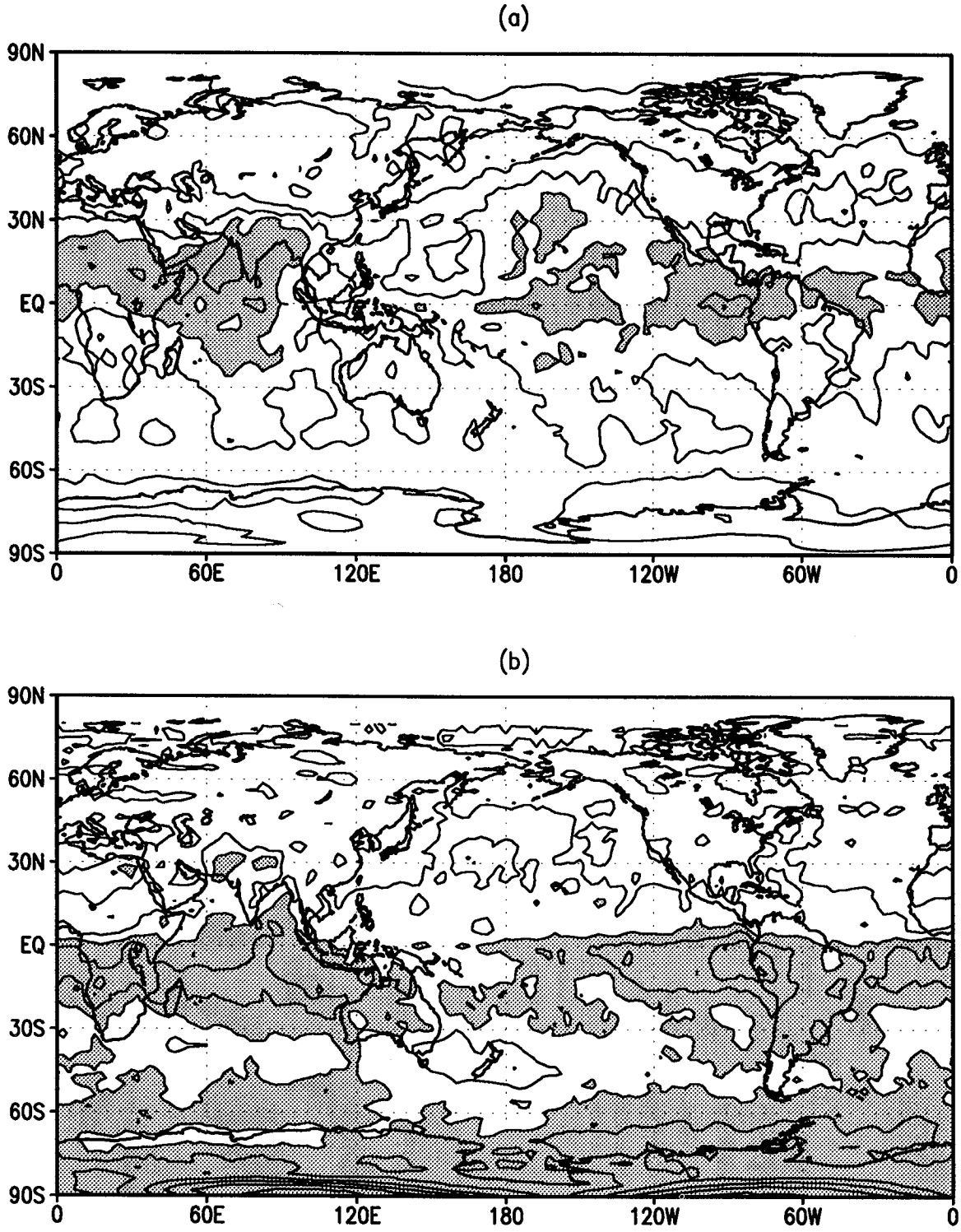


Figure 60: (a) DJF and (b) JJA 200 mb v–wind rms (non–systematic) difference between the GEOS–DAS and the ECMWF analysis based on daily time–series. Contour intervals are 1 m/sec. Values greater than 5 m/sec are shaded.

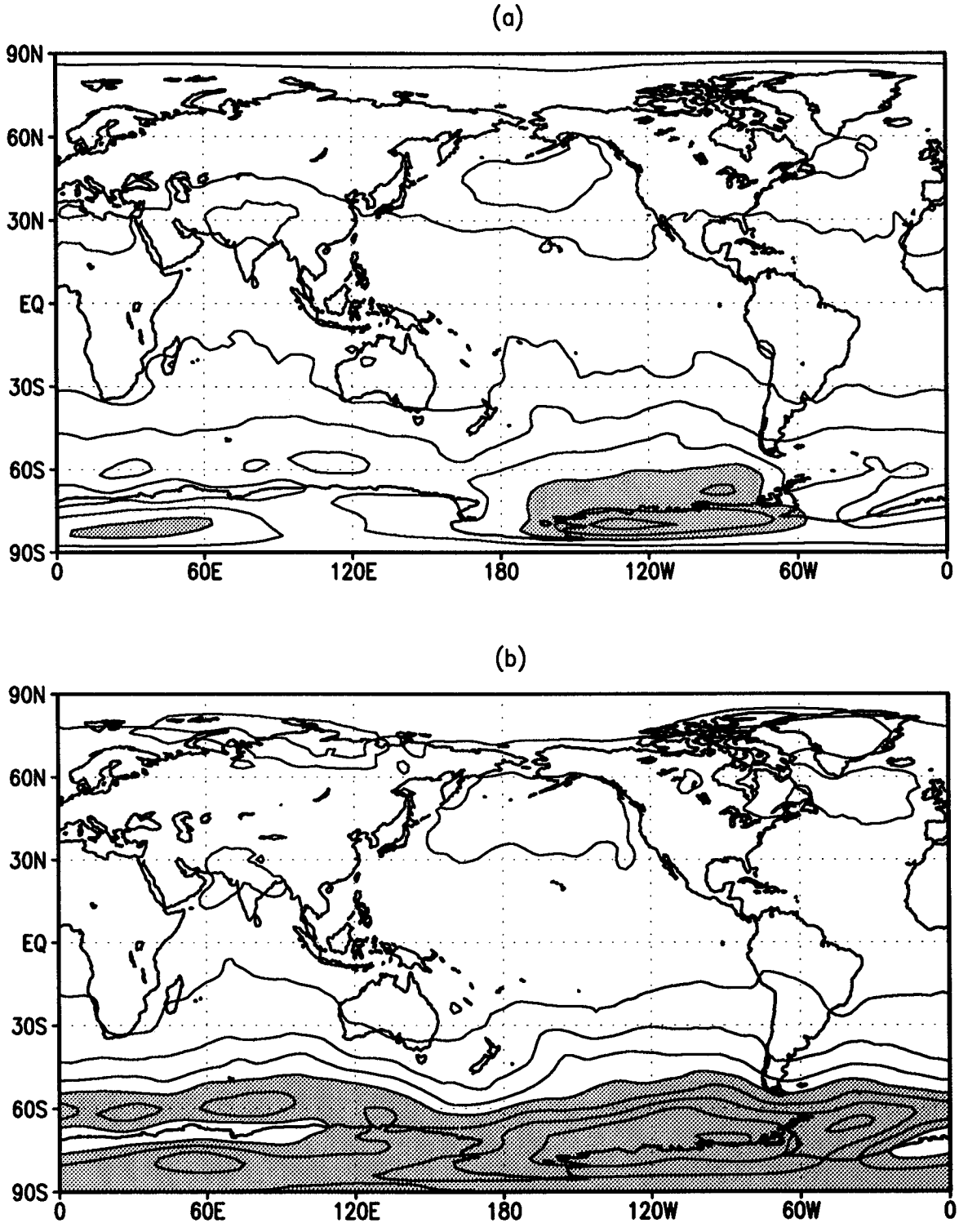


Figure 61: (a) DJF and (b) JJA 500 mb height rms (non–systematic) difference between GEOS–DAS and ECMWF analysis based on daily time–series. Contour intervals are 10 m. Values greater than 40 m are shaded.

4.3.2 The Tropics

In this section the velocity potential at 200 mb is used to look in more detail at the tropical variability. The velocity potential anomalies are produced as described in the previous section. Next three different time filters are applied to isolate time scales of 20– 70 days, 10–20 days, and 2–10 days. The algorithm for generating the filters is describe in Hamming (1989).

Figures 62–65 show time-longitude plots at the equator for 1986–89 for the 20–70 day time scales. The GEOS and ECMWF values show very similar features in this frequency band, which is dominated by the eastward propagating Madden-Julian Oscillation (MJO). In particular, the substantial interannual variability of the MJO is captured in both analyses: the El Niño year (1987) shows a very robust MJO, while the La Niña year (1988) show a very weak and disorganized oscillation.

Examples of the variations at the shorter time scales are shown in Fig. 66 for the 10–20 day filter, and Fig. 67 for the 2–10 day filter. Note that the 2–10 day fields are shown at 10°N, which is near the latitude of the ITCZ. The 10–20 day signal shows more differences between the analyses than at the lower frequencies, though the generally eastward propagation and apparent coupling with the MJO are evident in both. At the synoptic time scales (2–10 days) the fields begin to show very little resemblance, except for the overall patterns of weak and strong activity and some of the more prominent events. These differences are quantified in the next section in terms of correlations.

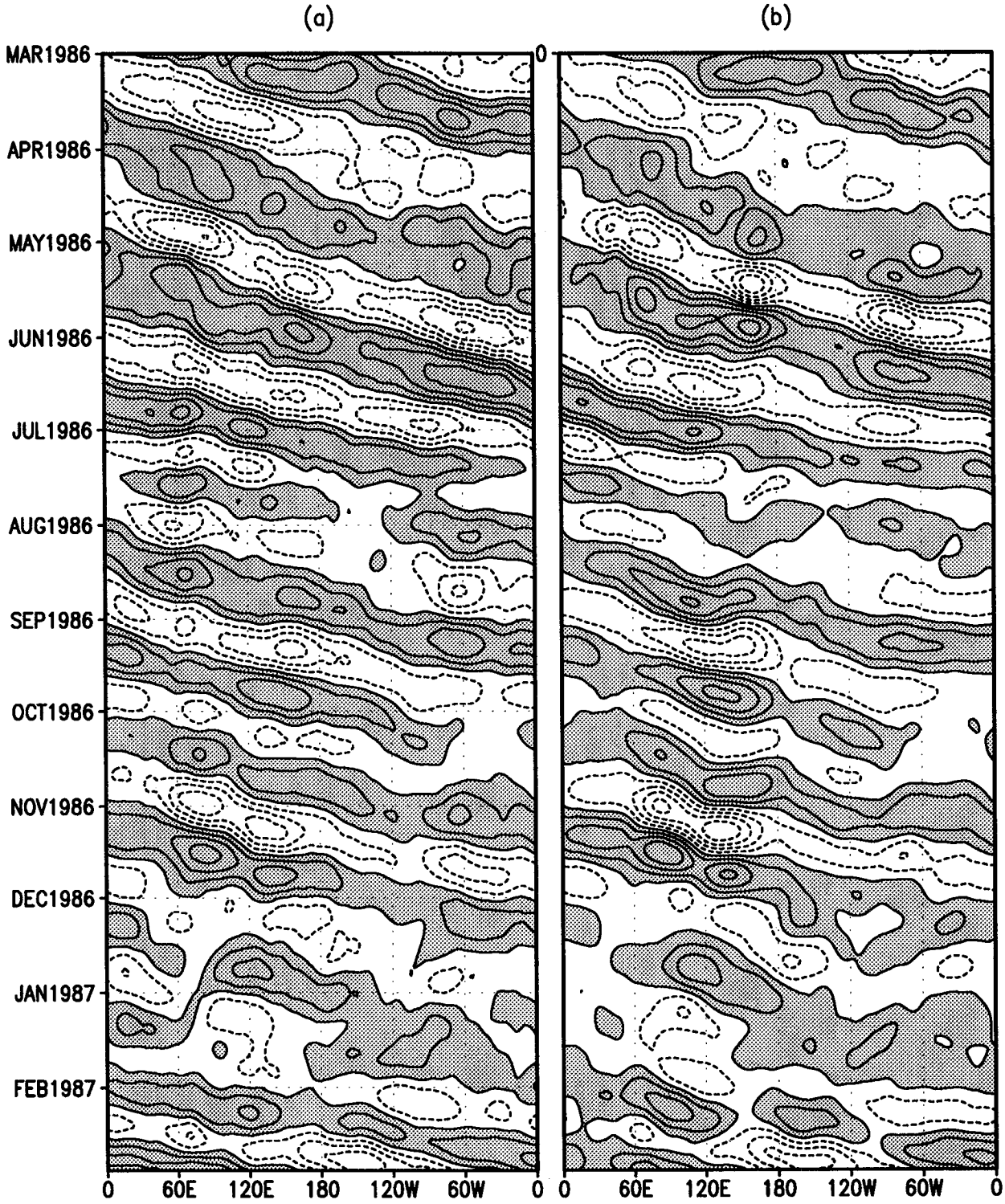


Figure 62: Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 20–70 days (1986). Contour interval is $2 \times 10^6\ m^2/sec$. The negative values are shaded.

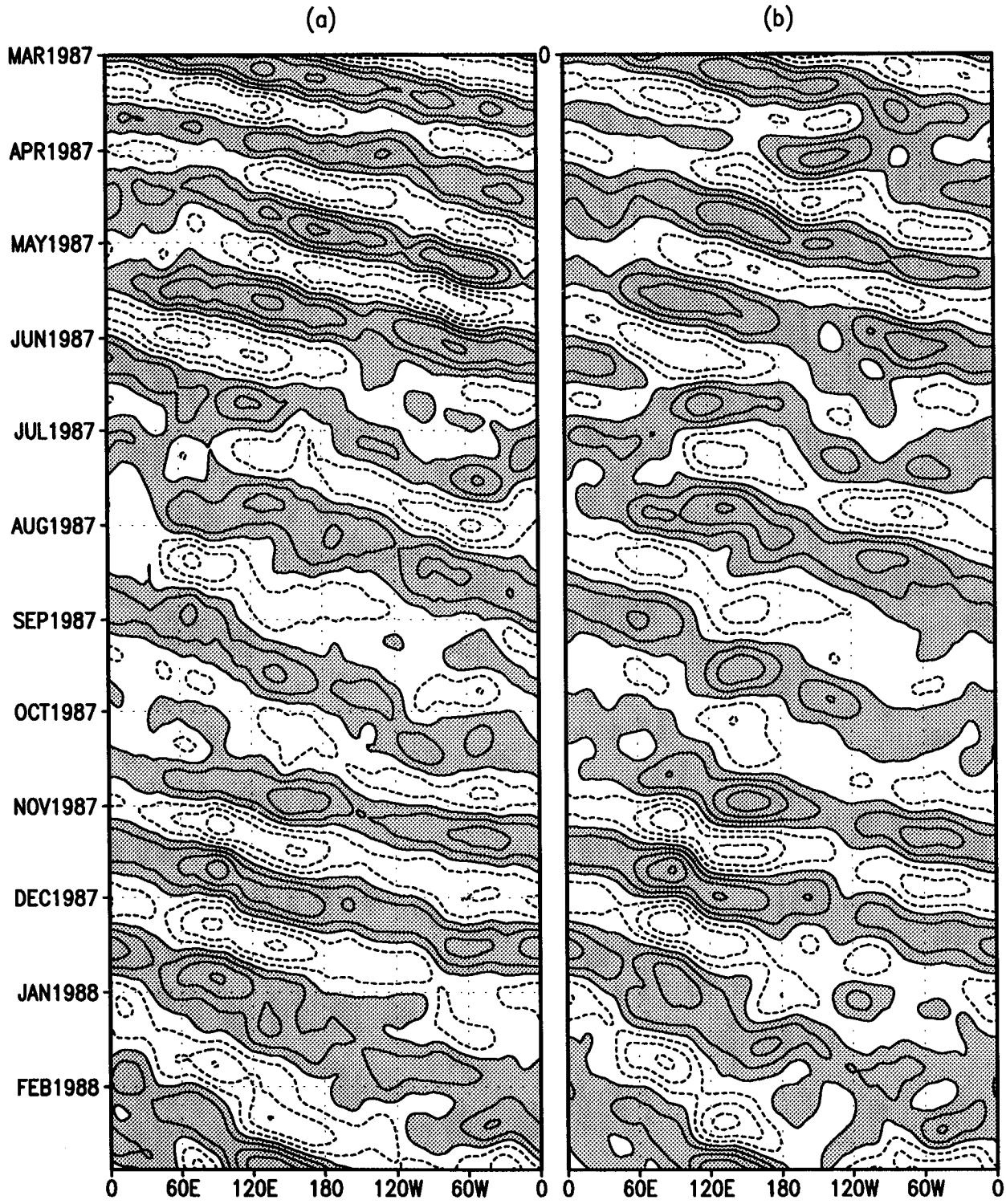


Figure 63: Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 20–70 days (1987). Contour interval is $2 \times 10^6\ m^2/sec$. The negative values are shaded.

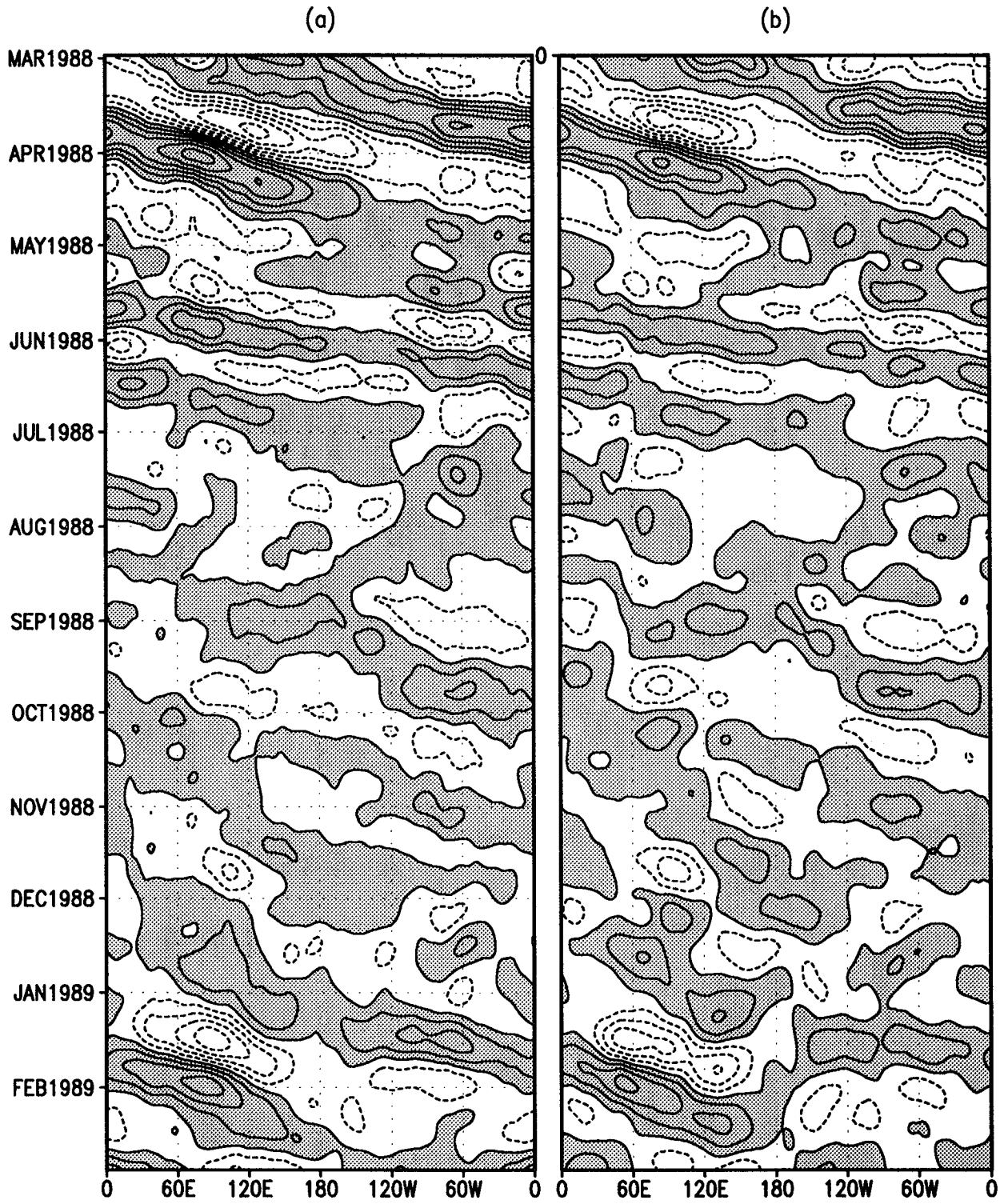


Figure 64: Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 20–70 days (1988). Contour interval is $2 \times 10^6\ m^2/sec$. The negative values are shaded.

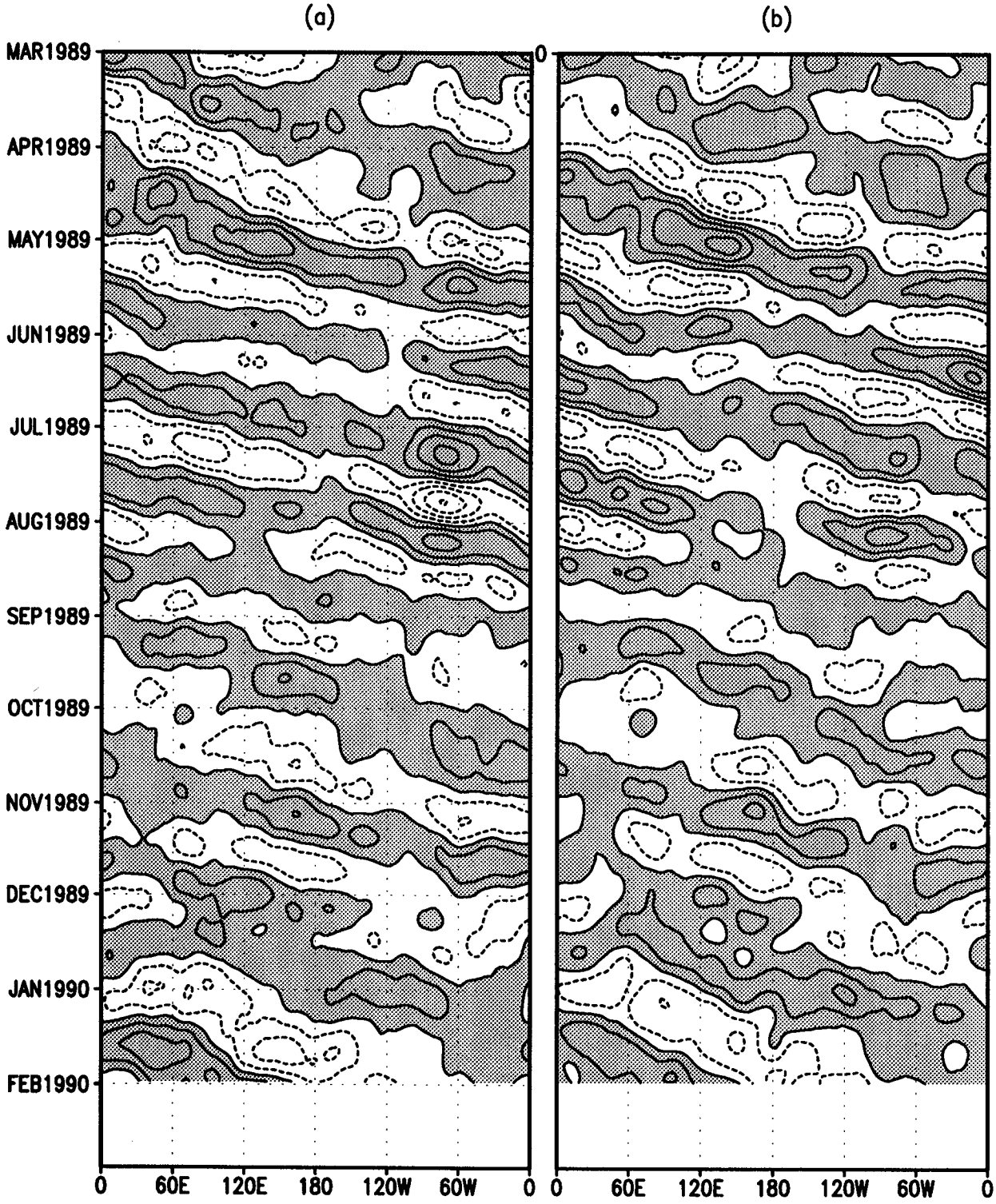


Figure 65: Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 20–70 days (1989). Contour interval is $2 \times 10^6\ m^2/sec$. The negative values are shaded.

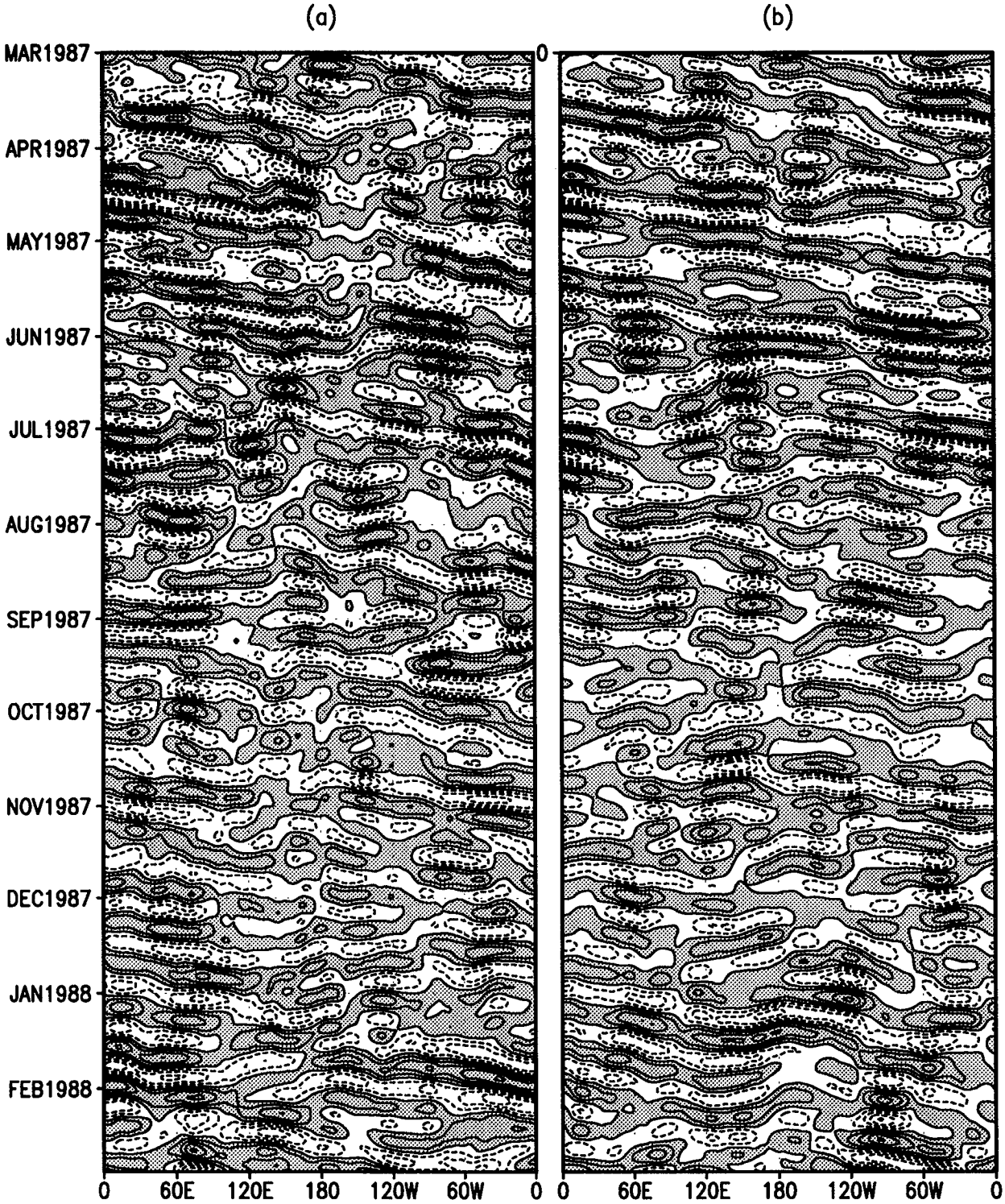


Figure 66: Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis averaged in the tropics [10S–10N]. The anomalies are departures from the seasonal cycle filtered to retain periods of 10–20 days (1987). Contour interval is $1 \times 10^6\ m^2/sec$. The negative values are shaded.

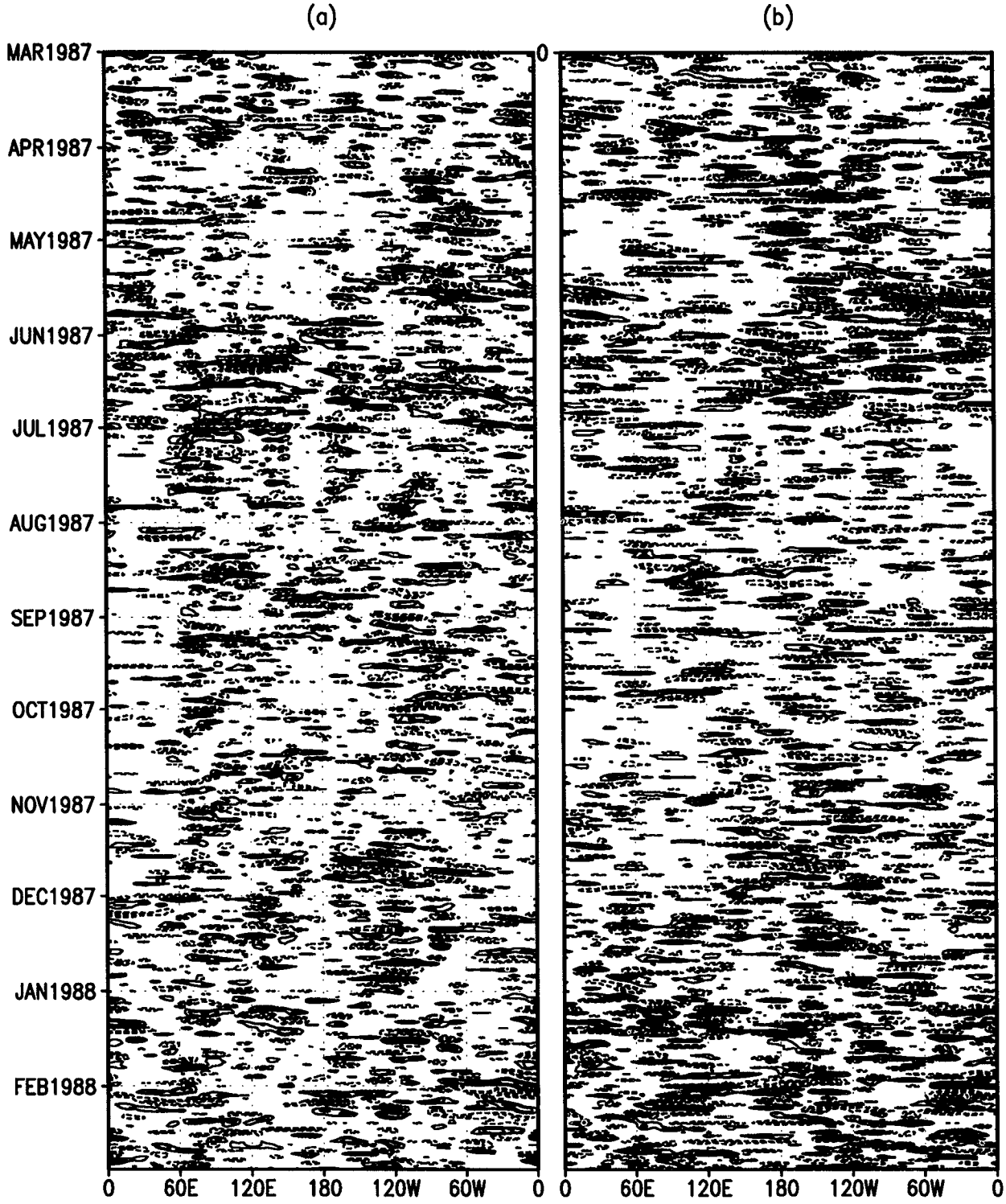


Figure 67: Longitude–time section of the 200 mb velocity potential anomalies from (a) the GEOS–DAS and (b) the ECMWF analysis at 10N. The anomalies are departures from the seasonal cycle filtered to retain periods of 2–10 days. Contour interval is $1 \times 10^6\ m^2/sec$ and contours between $-1 \times 10^6\ m^2/sec$ and $1 \times 10^6\ m^2/sec$ are not shown.

4.3.3 Correlations

Figures 68–71 show various correlations between the GEOS and ECMWF fields. The correlations are computed as follows: The fields were first averaged to produce daily mean values. Next, the seasonal cycle (computed separately for each analysis) was removed. The seasonal cycle was computed separately for each year by performing a least squares fit to the data of the first two harmonics (annual and semi-annual). The deviations from the seasonal cycle were filtered to produce three new time series with time scales of 2–10, 10–20 and 20–70 days. The quadratic terms (cross-products and variances) were then computed for each of the filtered fields over the period March 1986–February 1990. These were then used to produce the correlations.

Figures 68–69 show the correlations for the velocity potential field. The global distributions show that for the longest time scales the values are > 0.8 over most of the globe. This reflects the large scale and smoothed nature of the velocity potential and the ability of both analyses to capture these long period, planetary scale fluctuations in the divergence. At the shorter time scales, the largest correlations (> 0.7) occur over the storm track regions in both hemispheres. The line plots (Fig. 69) highlight the variations of the correlation with longitude and latitude. Most surprising is the similarity in magnitude of the correlations in the Northern and Southern Hemispheres. Note also that, for the shorter time scale tropical variations, the smallest correlations tend to occur over the Indian Ocean.

Figures 70–71 show the correlations for the 500 mb height field. The correlations are substantially higher than for the velocity potential field, with values exceeding 0.9 over most of the globe at all time scales. The largest variation is with latitude, especially for the shorter time scales, where the correlations drop to 0.7 and less; this is likely due to the weak signature of the height field in the tropics. The line plots show that the correlations are extremely high (exceeding 0.99) in middle latitudes for the intermediate and longer time scales. The drop in correlations near 100°E longitude at 40°N is likely associated with interpolation problems over the northern reaches of the Tibetan plateau.

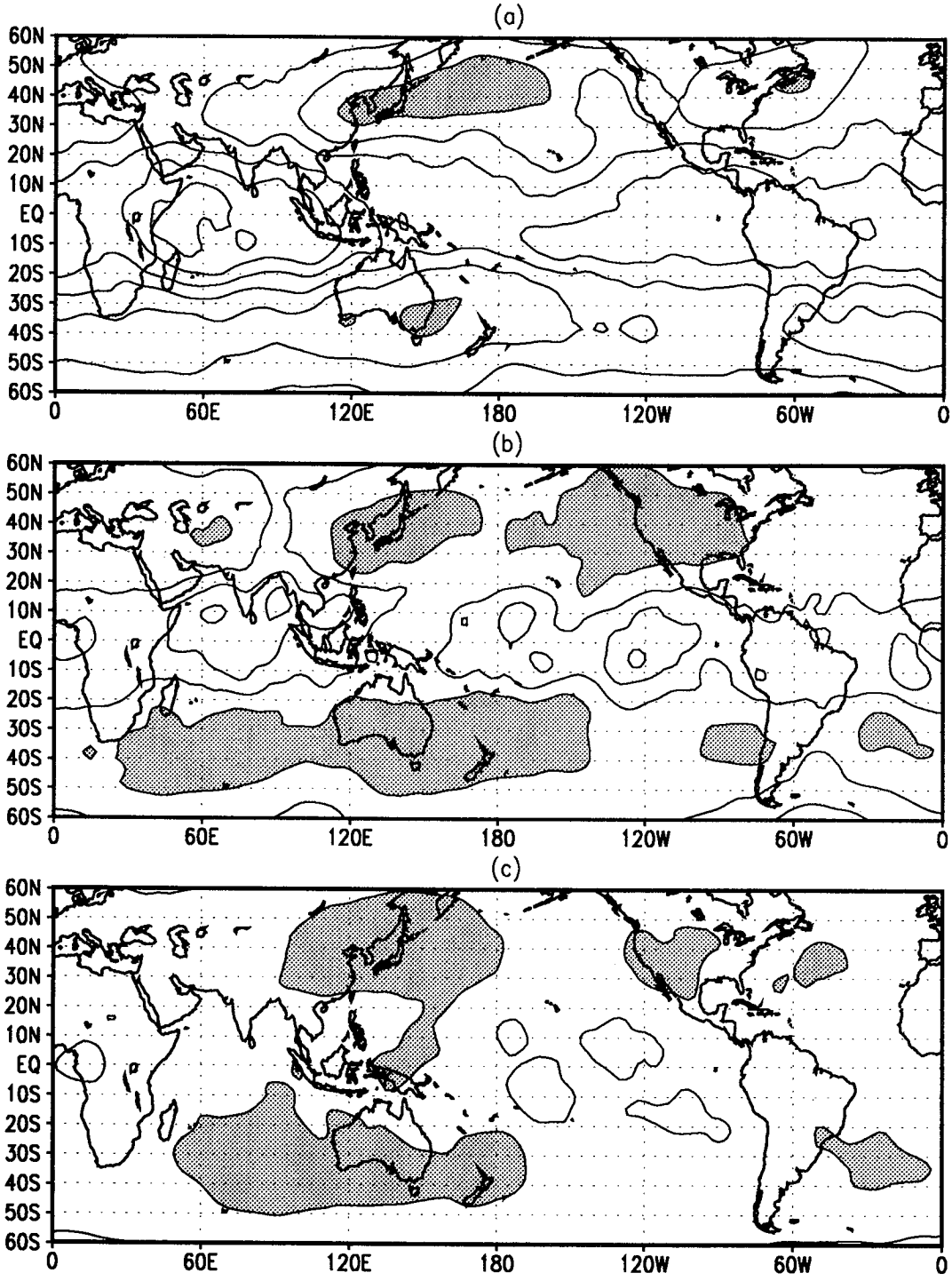


Figure 68: Correlation between the 200 mb velocity potential anomalies from the GEOS-DAS and the ECMWF analysis (Mar 1986–Feb 1990), filtered for the periods of (a) 2–10 days, (b) 10–20 days, and (c) 20–70 days. Contour intervals are 0.1. The shading represents the correlations greater than 0.7 in (a), 0.8 in (b), and 0.9 in (c).

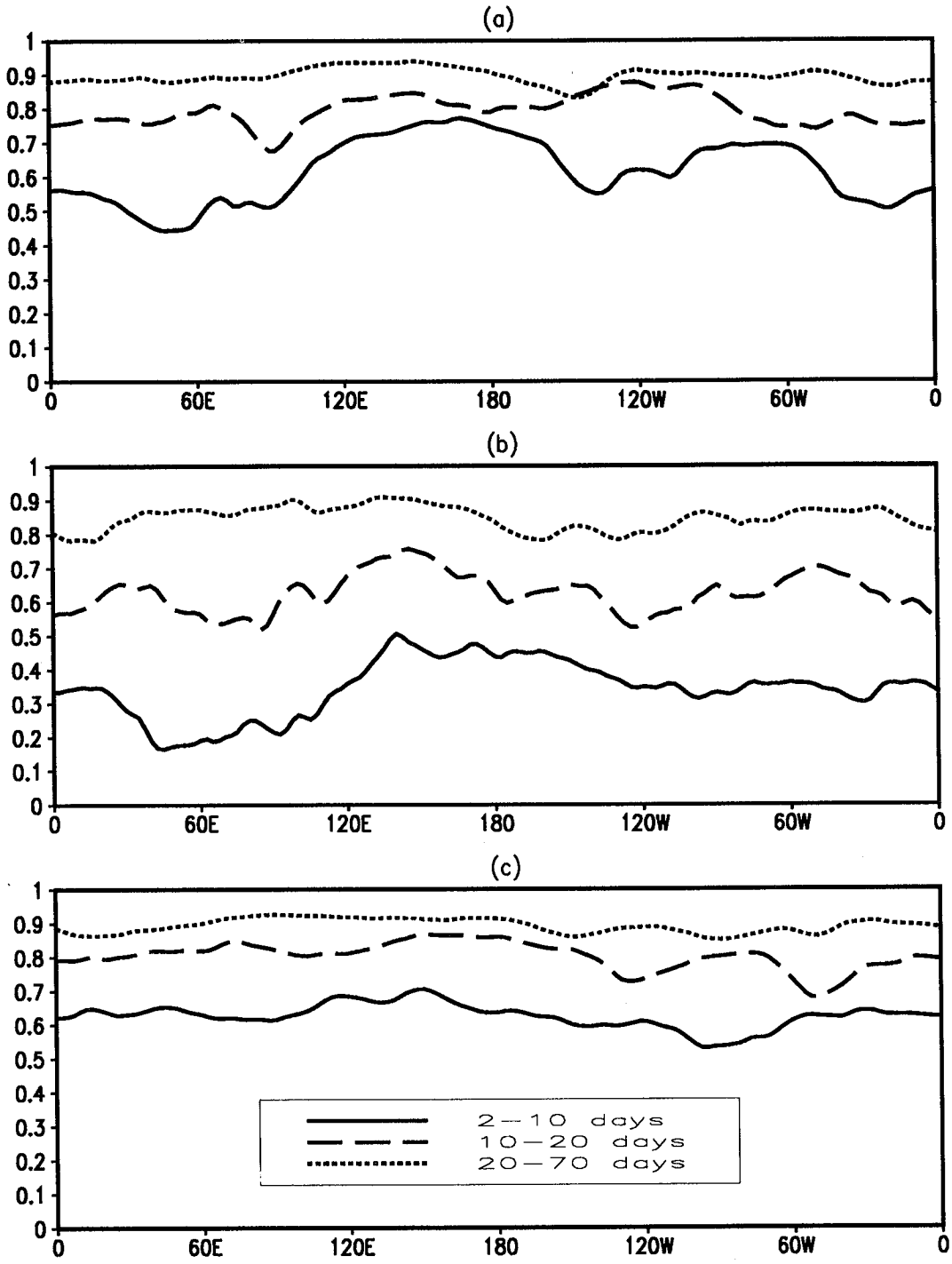


Figure 69: Correlation between the 200 mb velocity potential anomalies from the GEOS–DAS and the ECMWF analysis (Mar 1986–Feb 1990), filtered for the periods of 2–10 days, 10–20 days, and 20–70 days at (a) 40N, (b) the equator, and (c) 40S.

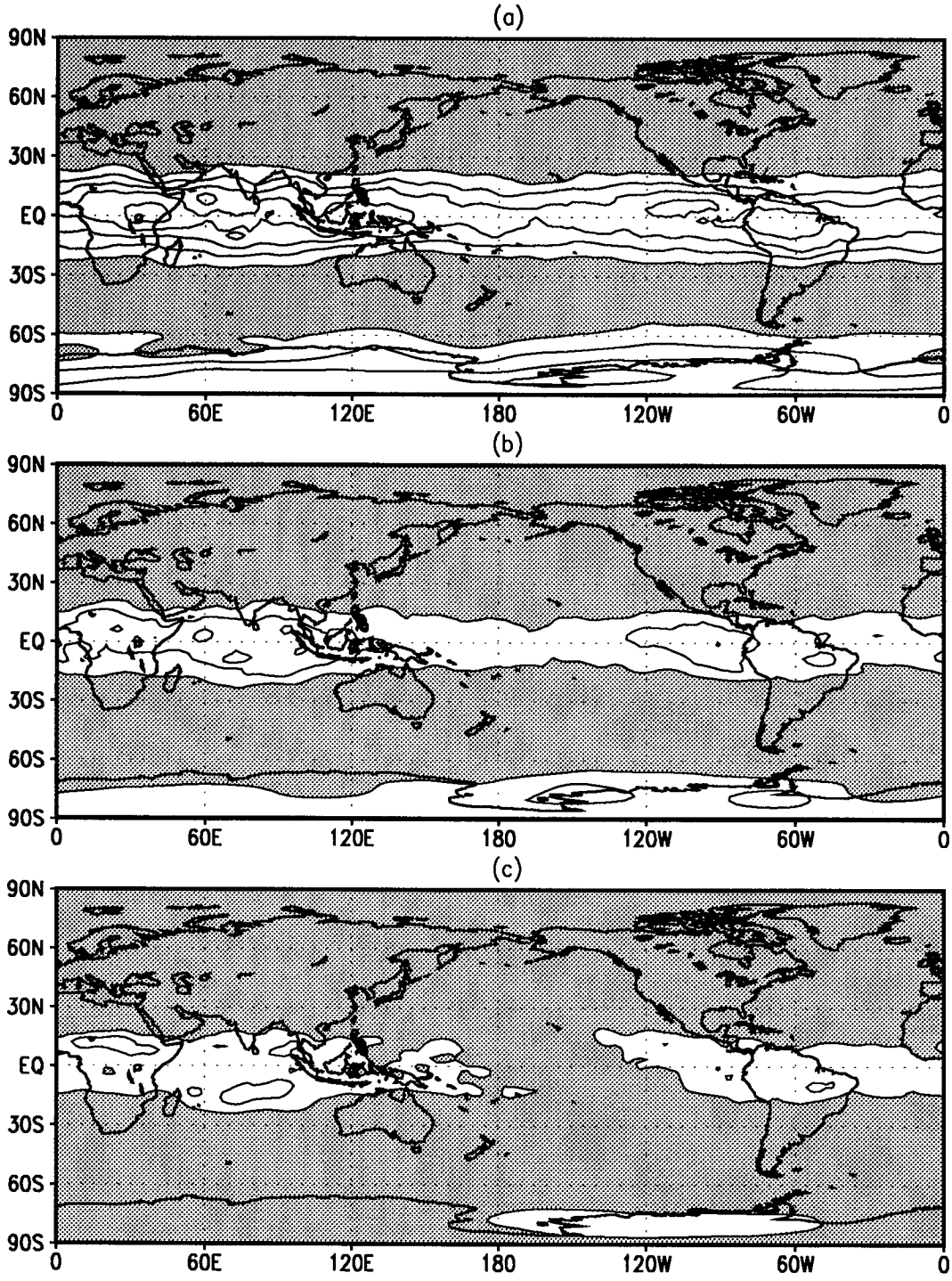


Figure 70: Correlation between the 500 mb height anomalies from the GEOS–DAS and the ECMWF analysis, filtered for the periods of (a) 2–10 days, (b) 10–20 days, and (c) 20–70 days. Contour intervals are 0.1. The shading represents the correlations greater than 0.9

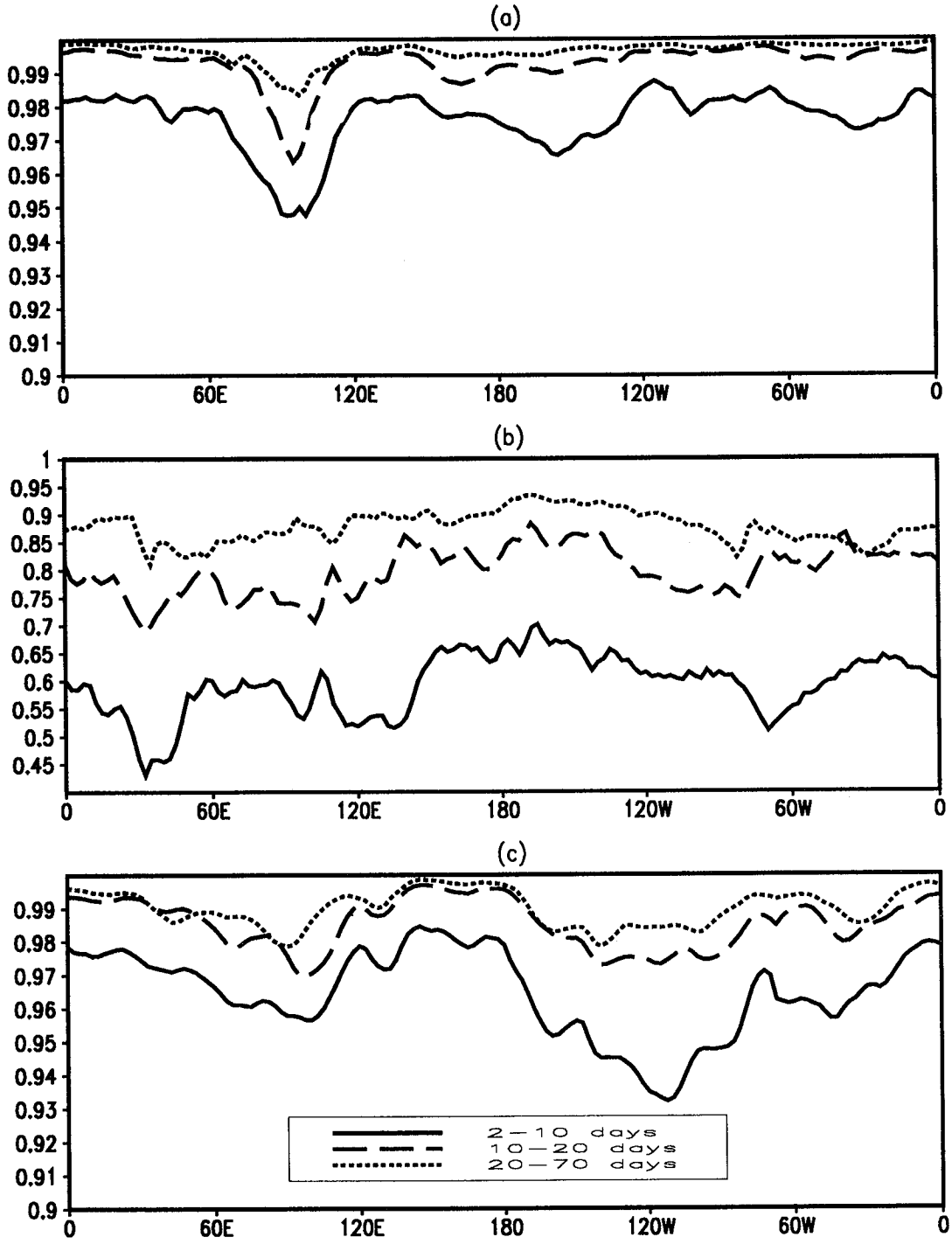


Figure 71: Correlation between the 500 mb height anomalies, filtered for the periods of 2–10 days, 10–20 days, from the GEOS–DAS and the ECMWF analysis and 20–70 days at (a) 40N, (b) the equator, and (c) 40S.

4.4 Interannual Variability

In this section we examine various aspects of the interannual variations that occurred during the 5–year period March 1985–Feb 1990. The emphasis is primarily on the intercomparisons with other estimates, including those from the operational ECMWF analyses, station observations, and satellite measurements.

4.4.1 Space–Time

Figures 72–74 compare the variations in the monthly mean zonal mean winds and height field over the five–year period for the DAO and ECMWF analyses. The variations are the anomalies from the seasonal cycle (computed separately for each analysis as a five–year average for each month). The 200 mb zonal mean zonal wind and 300 mb zonal mean height variations show a remarkable degree of similarity, suggesting that the changes made to the ECMWF operational system during this time period (see e.g. Arpe 1990) had little effect on these quantities. This is in contrast to the variations in the 200 mb zonal mean merional winds (Fig. 73), which show almost no correspondence in the two analyses. This quantity is a measure of the strength of the Hadley cell, which is poorly observed and strongly dependent on the convection scheme of the model. Arpe(1990) lists numerous changes to the ECMWF system during the period examined here, suggesting that many of the anomalies in the ECMWF 200 mb zonal mean meridional wind may be artifacts of these changes.

Figure 75 compares the outgoing longwave radiation at the equator over the Pacific region from NOAA satellite measurements, the GEOS-DAS, and the GEOS GCM simulation. All three estimates show relatively large negative anomalies (indicating enhanced deep convection) during 1987 and positive anomalies during most of 1988. The GEOS–DAS and NOAA comparison shows generally good agreement; however, the DAS results appear to overestimate the magnitude of the deep convection. Surprisingly, the GEOS–GCM results are similar in many of the main features, suggesting that much of the signal is in the SST anomalies. Some information does appear to come from the assimilated observations; for example, the anomalies in early 1986 are in both the NOAA and GEOS–DAS results, but not in the GEOS– GCM simulation. A further look at the spatial patterns of these anomalies is left for the next section.

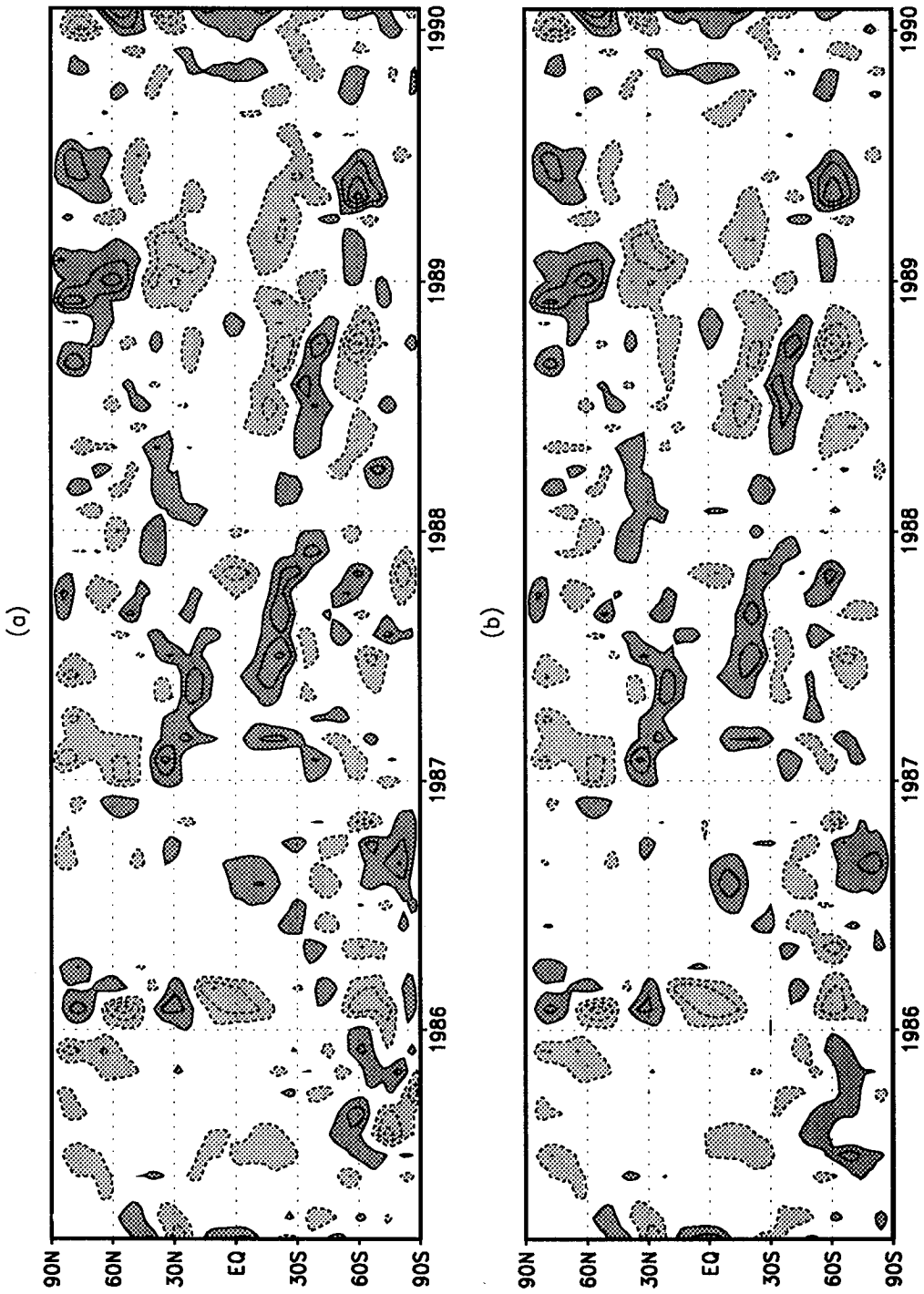


Figure 72: Departure from the seasonal cycle for the 200 mb zonal mean u-wind from (a) the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 2.0 m/sec. Zero contours are not drawn.

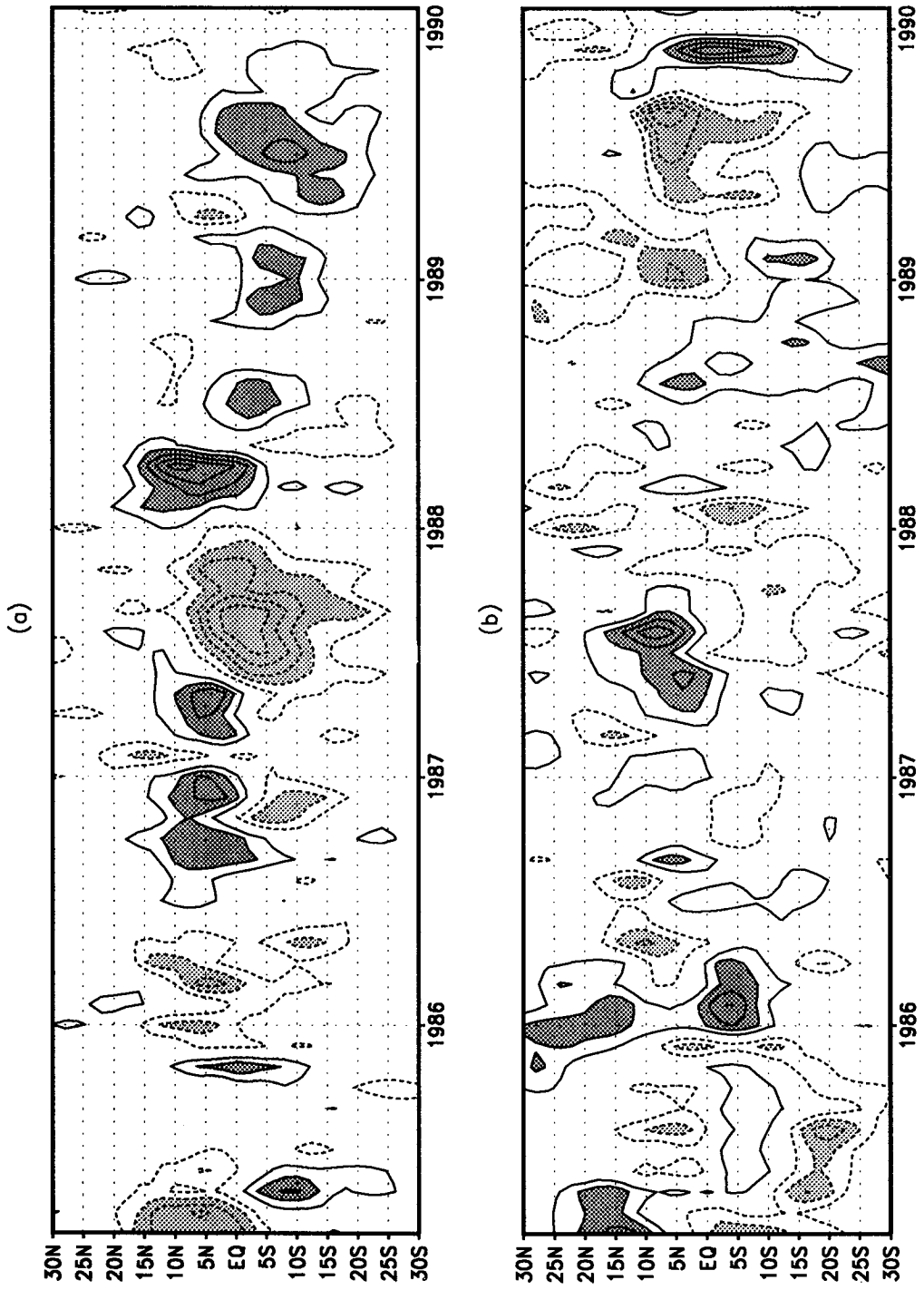


Figure 73: Departure from the seasonal cycle for the 200 mb zonal mean v-wind from (a) the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 0.2 m/sec. Zero contours are not drawn. The values greater than 0.4 m/sec and less than -0.4 m/sec are shaded.

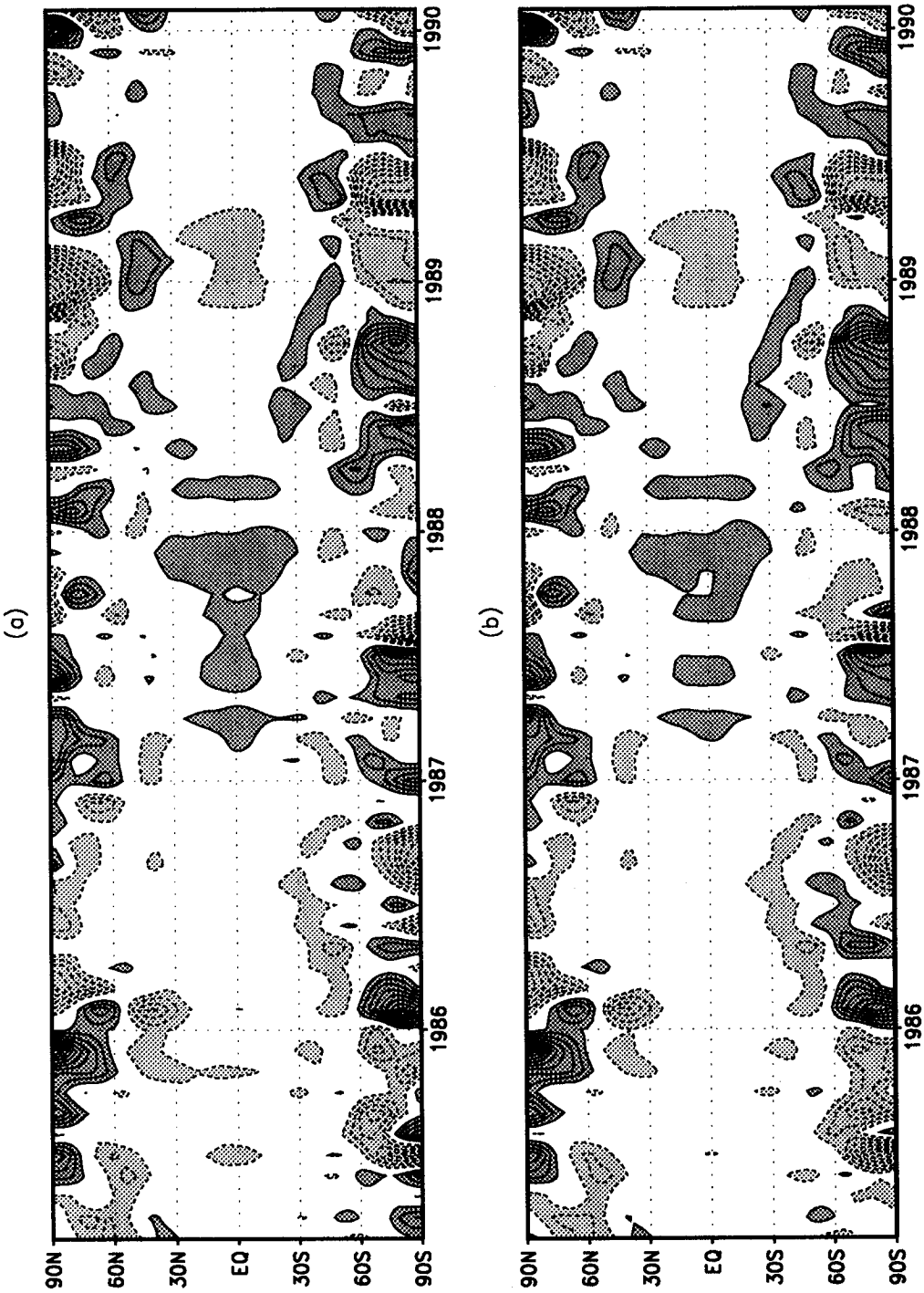


Figure 74: Departure from the seasonal cycle for the 300 mb zonal mean height from (a) the GEOS–DAS and (b) the ECMWF analysis. Contour interval is 20.0 m. Zero contours are not drawn. The values greater than 20 m and less than -20 m are shaded.

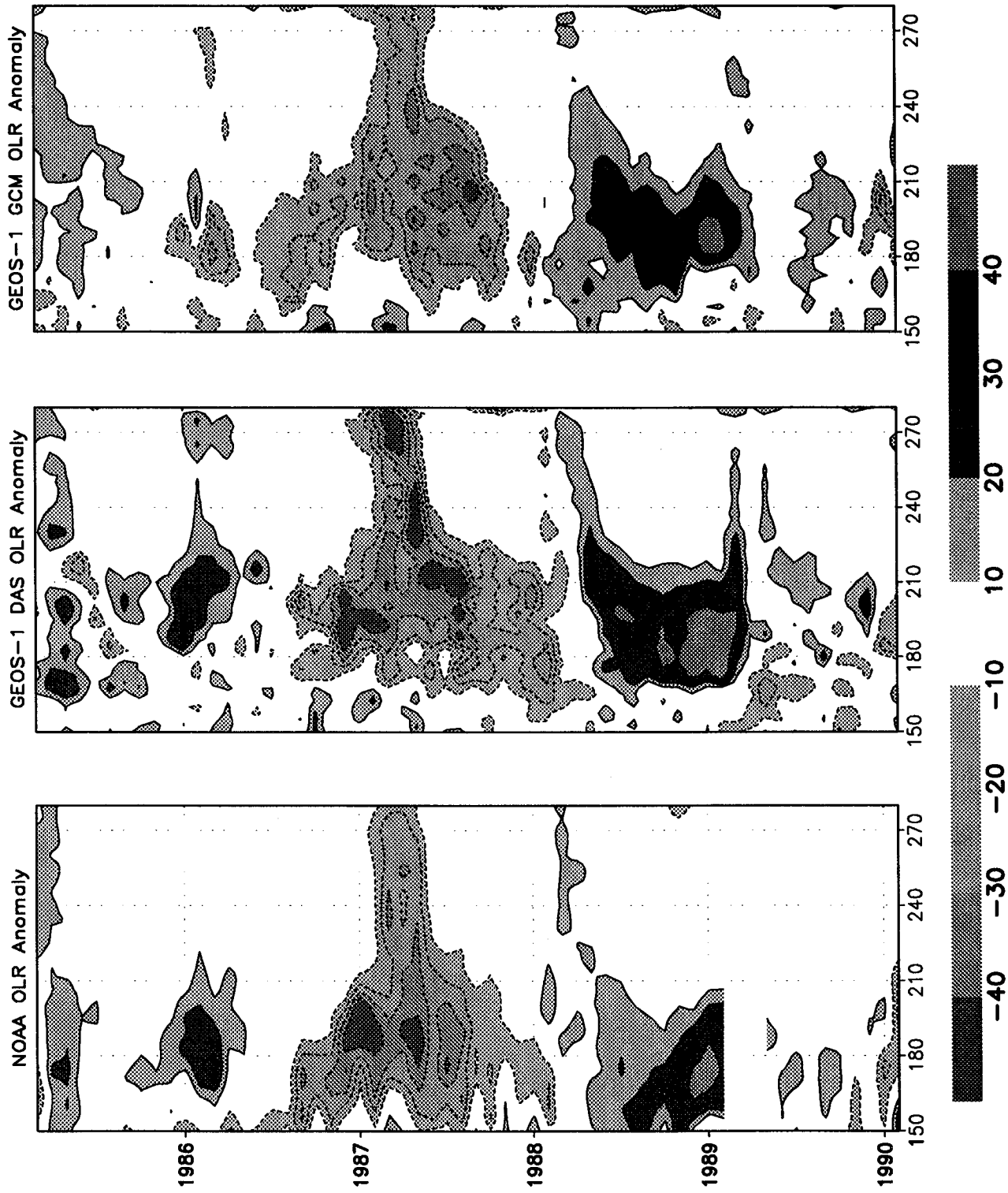


Figure 75: Longitude-time section of the outgoing longwave radiation anomaly [5S-5N] from NOAA satellite measurements, the GEOS-DAS and the GEOS-GCM simulation. Units: W/m^2

4.4.2 Global Distribution (1988–1987)

Figures 76–83 show maps of differences between 1987 and 1988 for various quantities from the GEOS-DAS, ECMWF analyses, SSM/I data, and the GEOS-GCM simulation. These two years were characterized by a transition from an El Niño to a La Niña event in 1988, and therefore provide good examples of the type of interannual variations that assimilated data must be able to capture. Figures 76–77 show the 200 mb streamfunction differences for May, June and July for the GEOS-DAS and ECMWF analyses, respectively. The fields are quite similar in the two analyses with both showing, for example, anomalous wave trains apparently emanating from the Indian Ocean (propagating southward) and from the SH middle latitudes (propagating equatorward) during May. Comparisons with Fig. 81 suggests these may be associated with anomalous convection in these regions. In general, all of the main features of the streamfunction differences are quite similar in both analyses.

The velocity potential differences (Figures 78–79) for the GEOS-DAS and ECMWF analyses are also in general agreement in the main features. For example, in July both analyses show a strong negative anomaly over the Indian Ocean and a strong postive anomaly over the eastern Pacific and Central America. However, unlike for the streamfunction differences, a closer inspection of the velocity potential anomalies reveals significant differences between the GEOS-DAS and ECMWF analyses in the detailed shape, location and magnitude of the features. These differences between the two results reflect the sensitivity of the divergence field to the convection schemes employed in the assimilation systems, as well as the fact that the observations are of insufficient quality to adequately constrain the divergence field.

Figure 80 shows the July 1987/88 differences in the total precipitable water from the GEOS-DAS and the SSM/I. The SSM/I results are only available over water surfaces (see section 4.1). Both estimates show similar features with much reduce TPW over the central and eastern tropical Pacific, and a region of enhanced TPW over the Indian Ocean extending northward over the North Pacific, and southward in the region of the South Pacific convergence zone. The SSM/I anomalies are, however, significantly stronger than the GEOS-DAS values, especially over the tropics and northern middle latitudes of the Pacific.

The OLR differences (1988-87) estimated from the GEOS-DAS (Fig. 81), NOAA satellite measurements (Fig. 82), and the GEOS-GCM simulation (Fig. 83) all show the basic La Niña (El Niño) signature consisting of enhanced OLR over the eastern tropical Pacific, flanked by a v-shaped region of reduced OLR to the west, north and south. The GEOS-DAS and NOAA OLR also show large negative anomalies over India during July associated with the strong 1988 monsoon: this feature is absent from the GEOS-GCM simulation. Other features absent from the GEOS simulation include the signature of the drought over the U.S. (June), and the large negative anomaly over the eastern North Pacific in July. The simulation also shows an apparently unrealistic tendency for a double maximum in the OLR over the central and eastern tropical Pacific. These results show that while a substantial amount of information about the interannual variations comes from the SST anomalies, the

influence is primarily local. An accurate estimate of anomalies away from the immediate influence of the SSTs requires the additional information provided by the assimilation of atmospheric observations. The lack of a double maximum in the DAS results over the central and eastern tropical Pacific suggests that the assimilation of observations also helps in providing a better estimate of the local response.

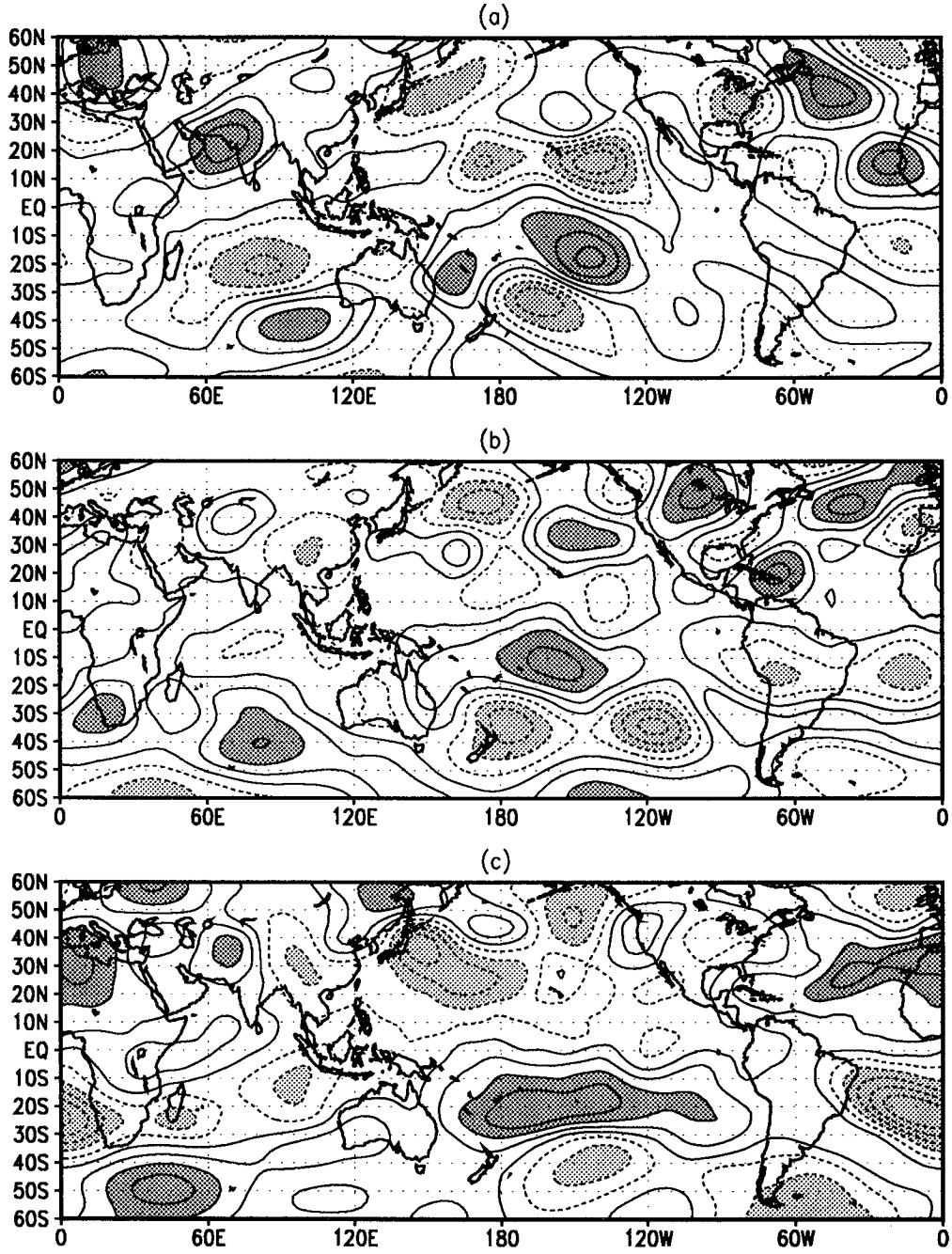


Figure 76: The 200 mb streamfunction difference (1988-87) for (a) May, (b) June, and (c) July from the GEOS–DAS. Contour intervals are $5.0 \times 10^6\ m^2/sec$. The values greater than $10.0 \times 10^6\ m^2/sec$ and less than $-10.0 \times 10^6\ m^2/sec$ are shaded.

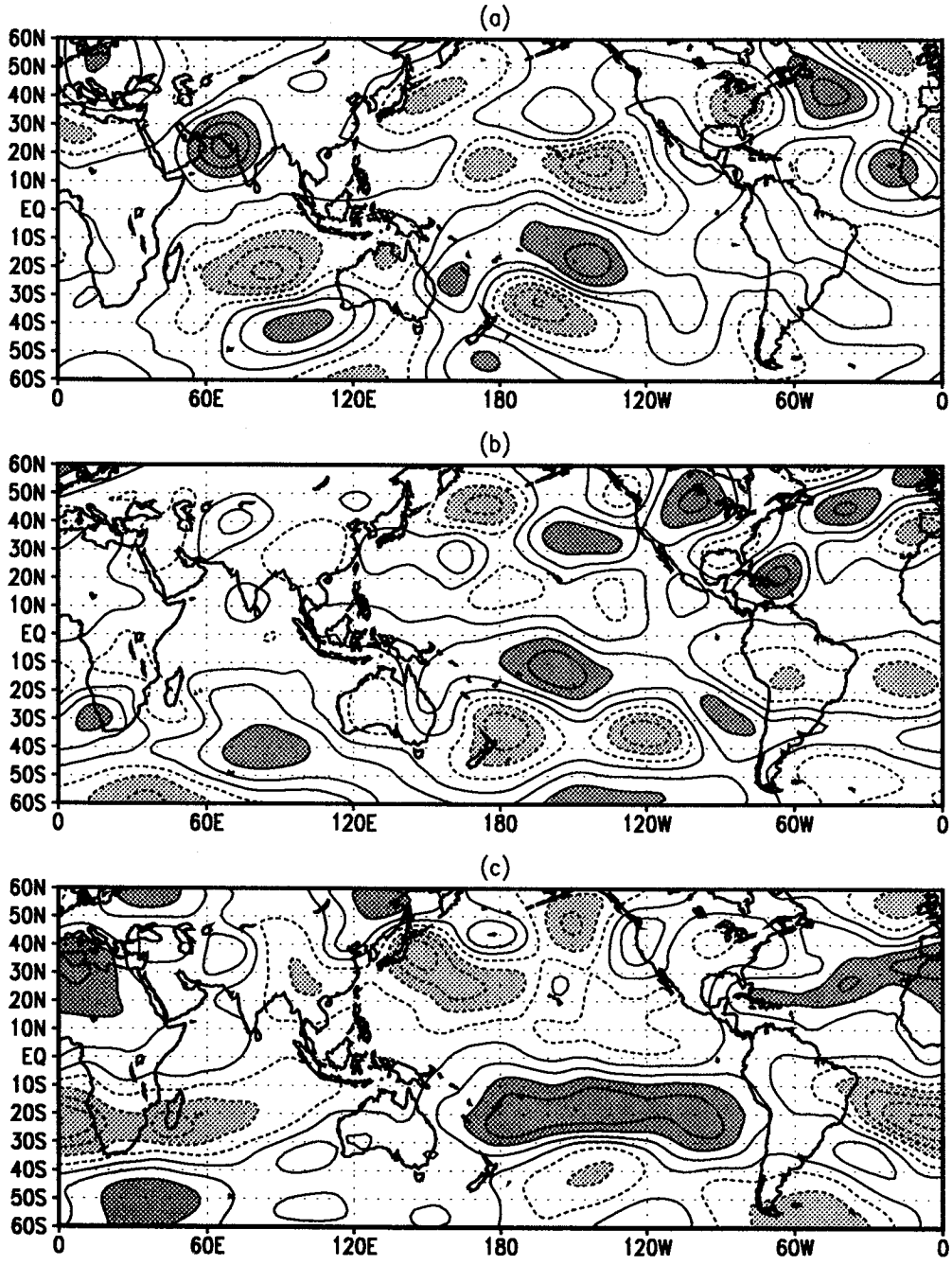


Figure 77: The 200 mb streamfunction difference (1988-87) for (a) May, (b) June, and (c) July from the ECMWF analysis. Contour intervals are $5.0 \times 10^6\ m^2/sec$. The values greater than $10.0 \times 10^6\ m^2/sec$ and less than $-10.0 \times 10^6\ m^2/sec$ are shaded.

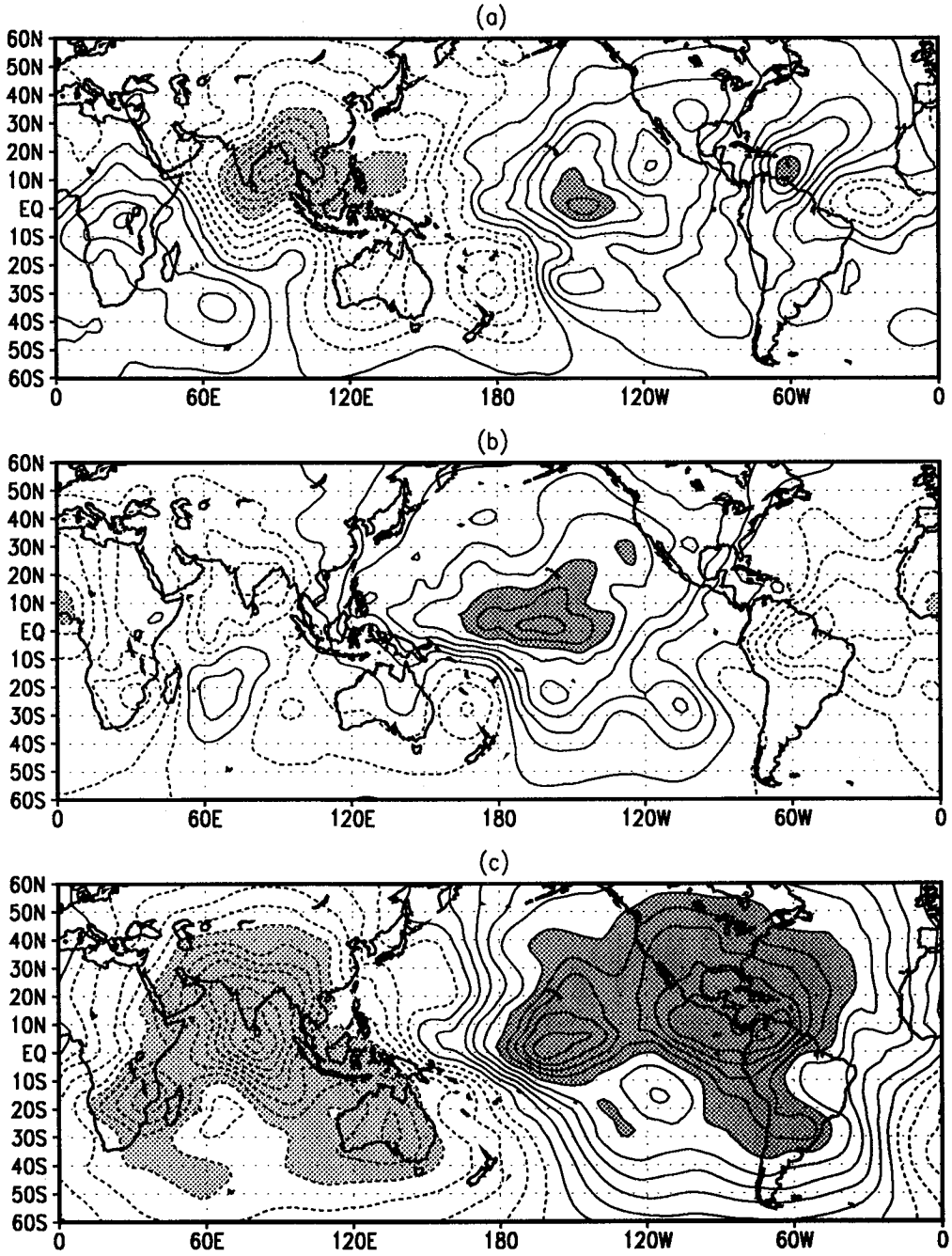


Figure 78: The 200 mb velocity potential difference (1988-87) for (a) May, (b) June, and (c) July from the GEOS–DAS. Contour intervals are $1.0 \times 10^6\ m^2/sec$. The values greater than $5.0 \times 10^6\ m^2/sec$ and less than $-5.0 \times 10^6\ m^2/sec$ are shaded.

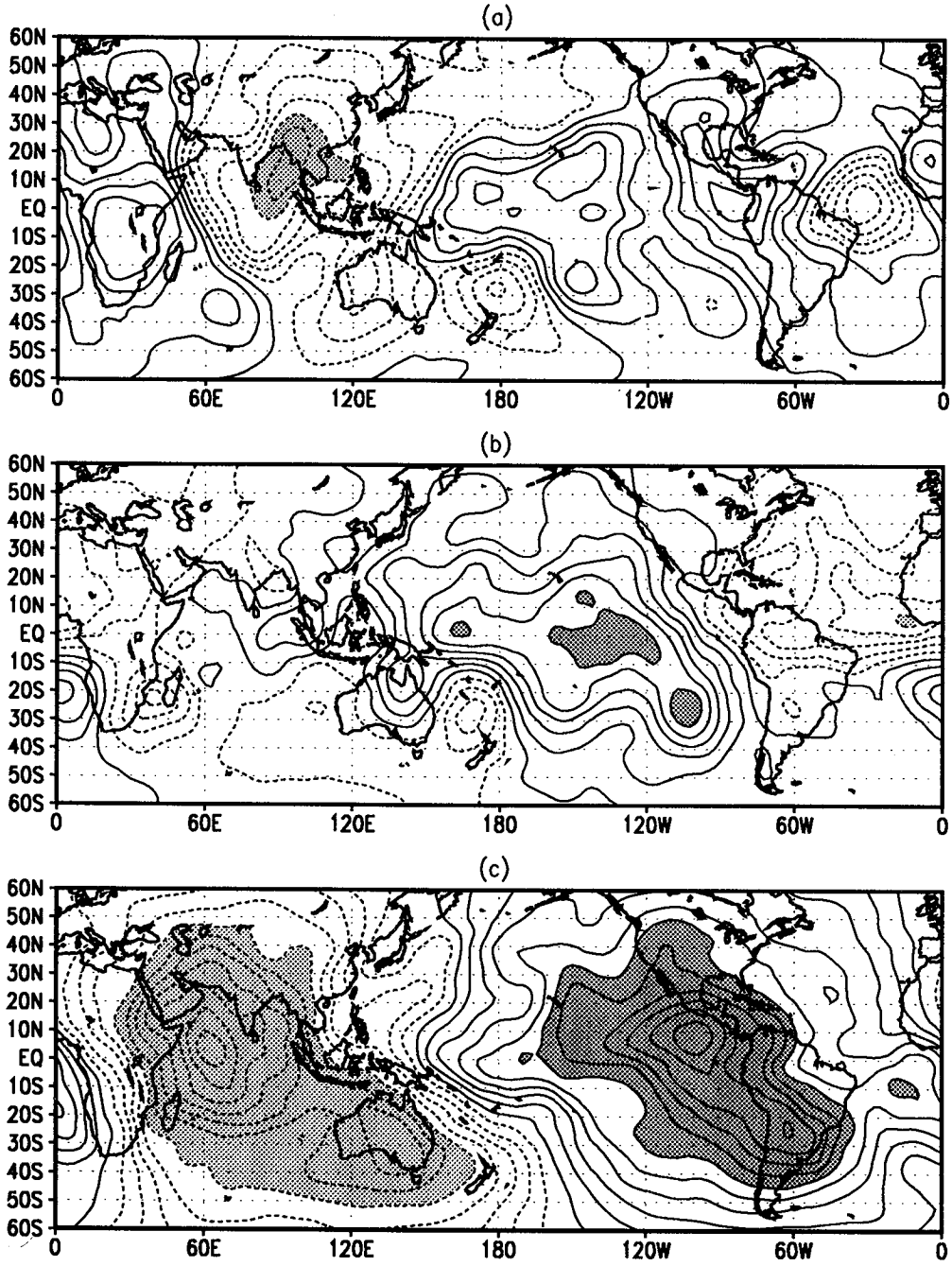


Figure 79: The 200 mb velocity potential difference (1988-87) for (a) May, (b) June, and (c) July from the ECMWF analysis. Contour intervals are $1.0 \times 10^6\ m^2/sec$. The values greater than $5.0 \times 10^6\ m^2/sec$ and less than $-5.0 \times 10^6\ m^2/sec$ are shaded.

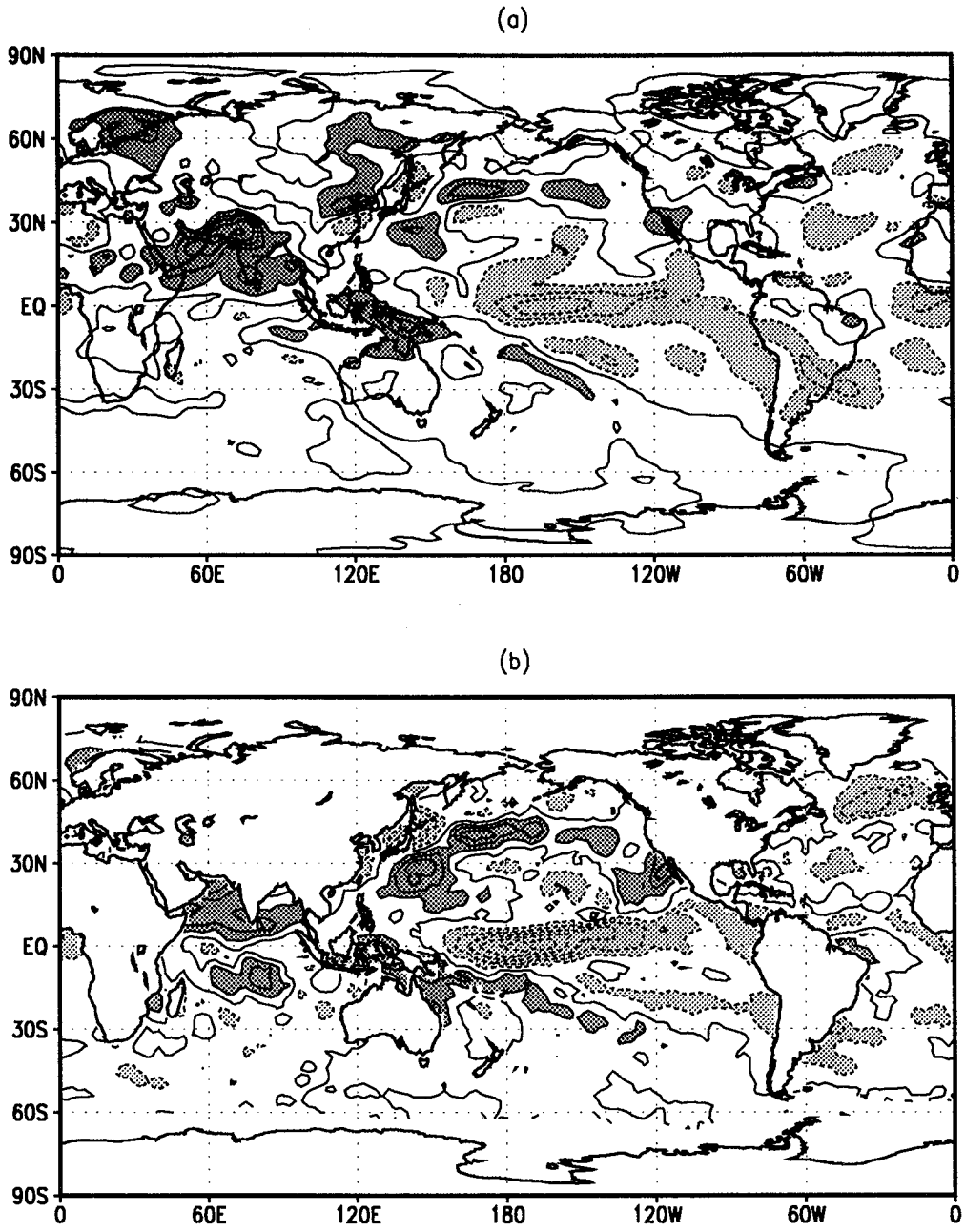


Figure 80: The total precipitable water difference (88-87) between July 1987 and July 1988 from (a) the GEOS–DAS and (b) SSM/I. Contour intervals are 0.5 gm/cm^2. The zero contours are not drawn. The values greater than 0.5 gm/cm^2 and less than -0.5 gm/cm^2 are shaded.

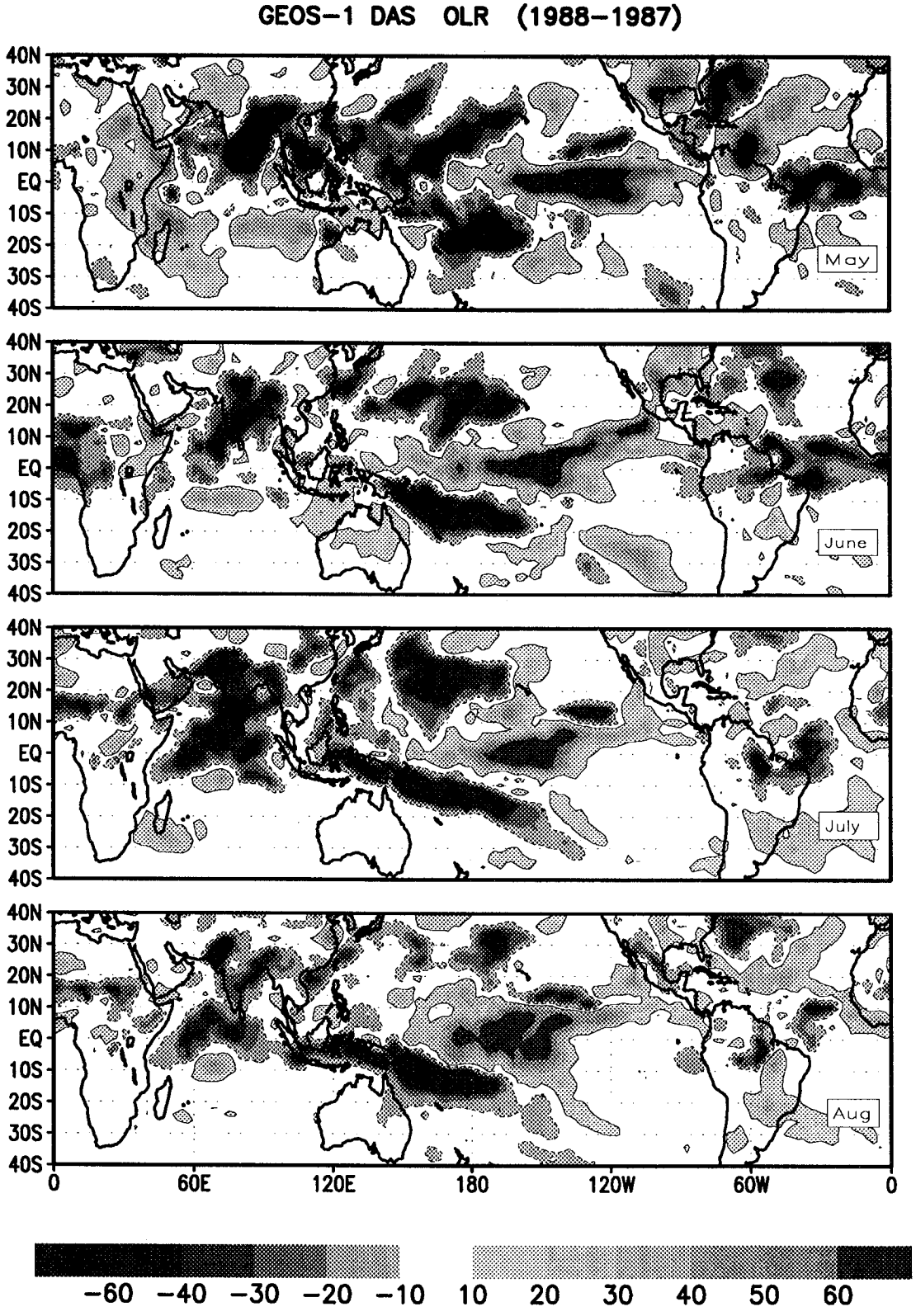


Figure 81: The outgoing longwave radiation difference (1988-87) for May, June, July, and August from the GEOS-DAS. Units: W/m^2

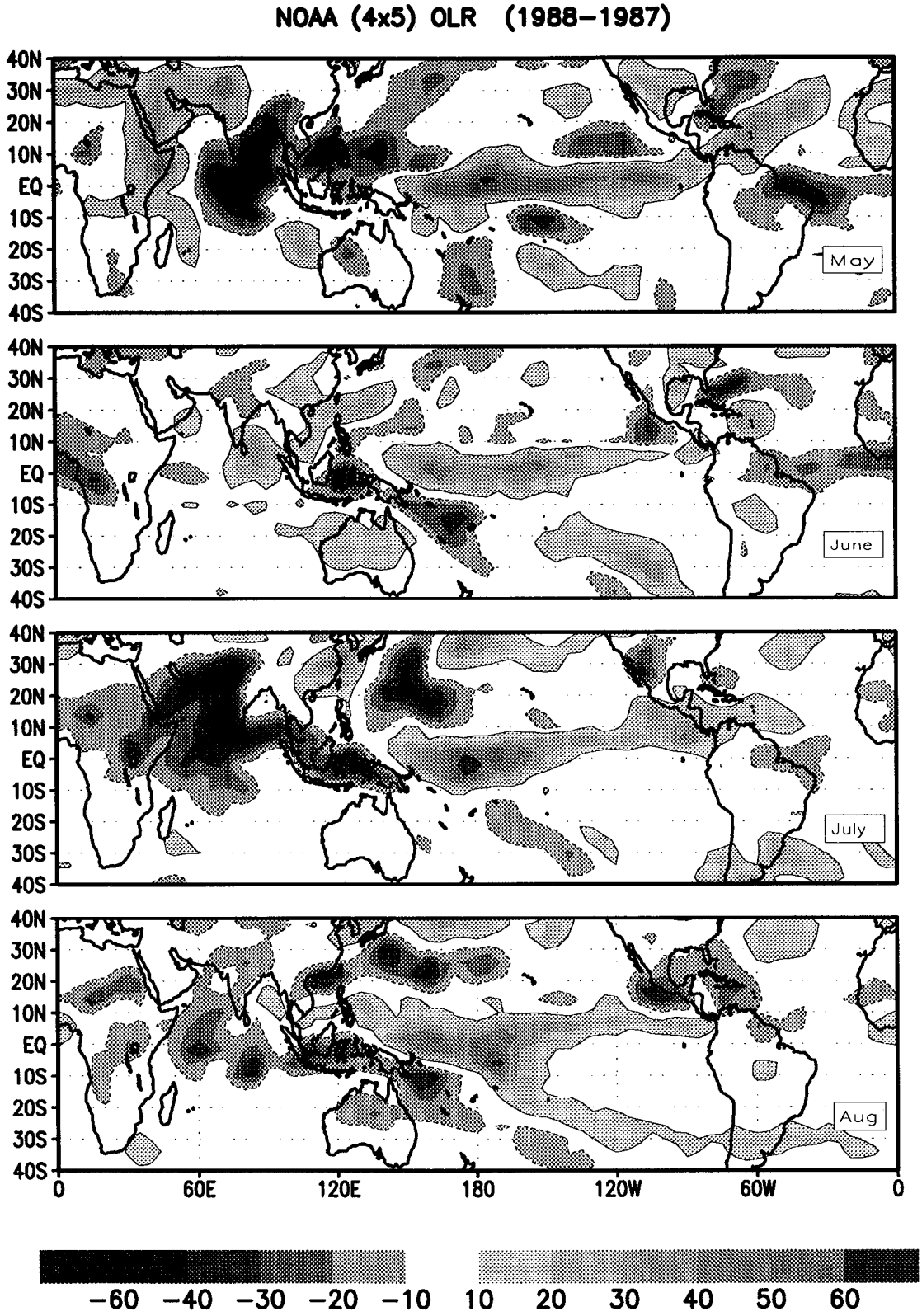


Figure 82: The outgoing longwave radiation difference (1988-87) for May, June, July, and August from the NOAA satellite observations. Units: W/m^2

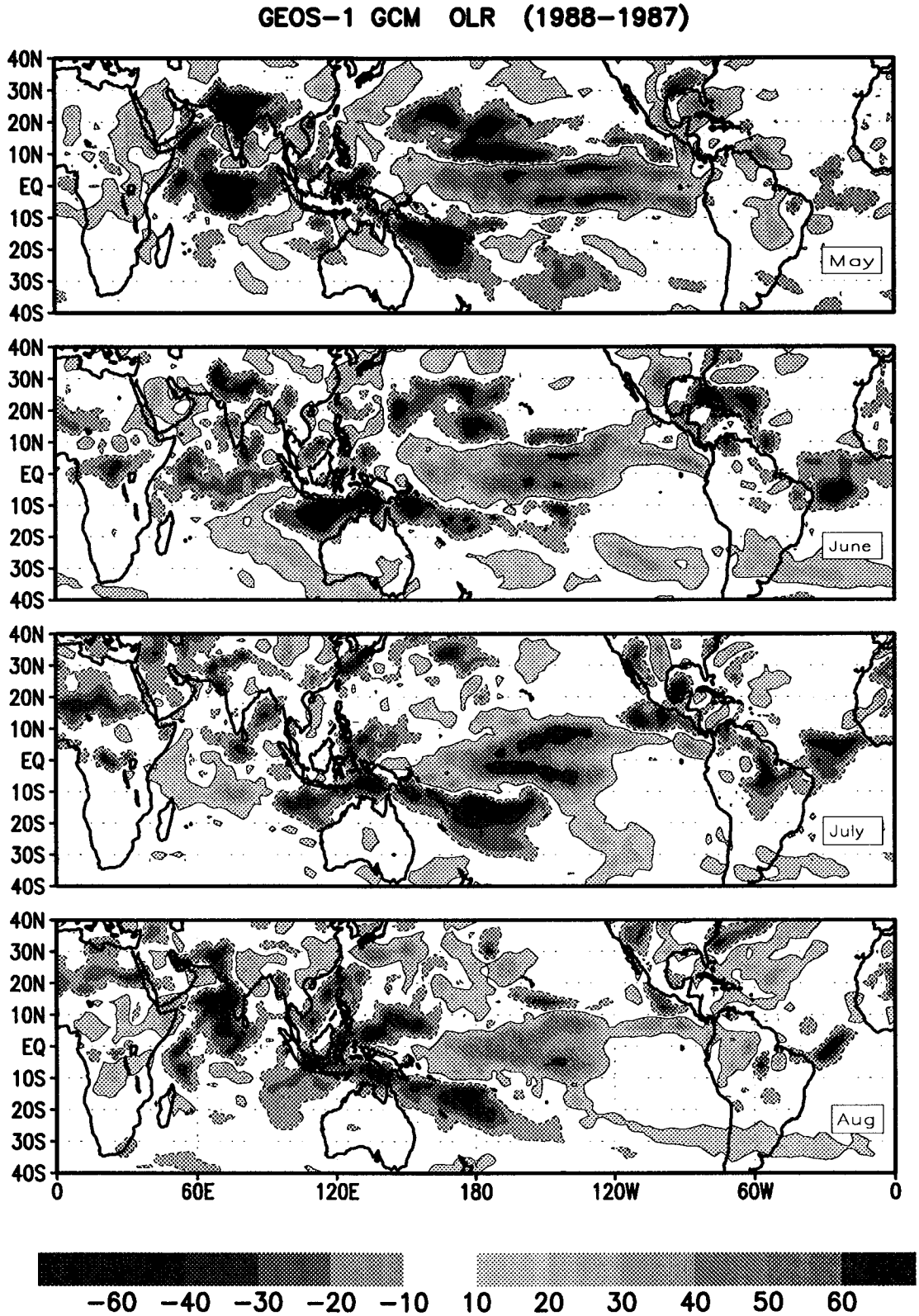


Figure 83: The outgoing longwave radiation difference (1988-87) for May, June, July, and August from the GEOS-GCM simulation. Units: W/m^2

4.4.3 Precipitation Anomalies

Area-Averages

Figures 84 and 85 compare the area-averaged precipitation from the GEOS-DAS and the world monthly station observations (see section 4.1) for the land regions outlined in Fig. 47. The comparison shows generally good agreement over North America. Over other regions there are substantial differences, with a tendency for the GEOS-DAS results to underestimate the station anomalies. It is unclear how much of these differences are the result of deficiencies in the GEOS-DAS or problems with the station observations (e.g., erroneous observations, sampling problems, errors of representativeness, etc.). These differences point to the need for further comparisons with other analyses and independent estimates of preciptation.

United States

The relatively good comparison between the monthly precipitation anomalies over North America found in the previous section warrants a closer look at these anomalies. This section presents the spatial maps of the U.S. precipitation anomalies from the GEOS-DAS and the observations based on the comprehensive U.S. climate division precipitation data set (see section 4.1). The results are shown in Figs. 86–95 for every other month over the entire five–year period. The comparison provides rather mixed results, with some months showing fairly good agreement while other others show very poor agreement. The assimilation tends to produce a more realistic precipitation climatology during the transition seasons (see section 4.2.4) and there is some indication that the anomalies are better reproduced during these months as well. The rather coarse resolution of the GEOS-DAS is likely one of the factors responsible for the differences, though further work needs to be done to try to assess any systematic problems in these results.

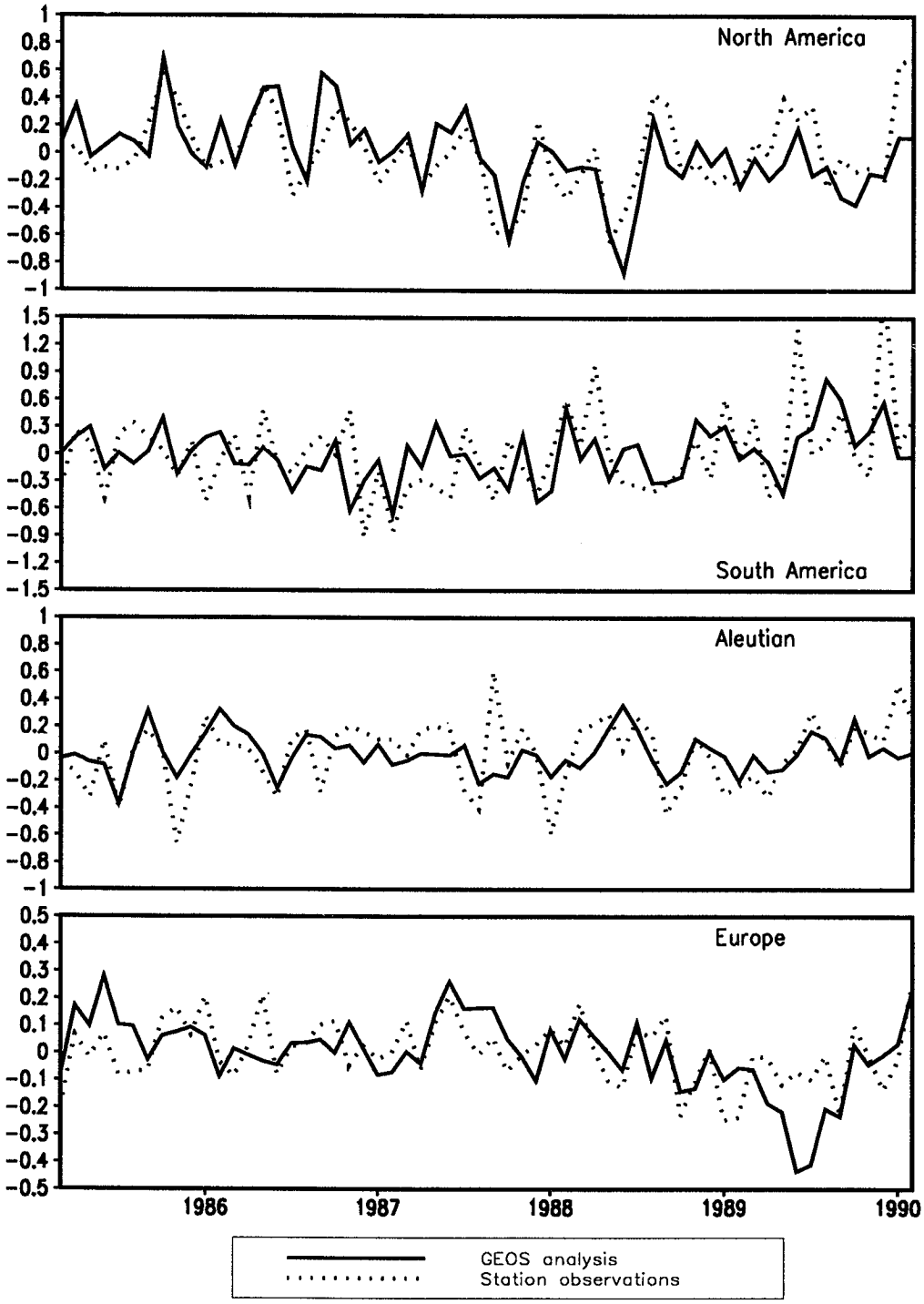


Figure 84: The regional precipitation anomalies over land from the GEOS analysis and the station observations. The anomalies are departure from the seasonal cycle. Units are *mm/day*.

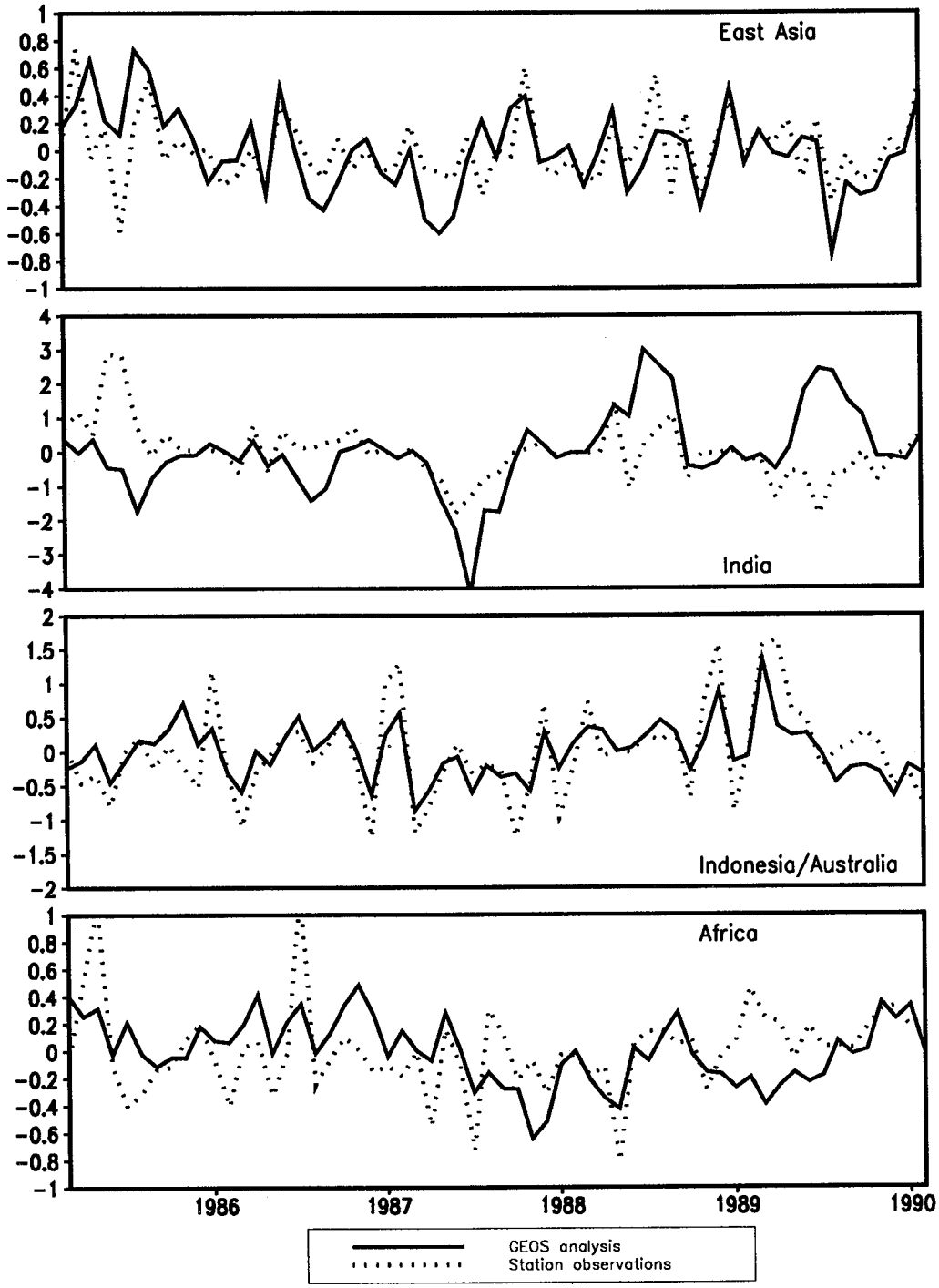


Figure 85: The regional precipitation anomalies over land from the GEOS analysis and the station observations. The anomalies are departure from the seasonal cycle. Units are mm/day.

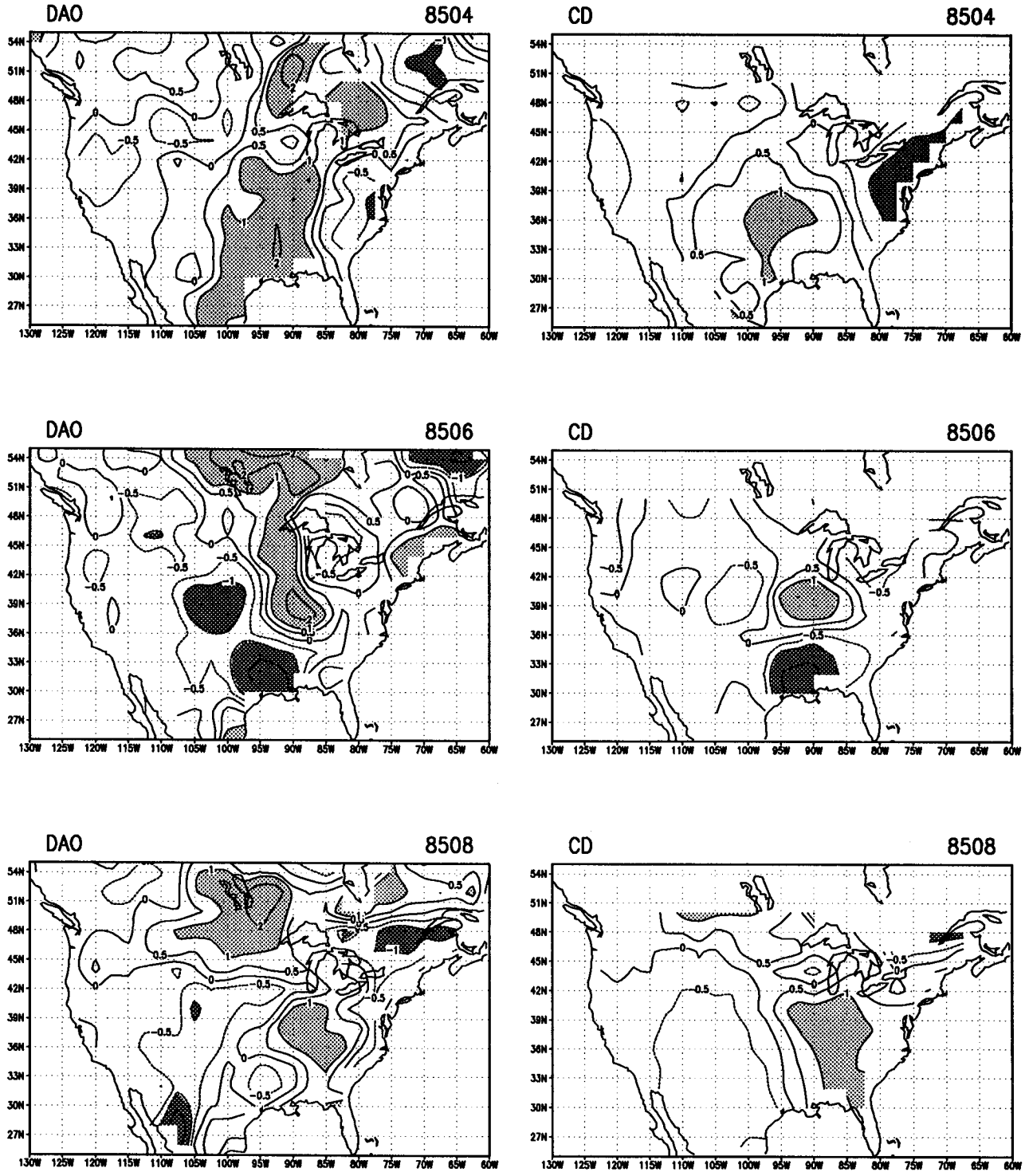


Figure 86: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1985. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

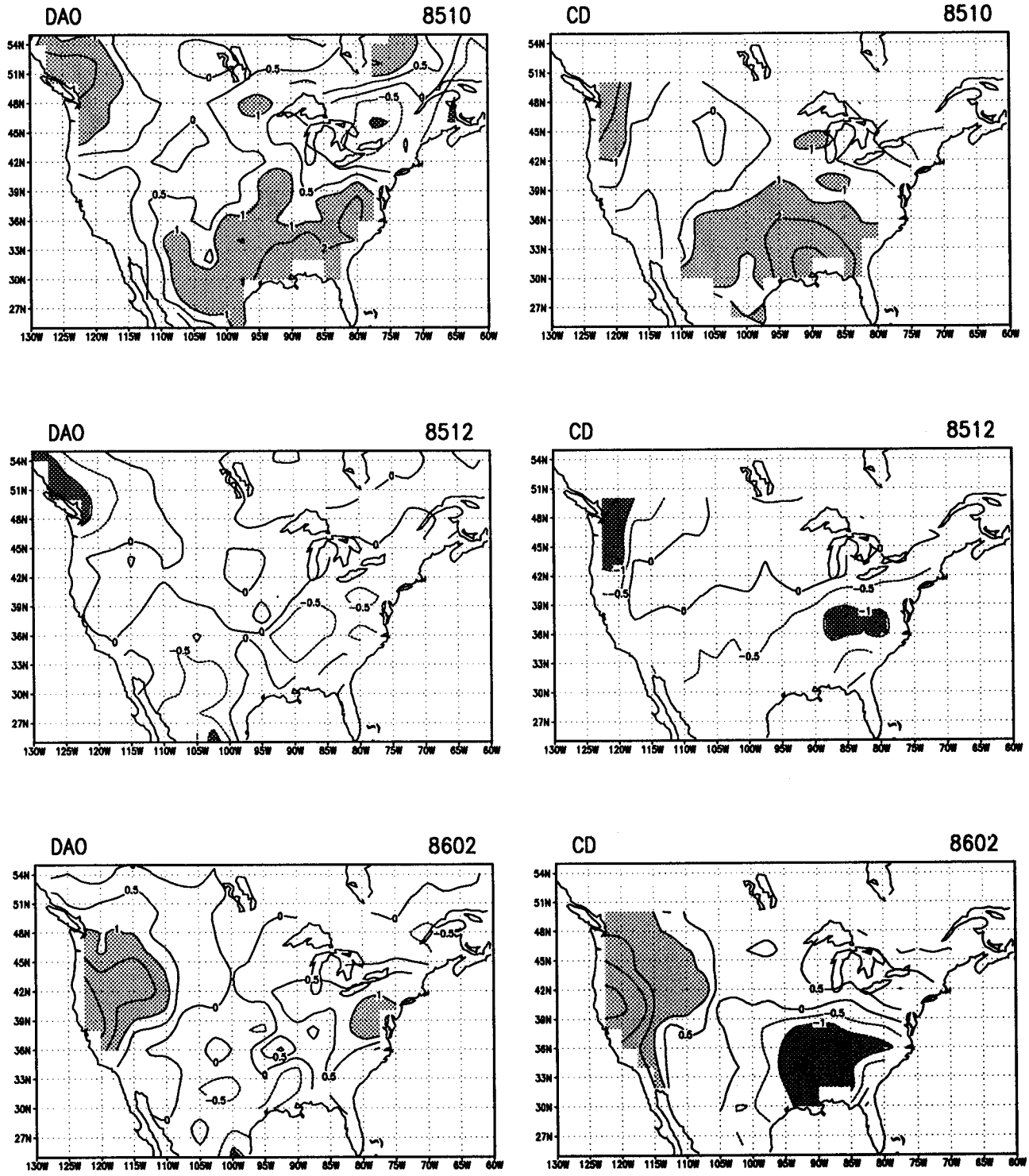


Figure 87: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1985, and February 1986. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

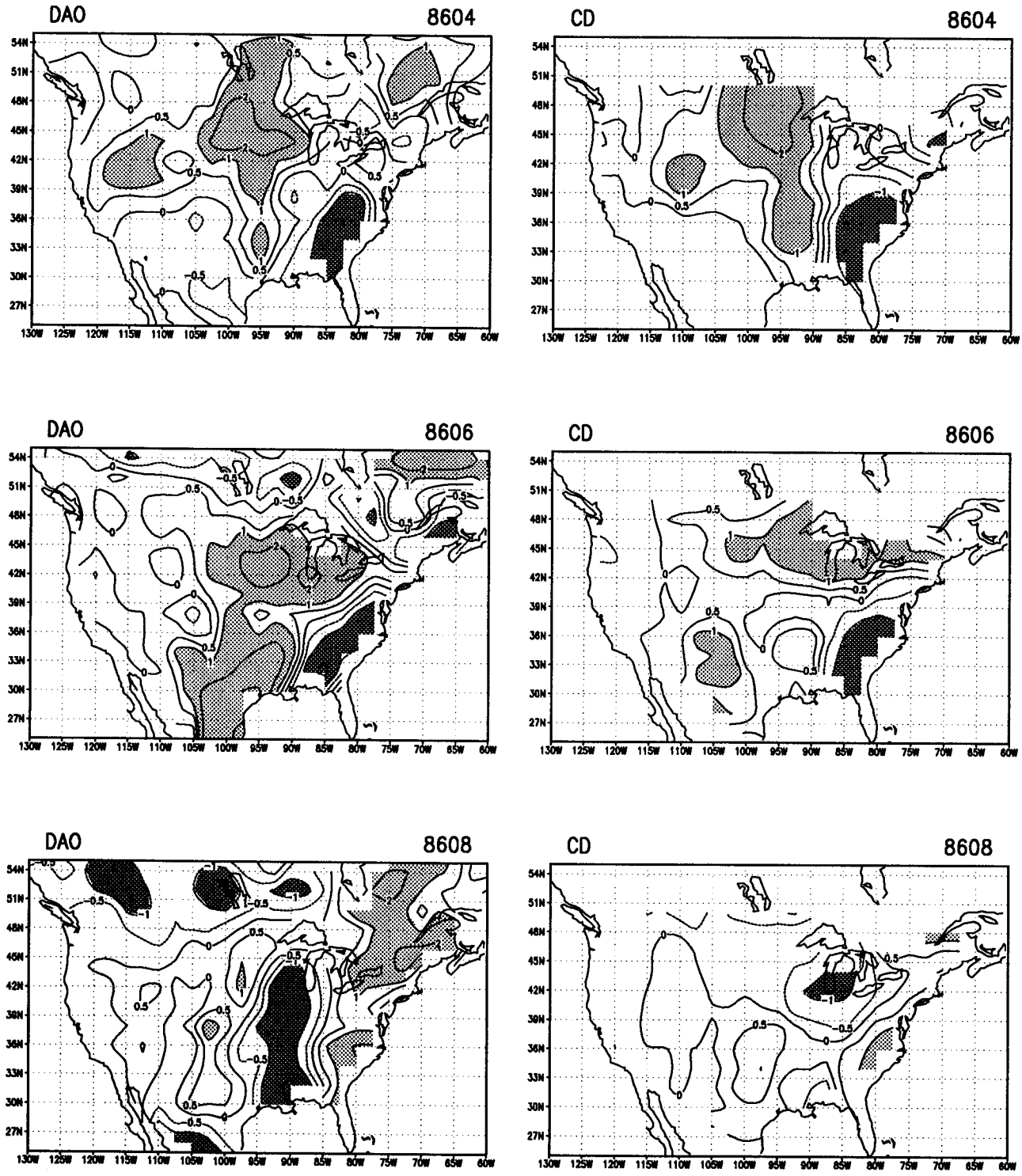


Figure 88: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1986. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

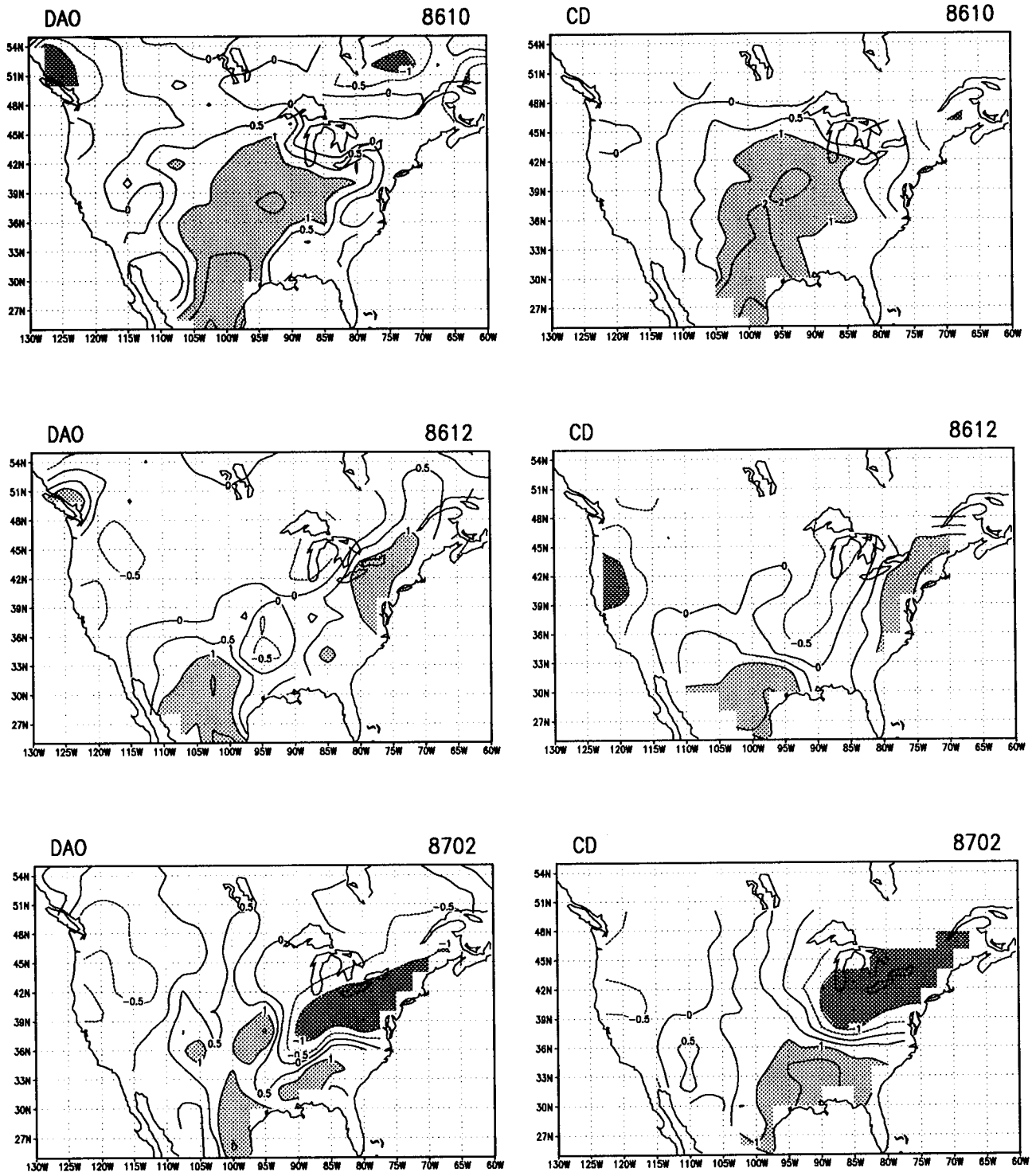


Figure 89: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1986, and February 1987. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

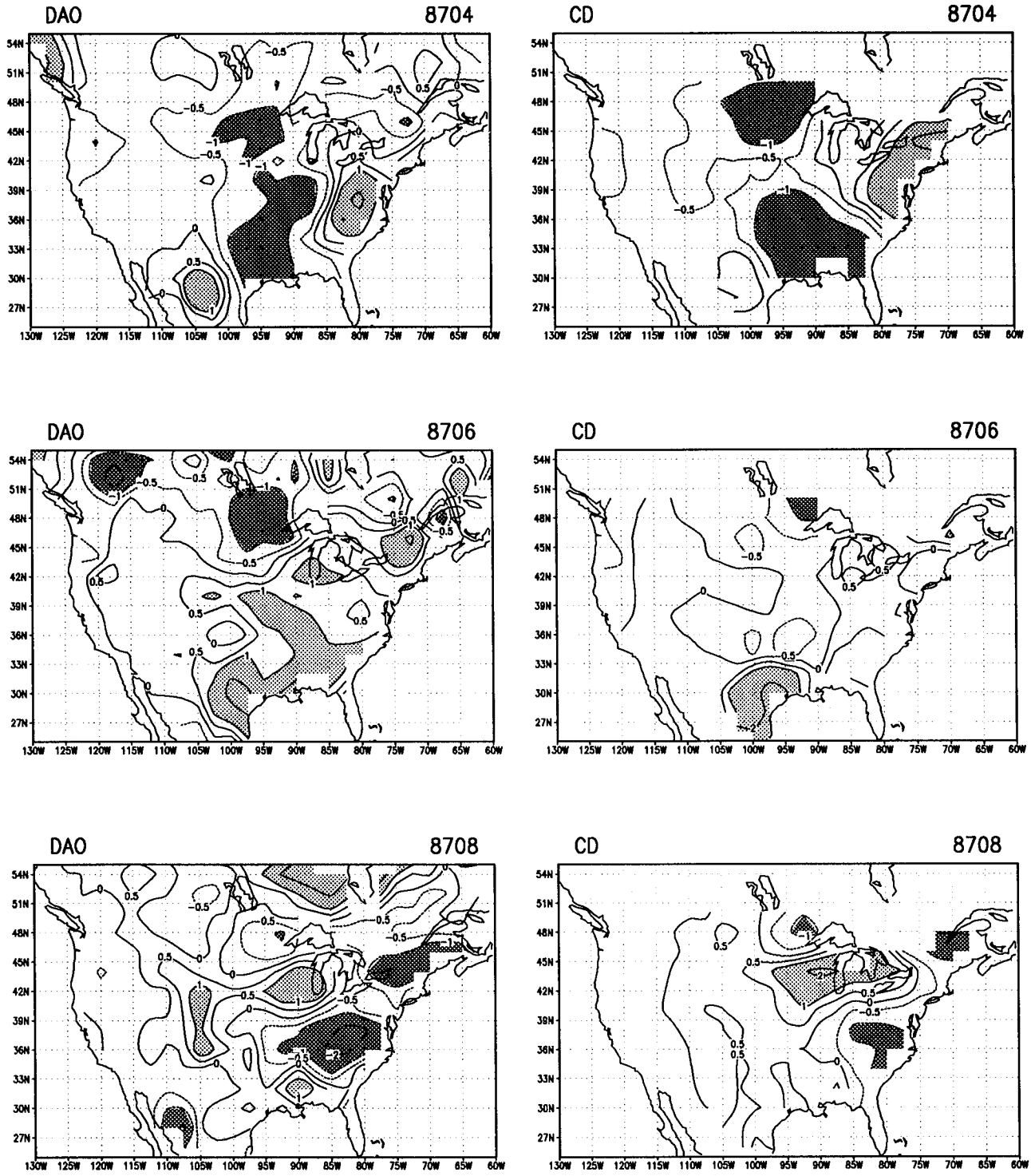


Figure 90: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1987. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

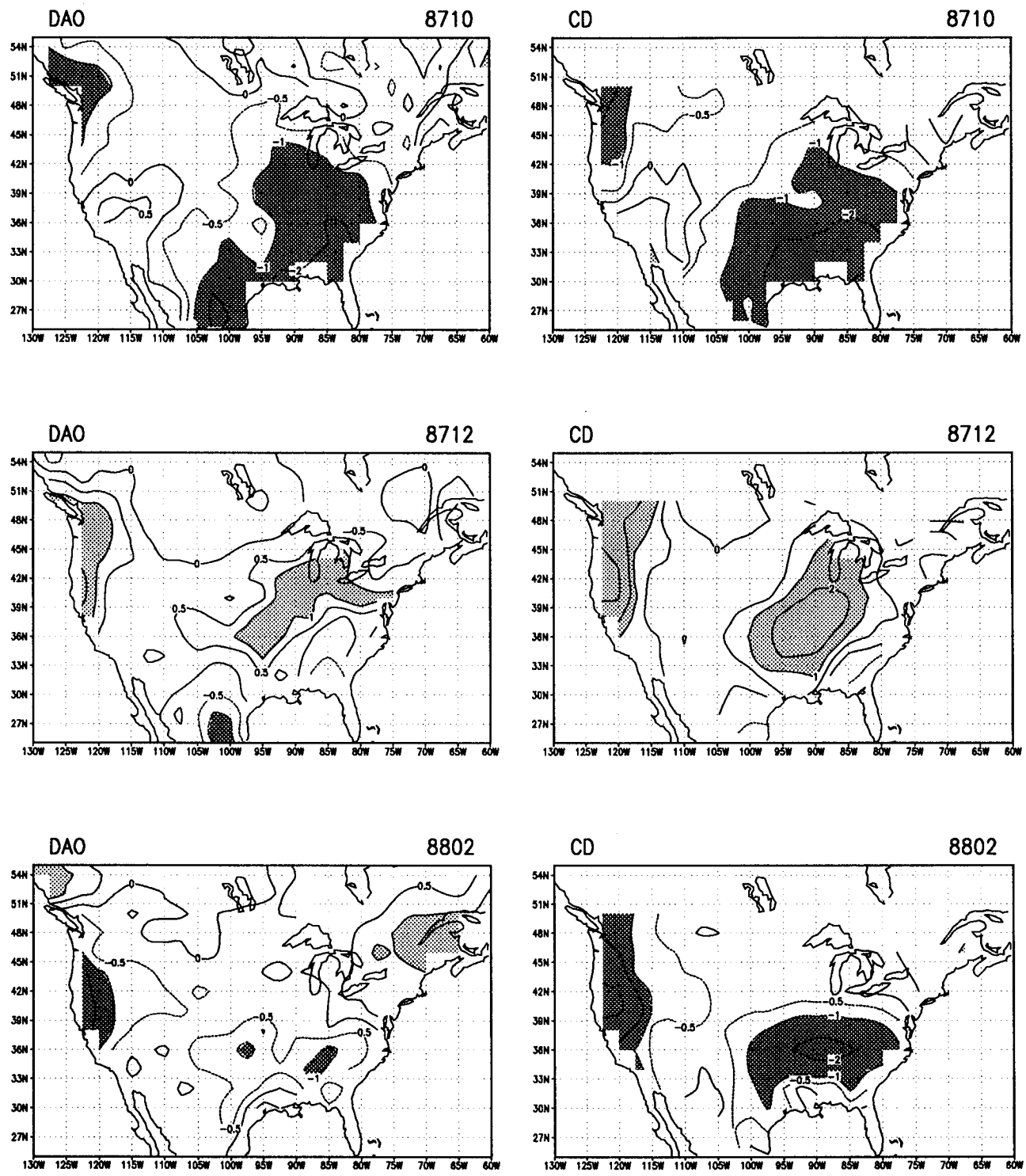


Figure 91: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1987, and February 1988. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

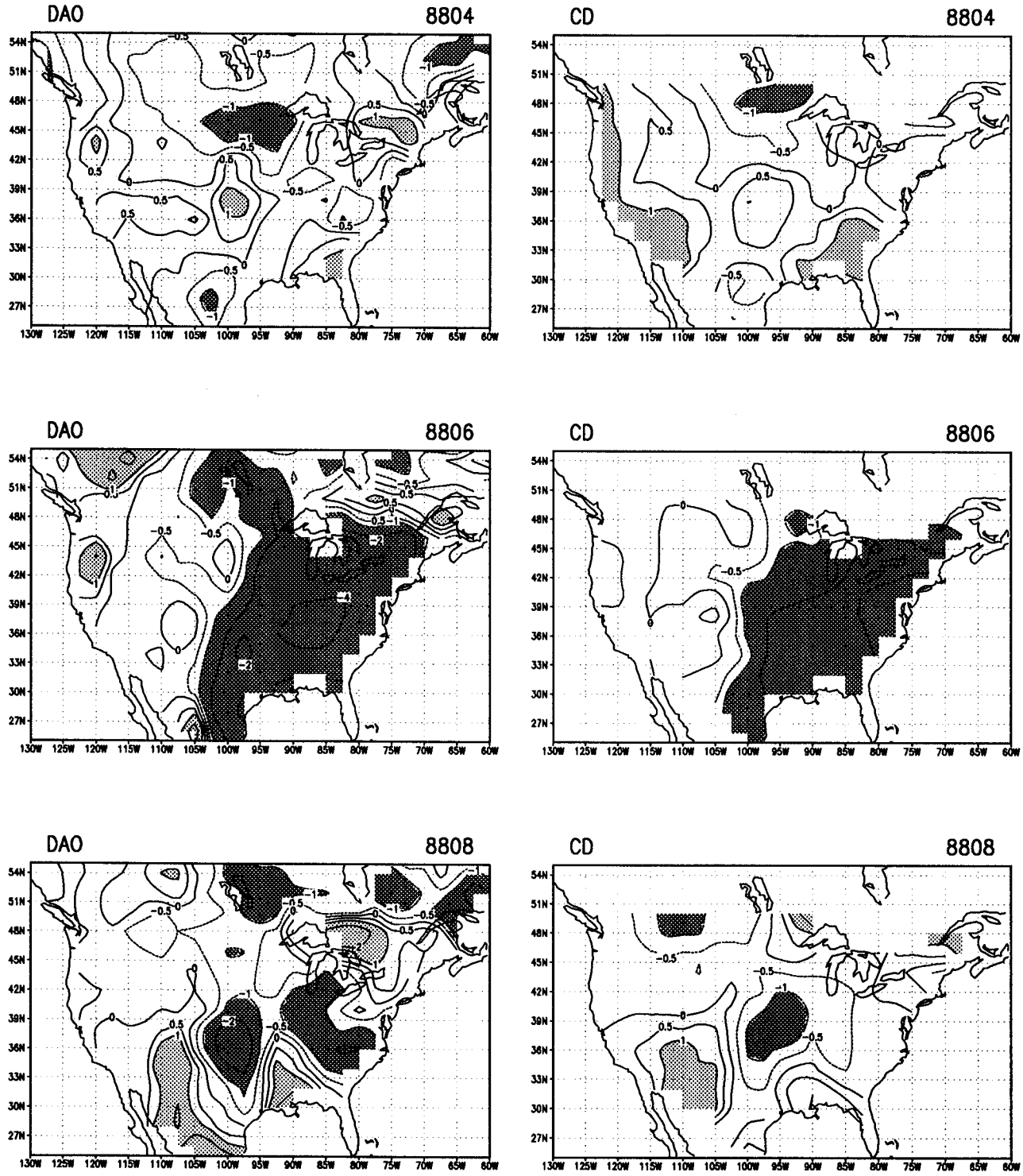


Figure 92: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1988. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

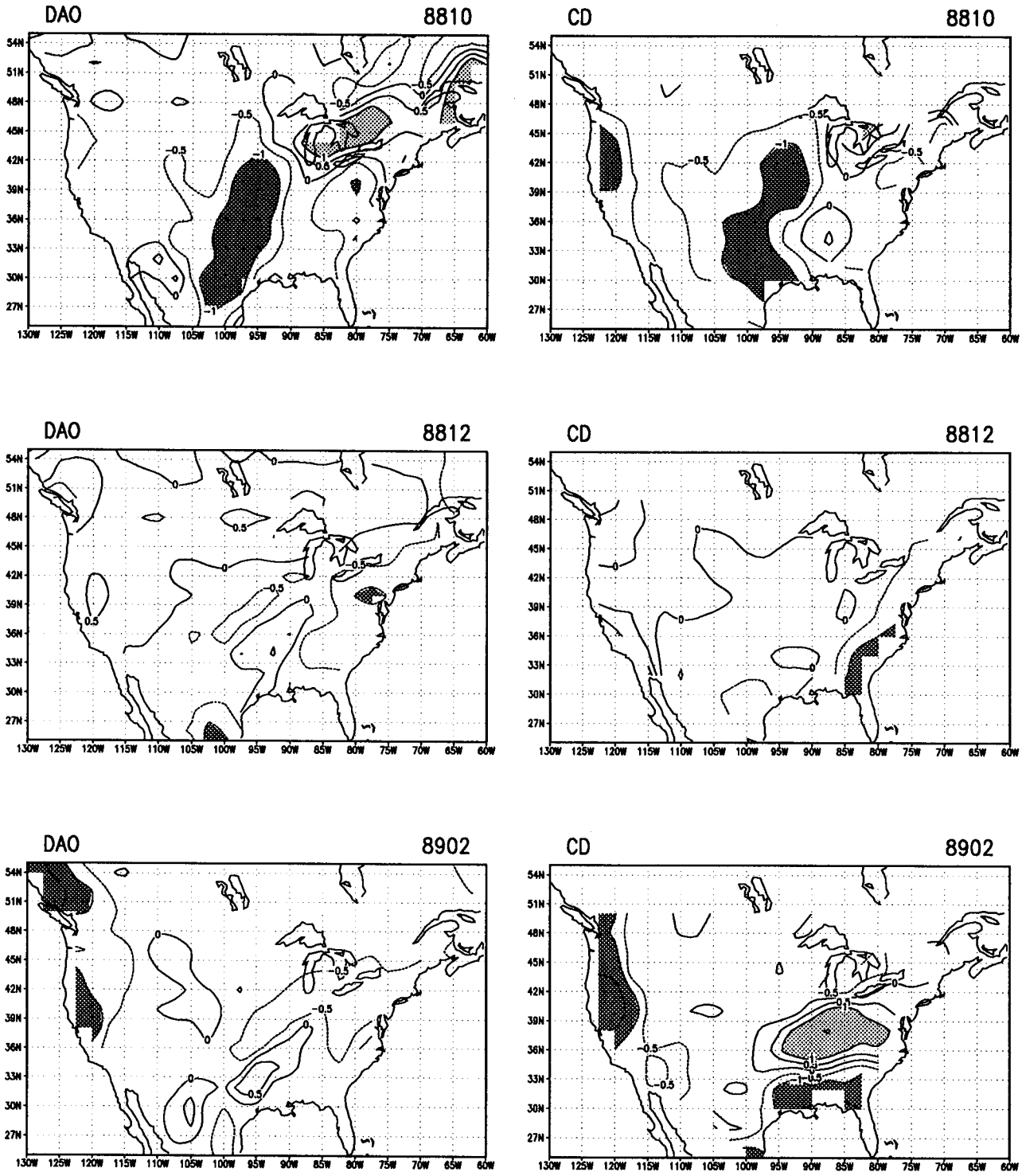


Figure 93: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1988, and February 1989. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

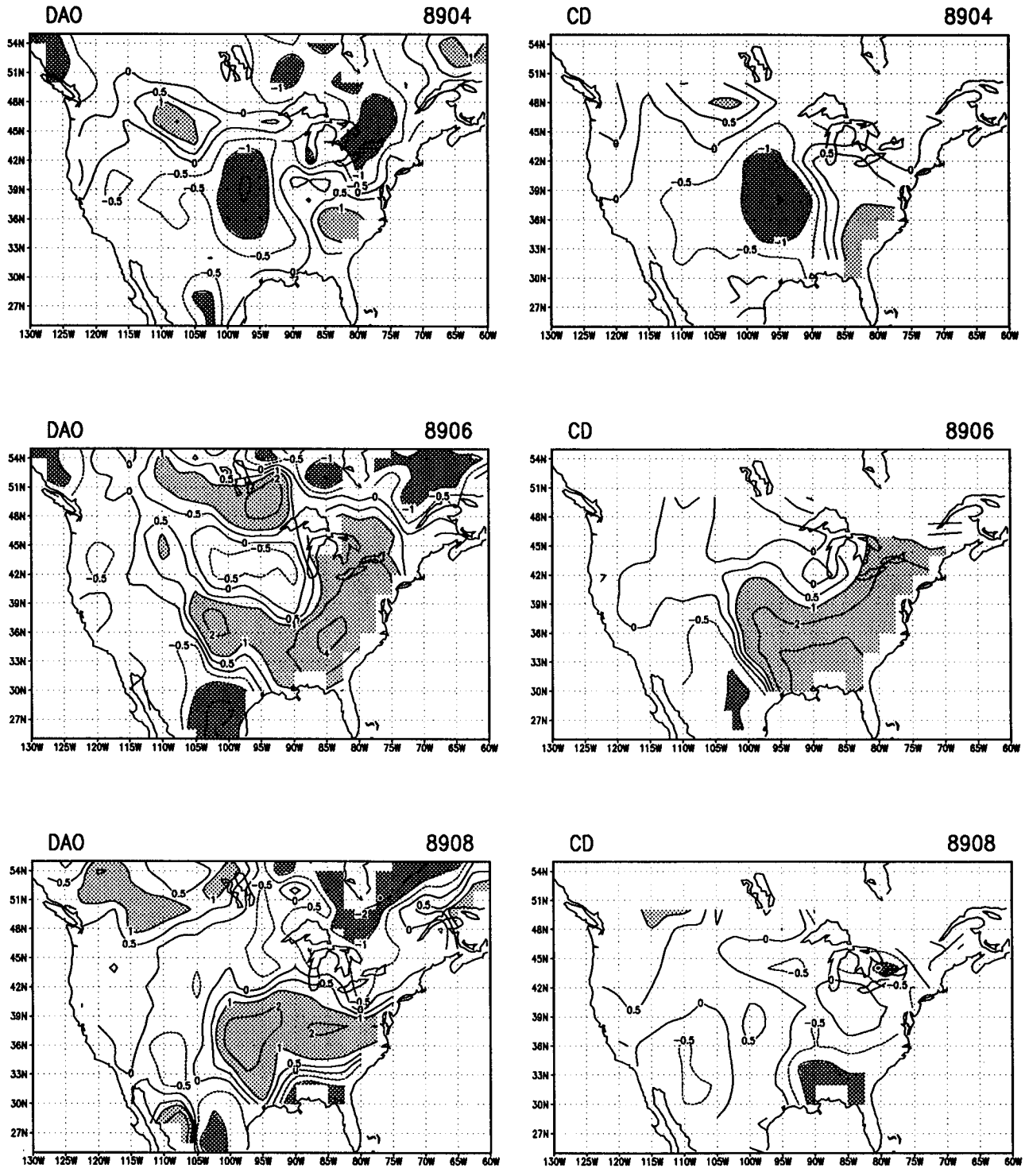


Figure 94: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for April, June and August of 1989. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

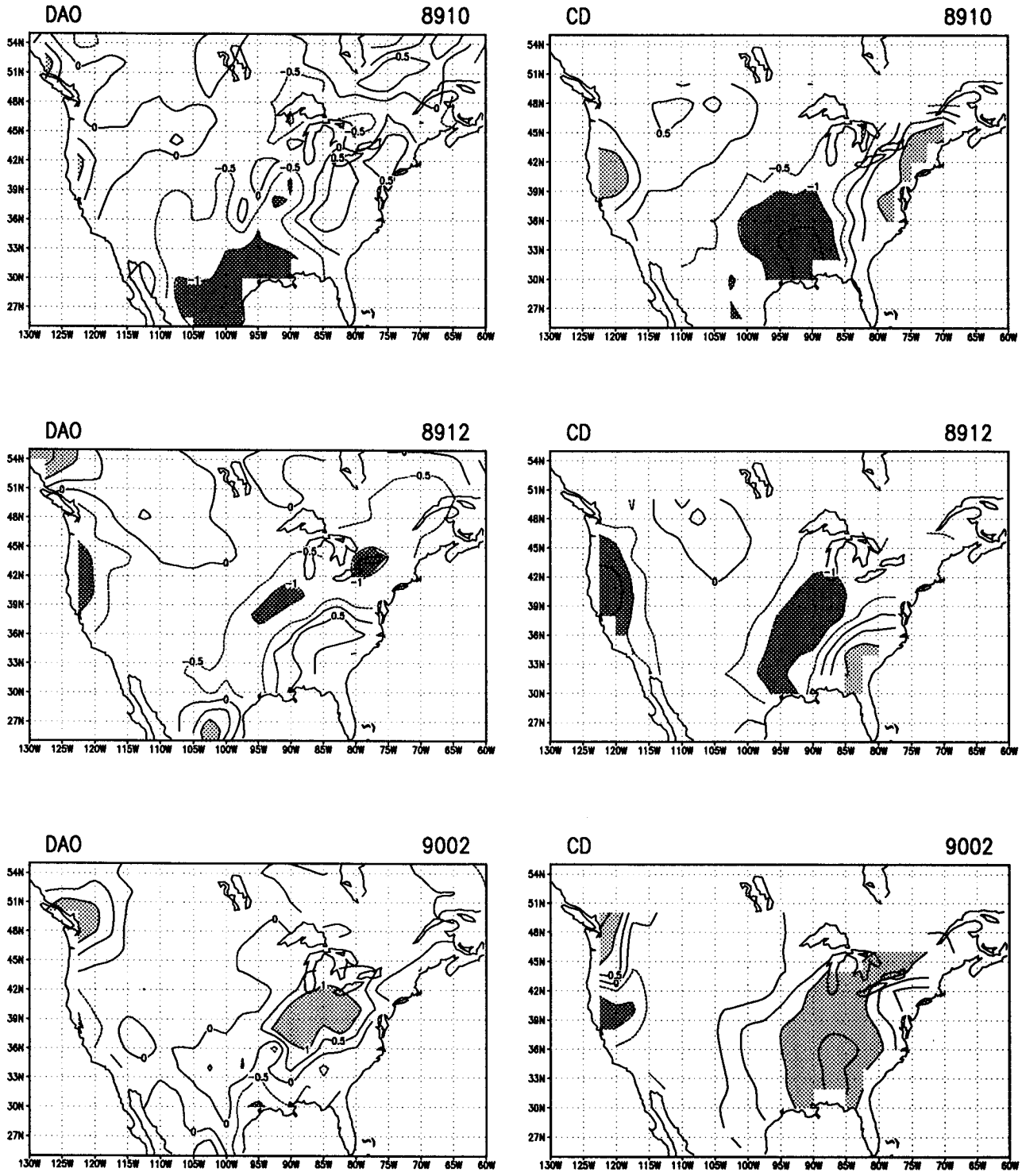


Figure 95: The precipitation anomalies from the GEOS–DAS compared with the climate division precipitation for October, December of 1989, and February 1990. Contour intervals are -10 -8 -6, -4, -2, -1, -0.5, 0, 0.5, 1, 2, 4, 6, 8, 10 mm/day. Absolute values greater than 2 are shaded.

4.5 Concluding Remarks

The results presented in this document are intended to give a general overview of the quality and consistency of the GEOS–1 DAS output. The results show that many of the time–mean prognostic fields are quite similar to those of the ECMWF analyses, though substantial systematic differences do occur in some regions. It is unclear how many of these differences are the result of model bias, or whether they reflect differences in the input observations, or how closely each system draws to the same observations. Other quantities, such as the zonal mean meridional wind, show substantial differences. These quantities are only weakly constrained by the observations and appear to reflect model biases. The impact of model bias is clearest in the hydrological cycle. This is reflected in the list of known deficiencies (summarized in Fig. 96, see also the Appendix), which are dominated by biases in the moisture field, clouds, and precipitation.

Perhaps the most promising new products provided by the current assimilation are the estimates of interannual variability. This includes both the prognostic fields and diagnostic fields such as precipitation and latent heating. It is for these products that the use of an unvarying assimilation system has clearly resulted in major improvements over operational products, though it must be kept in mind that any inhomogeneities in the input observations can also lead to spurious signals. The El Niño and La Niña events, the variability in the Indian summer monsoon, and the droughts are clearly evident in the assimilated fields during this period (1985–90). The availability of consistent forcing anomalies (e.g. precipitation) associated with these events should make this data set an extremely valuable resource for studying short–term climate variabilty.

Time scale	Quantity	Remarks
Time mean	T_s over land	Too cold during winter Weakly constrained by observations. Model bias
Winter mean	precipitation over U.S.	underestimate
Summer mean	precipitation over U.S.	overestimate
Time mean	lower troposphere q in tropics	too dry, suggest need for down drafts
Time mean	upper troposphere q over oceans	too wet, improve/tune convective scheme, reduce contention with analysis, analyze RH
Time mean	tropical clouds	Over estimate cloud forcing Feedback from analysis Improve cloud optical properties and cloud amount Reduce contention between convection scheme and analysis
Time mean	middle latitude clouds	underestimate cloud forcing improve cloud optical properties and cloud amount
Diurnal	precipitation over U.S.	Diurnal cycle too strong, improve interaction between convection scheme and PBL, currently no land surface model
model time step	noise in divergence and vertical motion near poles	also evident in other quantities, new version of GCM allows pole rotation
?	winds at equator	excessive divergence due to problem with OI wind-wind correlations (probably confined to high frequency, also appears to have a mean component)

Figure 96: A summary of known deficiencies in the GEOS–1 DAS output based on the analysis of the first 5 years (March 1985–February 1990).

5 Input Observations

This section summarizes the basic characteristics (spatial and temporal coverage) of the observations that were used as input to the assimilation. Details about the analysis scheme, including the quality control of the observations are described in Pfaendtner et al. (1995).

5.1 Spatial Coverage

An example of the spatial distribution of the various data sources for May 5th, 00Z is given in Figures 97–100. The GEOS-1 DAS incorporates the rawinsonde reports (upper panel Fig. 97) of horizontal winds, temperature (converted to heights), and dew point temperature (converted to mixing ratio). Horizontal winds are also obtained from a number of special wind observations (lower panel Fig. 97), commercial aircraft reports (lower panel Fig. 99), and satellite cloud-tracks (upper panel Fig. 100). Sea–level pressure information is incorporated from land surface (upper panel Fig. 98), ship (lower panel Fig. 98), and buoy (upper panel Fig. 99) observations. Surface winds are obtained from the ship and buoy reports. An example of the coverage of the NESDIS temperature retrievals (also converted to heights in the analysis) is shown in the lower panel of Fig. 100.

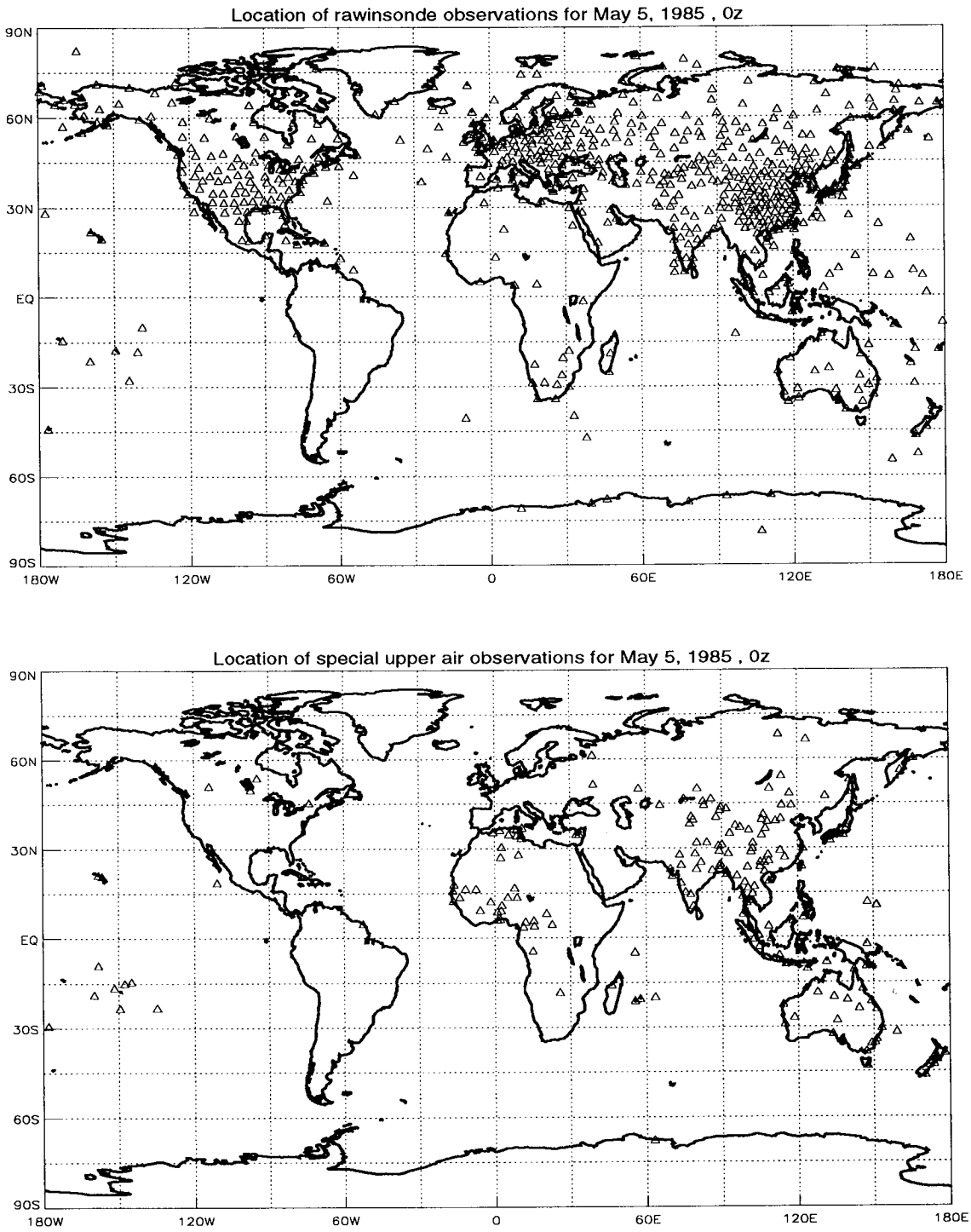


Figure 97: Locations of rawinsonde (upper panel) and special upper air (lower panel) observations for May 5th, 00Z. The latter include pilot balloon, dropwinsonde and reconnaissance aircraft and radar winds.

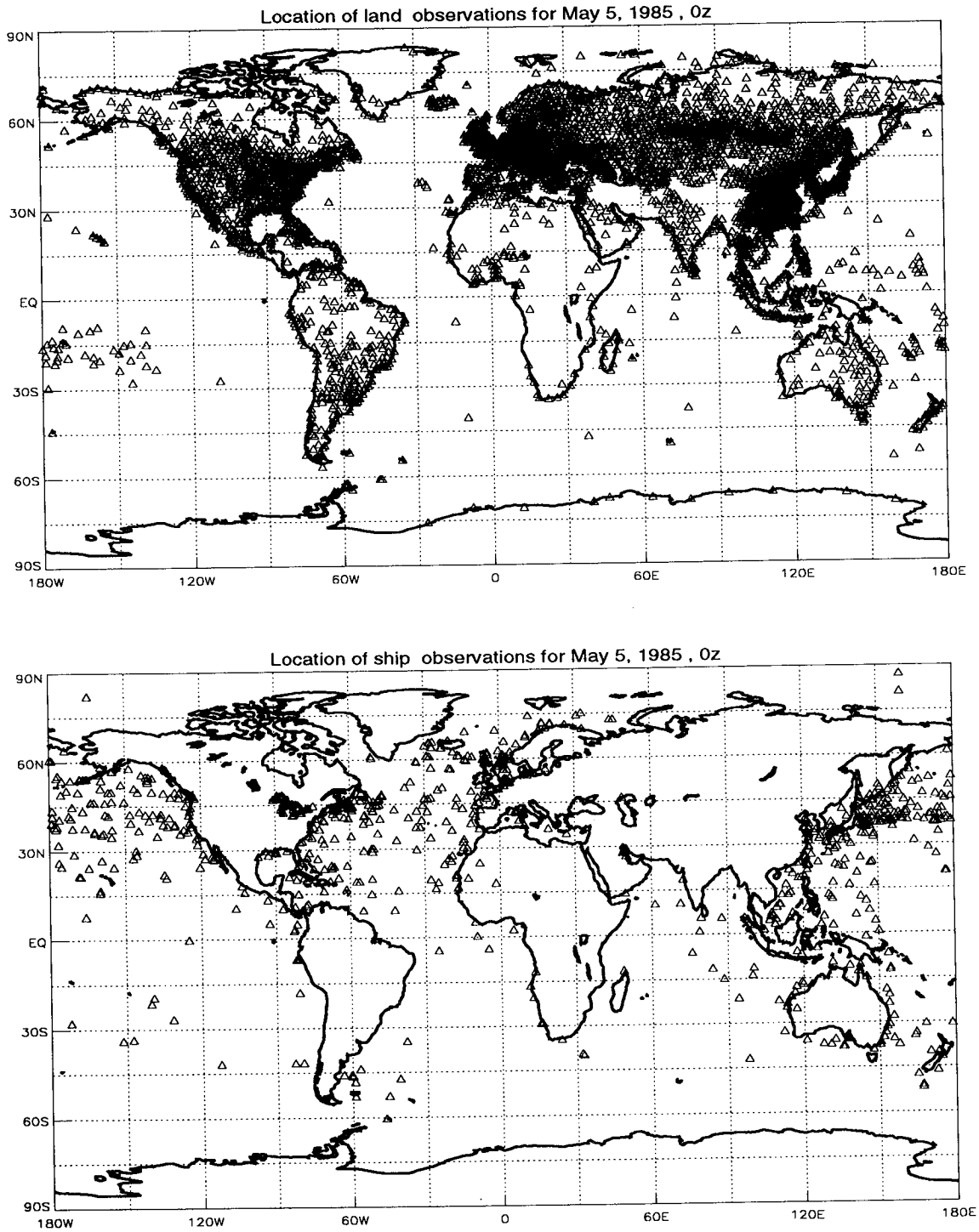


Figure 98: Locations of land (upper panel) and ship (lower panel) reports for May 5th, 00Z.

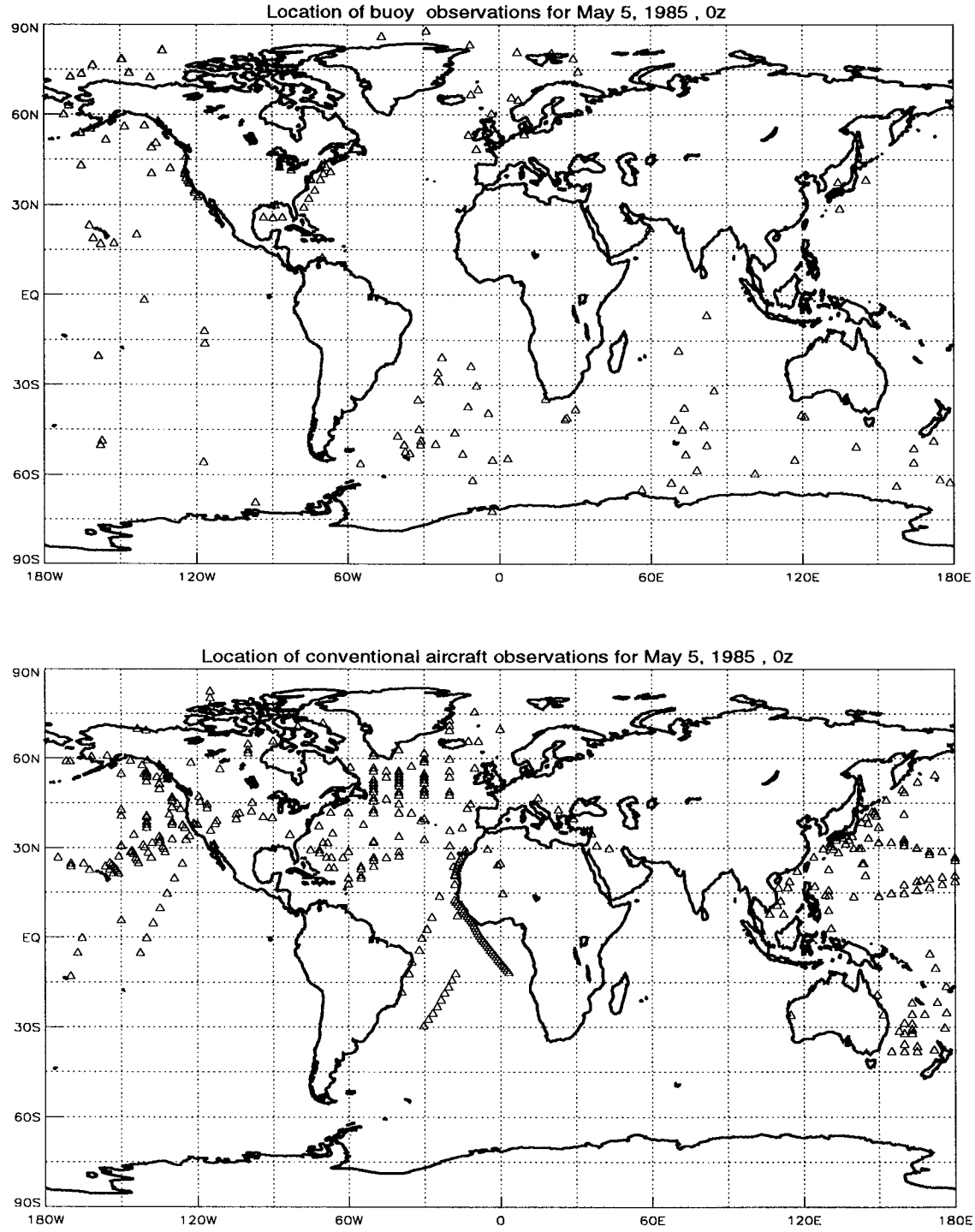


Figure 99: Locations of buoy (upper panel) and aircraft (lower panel) reports for May 5th, 00Z.

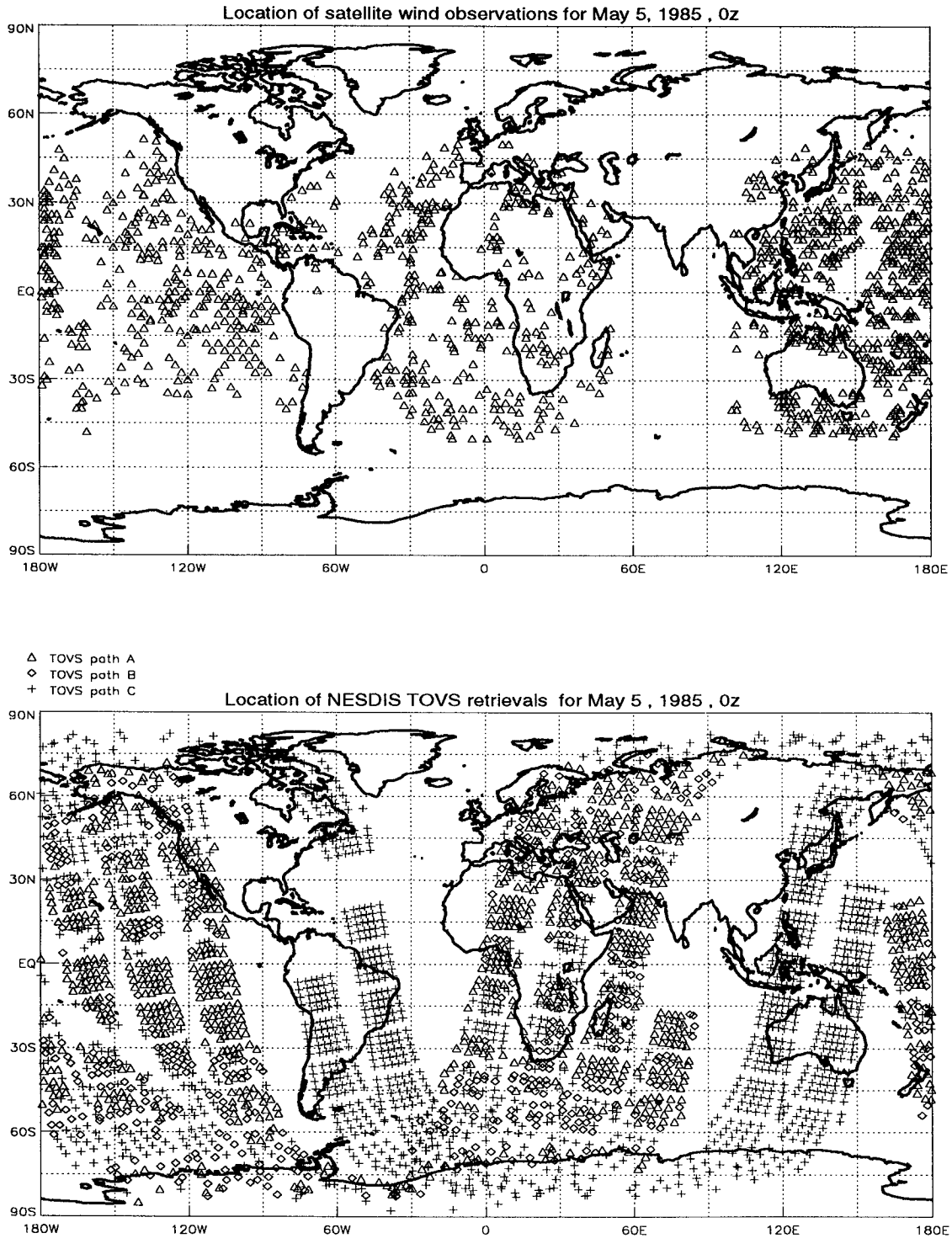


Figure 100: Locations of satellite wind (upper panel) and NESDIS temperature retrievals (lower panel) under clear (path A), partly cloudy (path B) and cloudy (path C) conditions for May 5th, 00Z.

5.2 Inventory

This section lists (in the form of tables) all significant gaps in the DAO input observational database for the period (1985–1990). These represent periods for which the tapes received by the DAO contained no data and initial efforts to fill these gaps from other sources were unsuccessful. Data for times prior to July 1987 were mostly acquired from the National Center for Atmospheric Research (NCAR). For later times, the data have been obtained directly from NMC and do not include data that came in after the cut–off time for the operational NMC system. In addition to these two sources, some TOVS temperature soundings have come directly from the National Oceanic and Atmospheric Administration (NOAA) National Environmental Satellite, Data and Information Service (NESDIS), and gaps have been filled with data from the National Climatic Data Center (NCDC) in Ashville, NC. For the global sea–level pressure and near surface wind analysis over the oceans, data from surface land synoptic reports (sea–level pressure only), ships and buoys are used. The upper–air analyses of height, wind, and moisture incorporate the data from rawinsondes, dropwindsondes, rocketsondes, aircraft winds, cloud tracked winds, and thicknesses from the historical TOVS soundings produced by NOAA NESDIS.

The actual counts for each of the observational types is also available in the form of bargraphs, via anonymous ftp to the machine: hera.gsfc.nasa.gov

The directory is: pub/assimilation/e0054A/observations

5.2.1 Sample Bargraphs of Data Counts

An example of data counts for May, 1985 is given in Figs. 101–104.

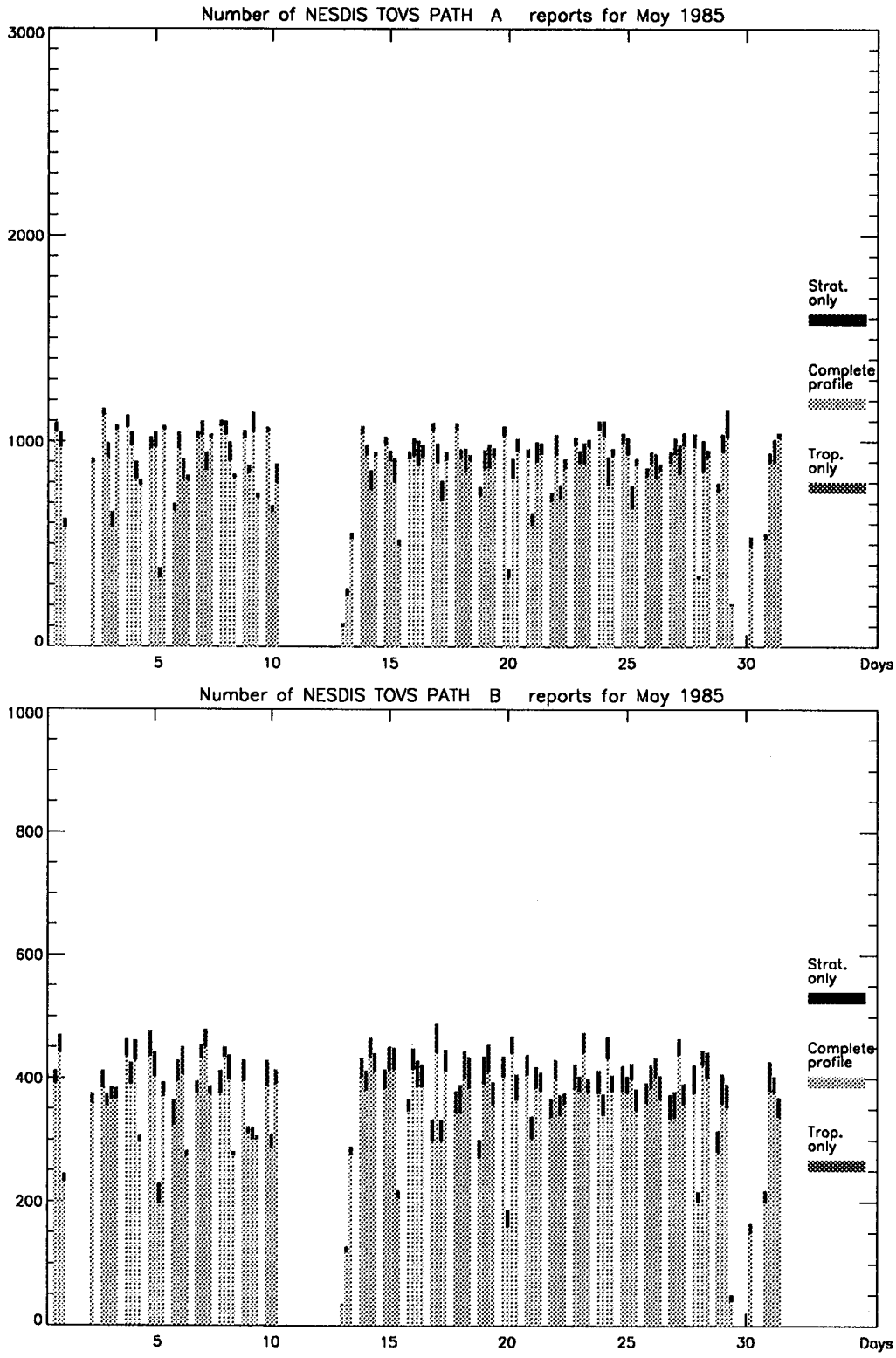


Figure 101: Number of TOVS clear (path A, upper panel) and TOVS partly cloudy (path B, lower panel) reports for May 1985. Stratospheric only soundings are defined as occurring above the 100 mb level.

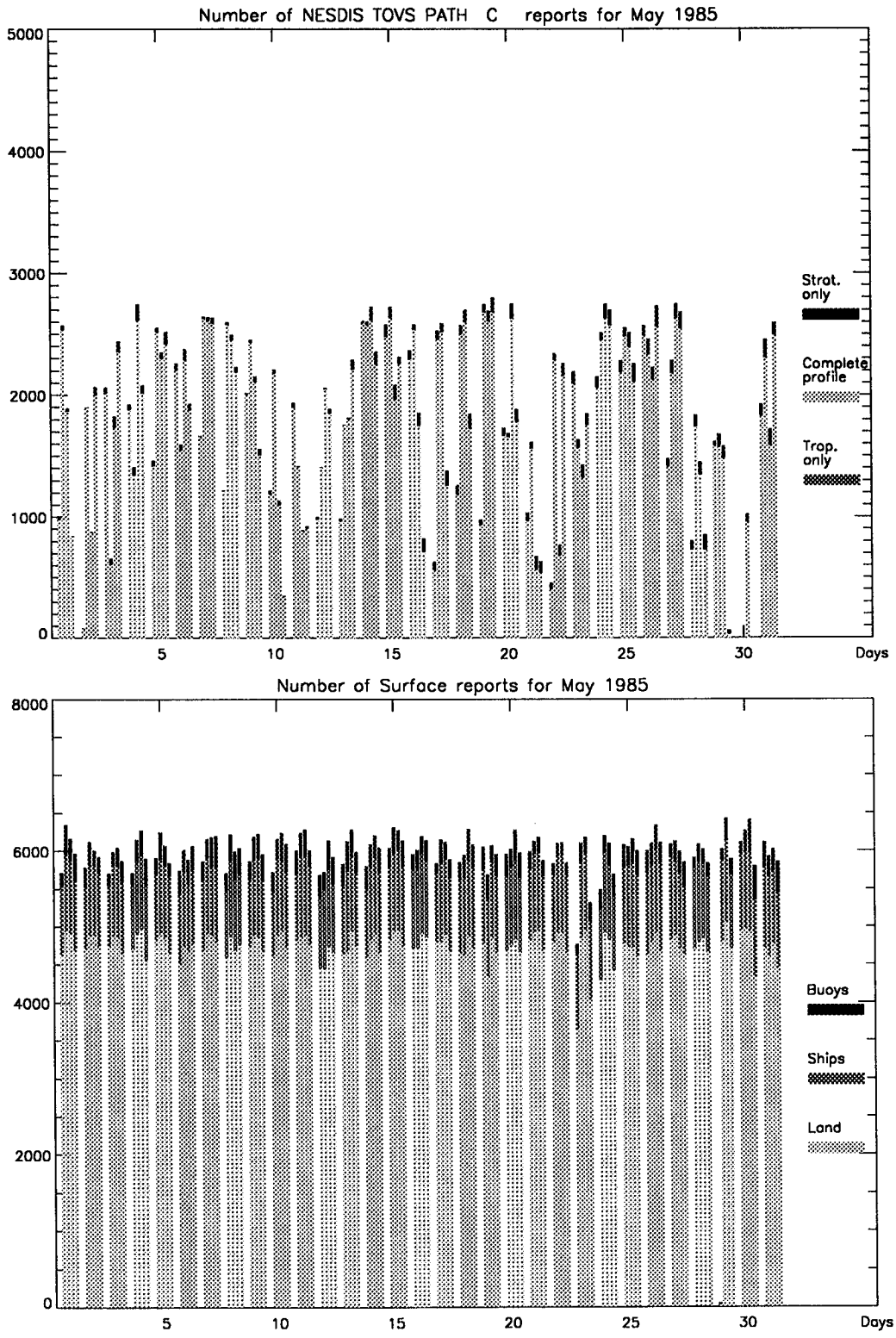


Figure 102: Number of TOVS cloudy (path C, upper panel) and surface reports (lower panel) for May 1985. In (a) stratospheric only soundings are defined as occurring above the 100 mb level.

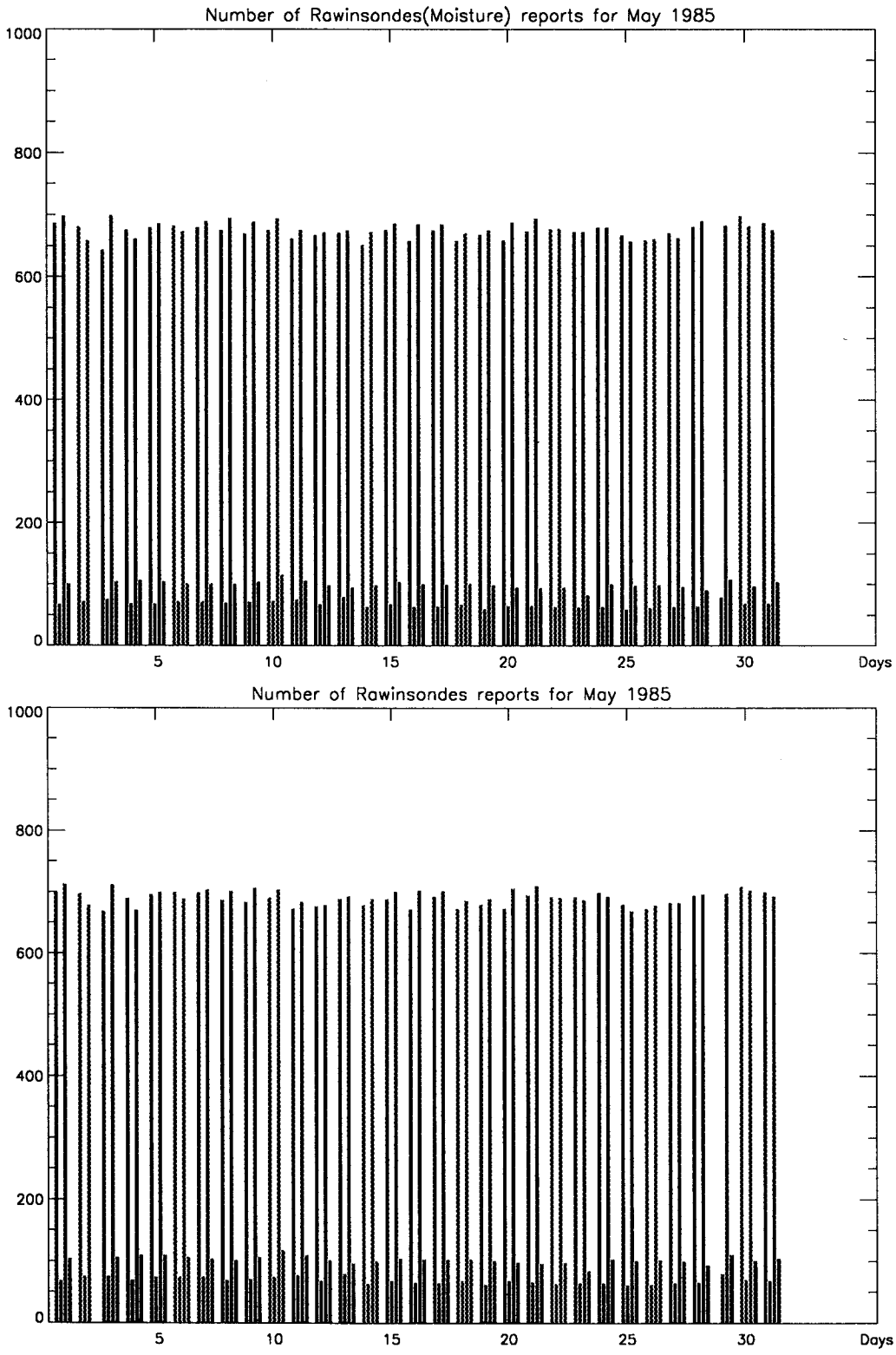


Figure 103: Number of rawinsonde (moisture only, upper panel) and rawinsonde (all, lower panel) reports for May 1985.

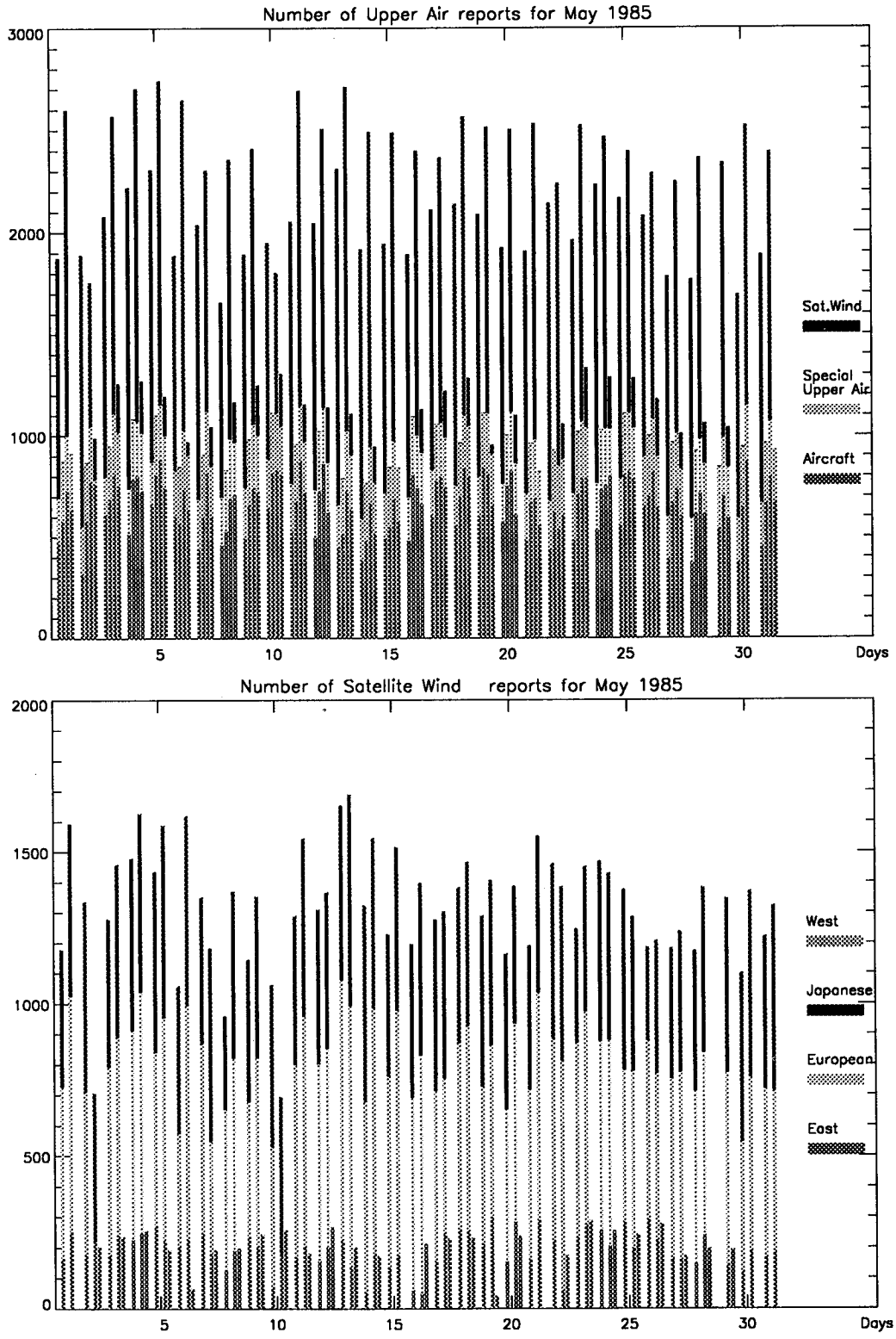


Figure 104: Number of upper air reports (upper panel), and satellite wind reports (lower panel) for May 1985.

5.2.2 Tables of Significant Data Gaps

1985	
TIME PERIOD	MISSING DATA
January 4 (18Z)	About 1/2 Land data
February 6 - February 25	No A and B TOVS data
March 16 (18Z)	No TOVS data
March 17	No TOVS data
April 1 (0Z)	About 1/2 Land data
April 3 (18Z)	No TOVS data
April 4 (0Z)	No TOVS data
May 2 (18Z)	No Rawins,No TOVS data
May 4 (0Z-12Z)	No TOVS data
May 10 (18Z)	No TOVS data
May 11 - May 12	No TOVS data
May 13 (0Z)	No TOVS data
May 29 (0Z)	Almost No Land,Ships and Buoys, No Rawins,Aircrafts, No Upper Air,No Sat.Wind
May 30 (18Z)	No Aircrafts,Upper Air,Sat.Wind
May 30 (0Z,6Z, 18Z)	No TOVS data
July 16 (0Z)	No Land,Ships and Buoys, No Rawins,Aircraftss,Upper Air,Sat.Wind
July 20 (12Z)	About 1/2 Land data
July 29 (18Z)	No TOVS data
July 30 (0Z)	No TOVS data
August 11 (6Z) - August 17	No TOVS
September 27 (6Z)	No TOVS
October 3 (12Z)	About 1/2 Land data
October 9 (18Z)	Almost No Land data

1986	
TIME PERIOD	MISSING DATA
January 2 (18Z)	No TOVS data
January 3 (0Z-12Z)	No TOVS data
February 13 (18Z)	No Sat.Wind
February 24 (6Z)	No Sat.Wind
February 25 (18Z)	Almost No Land
February 27 (18Z)	No Rawins,Upper Air,Sat.Wind
March 14 (6Z-18Z)	No TOVS data
March 15	No TOVS data
April 14 (0Z)	Almost No TOVS data (only 6 reports)
April 15 (18Z)	No TOVS data
April 21 (6Z)	No TOVS data
April 27 (18Z)	Almost No Rawins
April 30 (6Z)	No TOVS data
May 4 (0Z)	No Land, Ships, Buoys (almost), No Rawins, Aircrafts,Sat.Wind, almost no Upper Air
May 11 (0Z)	No TOVS data
May 23 (18Z)	No Sat.Wind
June 2 (6Z-18Z)	No TOVS data
June 3 (0Z-6Z)	No TOVS data
June 5 - 8	No TOVS data
June 11 (6Z-18Z)	No TOVS data
June 12 (0Z-12Z)	No TOVS data
June 14 (18Z)	No Land,No Ships,No Buoys data
June 14 (18Z)	No Rawins data
June 14 (18Z)	No Sat.Wind data
June 16 (6Z-18Z)	No TOVS data
June 17 (0Z-6Z)	No TOVS data
June 20	No Aircrafts data
June 21 (18Z)	No Land,No Ships,No Buoys data
June 21 (18Z)	No Rawins data
June 21 (18Z)	No Sat.Wind data
June 22 (0Z)	No TOVS data
June 26 (18Z)	No TOVS data

1986	
TIME PERIOD	MISSING DATA
July 19 (18Z)	No Sat.Wind
August 1 - August 5(6Z)	No TOVS data
August 11 (6Z - 12Z)	No TOVS data
August 21 (18Z) - August 23	No TOVS data
August 25 (6Z) - August 30(12Z)	No TOVS data
September 1 (6Z-18Z)	No TOVS data
September 2 - September 3	No TOVS data
September 7 (6Z)	No TOVS data
September 12	No TOVS data
September 13 (0Z - 12Z)	No TOVS data
September 14 (6Z-18Z)	No TOVS data
September 21 (6Z-18Z)	No TOVS data
September 25 (12Z-18Z)	No TOVS data
September 26 - 28	No TOVS data
October 2 (18Z)	No Sat.Wind data
October 4	Almost No Buoys data
October 13	No TOVS data
October 14 (0Z)	No TOVS data
November 2 - 8	No AIRCRAFT data
November 2 - 8	No Sat.Wind data
November 3 (6Z)	No TOVS data
November 10 (6Z, 12Z)	No TOVS data
November 17 (6Z, 12Z)	No TOVS data
November 24 (6Z)	No TOVS data
December 1 (18Z)	No Sat.Wind data
December 1 (6Z)	No TOVS data
December 8 (6Z)	No TOVS data
December 15 (6Z, 12Z)	No TOVS data
December 22 (6Z, 12Z)	No TOVS data
December 29 (6Z)	No TOVS data

1987	
TIME PERIOD	MISSING DATA
January 2 (6Z)	No TOVS data
January 5 (6Z- 12Z)	No TOVS data
January 9 (6Z- 18Z)	No TOVS data
January 12 (6Z)	No TOVS data
January 19 (6Z- 18Z)	No TOVS data
January 20 (0Z- 12Z)	No TOVS data
January 26 (6Z- 18Z)	No TOVS data
February 3 (6Z)	No TOVS data
February 10 (6Z)	No TOVS data
February 17 (6Z)	No Ships and Buoys data
February 17 (6Z)	No TOVS data
March 3 (6Z- 18Z)	No TOVS data
March 4	No TOVS data
March 8	No TOVS data
March 9 (0Z- 6Z)	No TOVS data
March 12 (6Z)	No TOVS data
March 17 (6Z)	No TOVS data
March 24 (6Z)	No TOVS data
March 31 (6Z)	No TOVS data
April 7 (6Z- 18Z)	No TOVS data
April 22 (0Z)	No Sat.Wind data
April 27 (6Z-12Z)	No TOVS data
April 28 (6Z-18Z)	No TOVS data
May 2 (6Z-18Z)	No TOVS data
May 4 (6Z-12Z)	No TOVS data
May 11 (6Z-12Z)	No TOVS data
May 24 (6Z)	No TOVS data
May 30 (6Z-18Z)	No Land ,Ships data
May 30 (6Z)	No TOVS data
June 20 (18Z)	No TOVS data
June 21 (0Z-6Z)	No TOVS data
June 28 (6Z)	No TOVS data
July 14 (6Z)	No Rawins data
July 18 (6Z)	No TOVS data
July 24 (6Z - 18Z)	No TOVS data
July 25 - July 26	No TOVS data
July 27 (12Z - 18Z)	No TOVS data
July 28 (6Z - 18Z)	No TOVS data
July 29 (0Z)	No TOVS data

1987	
TIME PERIOD	MISSING data
August 1 (6Z)	No Sat.Wind
August 1 (12Z)	No Aircrafts
August 1 (12Z- 18Z)	No Land,No Ships,No Buoys,No Rawins, No Sat.Wind,Upper Air
August 2 (6Z- 18Z)	No TOVS
August 3 (0Z- 6Z)	No TOVS
August 9 (6Z)	No TOVS
August 16 (6Z)	No TOVS
August 31 (6Z)	No Sat.Wind
August 31 (12Z)	No Land,No Sat.Wind,Upper Air
August 31 (18Z)	No Land,Ships,Buoys,Rawins, No Sat.Wind,Upper Air
September 6 (6Z-18Z)	No TOVS
September 7	No TOVS
September 12 (18Z)	No Sat.Wind
September 22 (6Z)	No TOVS
September 24	No TOVS
September 29 (0Z-6Z)	No TOVS
October 6 (6Z)	No TOVS
October 11 (6Z,18Z)	No Land,No Ships,No Buoys,No Rawins, No Sat.Wind
October 12 (6Z,18Z)	No Land,No Ships,No Buoys,No Rawins, No Sat.Wind
October 13 (6Z,18Z)	No Land,No Ships,No Buoys,No Rawins, No Sat.Wind
October 13 (6Z-12Z)	No TOVS
October 16 (18Z)	No TOVS
October 17 (18Z)	No Sat.Wind
October 20 (6Z)	No TOVS
October 27 (6Z)	No TOVS
November 3 (6Z)	No TOVS
November 10 (6Z)	No TOVS
November 17 (6Z)	No TOVS
November 18 (6Z-18Z)	No Rawins,No Aircrafts, No Upper Air, No Sat.Wind
November 24 (6Z)	No TOVS
December 1 (6Z)	No TOVS
December 8 (6Z-12Z)	No TOVS
December 15 (6Z)	No TOVS
December 23 (6Z)	No TOVS
December 29 (6Z)	No TOVS

1988	
TIME PERIOD	MISSING DATA
January 2 (6Z)	No TOVS data
January 5 (6Z- 12Z)	No TOVS data
January 12 (6Z- 12Z)	No TOVS data
January 19 (6Z- 12Z)	No TOVS data
January 26 (6Z)	No TOVS data
February 2 (6Z)	No TOVS data
February 9 (6Z- 12Z)	No TOVS data
February 13 (0Z- 6Z)	Almost No Rawins data
February 14 (18Z)	No Rawins data
February 15 (18Z)	No Land, No Ships, No Buoys data
February 15 (6Z- 18Z)	Almost No Rawins data
February 16 (6Z)	No TOVS data
February 23 (6Z)	No TOVS data
March 1 (6Z)	No TOVS data
March 8 (6Z-12Z)	No TOVS data
March 15 (6Z)	No TOVS data
March 22 (6Z)	No TOVS data
March 29 (6Z)-18Z	No TOVS data
March 30 (0Z- 6Z)	No TOVS data
April 5 (6Z-18Z)	No TOVS data
April 6 (0Z- 6Z)	No TOVS data
April 12 (6Z)	No TOVS data
April 19 (6Z- 12Z)	No TOVS data
April 26 (6Z-18Z)	No TOVS data
April 27 (0Z- 6Z)	No TOVS data
May 3 (6Z)	No TOVS data
May 5	No TOVS data
May 10 (6Z)	No TOVS data
May 18 (0Z- 6Z)	No TOVS data
May 24 (6Z)	No TOVS data

1988	
TIME PERIOD	MISSING DATA
June 10 (0Z- 6Z)	No TOVS data
June 21 (12Z-18Z)	No TOVS data
June 29 (0Z- 6Z)	No TOVS data
July 5 (6Z)	No TOVS data
July 11 (6Z)	No TOVS data
July 19 (6Z)	No TOVS data
July 26 (6Z-12Z)	No TOVS data
August 2 (6Z-18Z)	No TOVS data
August 3	No TOVS data
August 4 (0Z- 6Z)	No TOVS data
August 9 (6Z)	No TOVS data
August 16 (0Z- 6Z)	No TOVS data
August 23 (6Z)	No TOVS data
August 30 (6Z)	No TOVS data
September 6 (6Z)	No TOVS data
September 13 (6Z)	No TOVS data
September 20 (6Z)	No TOVS data
September 27 (6Z)	No TOVS data
October 4 (6Z)	No TOVS data
October 5	No TOVS data
October 11 (6Z)	No TOVS data
October 18 (0Z)	No TOVS data
October 25 (6Z)	No TOVS data
October 30 (0Z)	No TOVS data
November 1 (6Z)	No TOVS data
November 8 (6Z)	No TOVS data
November 15 (0Z)	No TOVS data
November 22 (6Z)	No TOVS data
November 29 (6Z)	No TOVS data

1989

TIME PERIOD	MISSING DATA
January 29 (18Z)	Almost No TOVS data
February 27	No Land
March 18 (18Z)	about 1/2 Land
APRIL 16 (12Z)	about 1/2 Land
April 21	No Ships, No Buoys
May 3 (18Z)	No Ships, No Buoys
June 28 (6Z, 18Z)	Almost No Rawins (Moisture)
July 26 (18Z)	about 1/2 Land
July 27 (6Z- 12Z)	about 1/2 Land
August 2 (6Z)	No Rawins data
August 3 (6Z)	No Rawins data
August 4 (6Z)	No Rawins data
August 5 (6Z)	No Rawins data
August 6 (6Z)	No Rawins data
August 7 (6Z)	No Rawins data
October 9 (6Z-18Z)	No TOVS data
November 5 (6Z)	No TOVS data
November 9 (18Z)	No Rawins
December 14 (6Z-18Z)	No TOVS data

1990

TIME PERIOD	MISSING DATA
February 7 (18Z)	No Rawins
May 18 (18Z)	No Rawins
June 2 (12Z-18Z)	No TOVS
June 6 (18Z)	No Rawins
June 10	No TOVS
July 13 (6Z)	No Rawins,No Upper Air,less Aircrafts Land, Buoys, and Sat.Wind
July 21 (18Z)	No Land,Ships,Buoys
July 9 (6Z) - July 15	No TOVS
August 20 (12Z-18Z)	No TOVS
September 21 - 30 (12Z)	about 1/2 of Rawins
September 3 (6Z) - September 11	No TOVS
September 17 (6Z- 12Z)	No TOVS
September 24 (12Z-18Z)	No TOVS
October 2 (12Z)	about 1/2 of Rawins
October 13 (12Z)- October 14	No TOVS
November 26 - 31 (except 30,0Z)	No TOVS
December 2 (6Z-18Z)	No TOVS

6 Datasets and Data Access

All output from the assimilation is sent to the local Unitree mass storage system of the NASA Center for Computational Sciences (NCCS). A subset of the data stored in a time series format are also available from the Goddard Distributed Active Archive Center (GDAAC). Access to the data for outside users will be primarily through the GDAAC. Some processed fields (e.g., monthly means) are available via anonymous ftp from a local (DAO) server. Details of the archives are presented below. Further information about methods of accessing the assimilated data may be obtained by sending e–mail to

data@dao.gsfc.nasa.gov

The primary products are the assimilated data (A–D below) which consist of the model output generated during the assimilation cycle and the analysis increments. The assimilation consists of the GCM integration forced by the analysis increments (the increments are recomputed every 6 hours as part of the OI analysis). This assimilation method is called the Incremental Analysis Update (IAU) procedure (Bloom et al. 1991, 1995). All upper air fields are saved at the sigma levels of the assimilating GCM. The prognostic fields are saved every 6 hours as instantaneous quantities. All the upper air diagnostics are saved 4 times daily as 6 hour averages centered on the output time. This is consistent with the IAU procedure in which the increments forcing the model are applied over the 6 hour window spanning the analysis times (0Z, 6Z, 12Z, 18Z). The single level and vertically–integrated fields are saved every 3 hours accumulated over the previous 3 hours.

Separate pressure–level data sets are generated for selected quantities. These are currently stored in two forms. The standard format groups many quantities together, and should be most useful for case studies and multivariable analysis. The time series format groups one month of a single upper–level quantity into one data set; these data sets should be convenient for longer term (multi–year) statistical analyses of selected quantities. We believe the time series data sets will initially be the most popular and these are currently available through the GDAAC.

We also save many of the analysis quantities (E–G below). These include the standard OI analysis and first guess fields produced every 6 hours on the standard pressure levels. They also include information about the input data (first guess–observations at the observation location), quality control information, and the files needed to restart the assimilation or perform model forecasts.

6.1 NCCS Datasets

6.1.1 Overview

Generally, all quantities are saved in IEEE format and on an A– grid. The exceptions are the restart ("r" files, see below) and the analysis increment ("iau" files, see below) files which are saved in full–precision Cray words and on the GCM's C–grid. Saving the increments in this way allows us to reproduce the assimilation exactly by running the GCM with the increments as an additional model forcing. Also, the del files (see below) consist of ungridded observational data. In the following, the date is given as yymmdd (e.g., 850305, is March 5, 1985).

A. Sigma Level Output (4X daily and one day/file)

prognostic	e0054A.prg.tyymmdd	27061584 (27MB)
primary diagnostic	e0054A.pri.tyymmdd	38597280 (39MB)
secondary diagnostic	e0054A.sec.tyymmdd	26012016 (26MB)

The prognostic files consist of the state variables and various lower boundary fields. The primary diagnostics are all the source/sink terms necessary to compute three–dimensional heat, momentum, and moisture budgets. The secondary diagnostics are additional derived quantities, such as clear sky radiances (e.g., for comparing with ERBE data), cloud fraction, and cloud mass flux needed for computing tropospheric transports of tracers.

B. Surface/Vertically–Integrated (8X daily and one day/file)

	e0054A.sfc.tyymmdd	17639328 (18MB)

These are various surface and vertically–integrated fields. Most have a strong diurnal signal and thus need to be saved at least 8 times a day. The vertically–integrated moisture flux is computed "on the fly" at each time step to minimize sampling and interpolation errors. This is important for computing accurate moisture budgets.

C. Analysis Increments (4X daily; one day/file)

	e0054A.iau.tyymmdd	34045952 (34MB)

These are the forcing functions (increments) computed by the OI at each analysis time (00,06,12,18UTC). The assimilation consists of the GCM run with these increments. The assimilation can be reproduced using a single restart file and a history of the increments.

D. Pressure Level Output

1. Standard Format (4X daily; one day/file)

prognostic	e0054A.prog.prs.tyymmdd	24544624 (25MB)
primary diagnostic	e0054A.pri.prs.tyymmdd	34821856 (35MB)

2. Time Series Format (4X daily)

Includes D1 and B	80000000 (80MB)

Each upper level field is in a separate monthly data set. Also includes the estimated height analysis errors. These data are available from the Goddard DAAC.

E. Analysis Fields (4X daily; one day/file)

first guess/analysis	e0054A.prs.tyymmdd	61673600 (62MB)

F. First Guess – Observations (4X daily; one day/file)

Also QC info	e0054A.del.tyymmdd	(approx.) 10000000 (10MB)

G. Restarts (one/file; every 2 days at 03Z)

e0054A.rs.tyymmdd	29941760/2 (15MB)

H. Monthly Means Size (5 years)

Selected fields	4.33 GB

Total output per assimilated day is 370365072 (370 MB) or 0.676 Terabytes/5 years

6.1.2 Details of Assimilation Output

```
DATA DIMENSION: (144, 91)
-180.,-177.5, ....,177.5  (longitude)
-90.,-88., .....,90.  (latitude)


A. Sigma level output:

Definition:  sigma=(p-pt)/(ps-pt) where pt=pressure at top (=10 mb), and
ps=surface pressure.

The quantities are saved at the following 20 sigma levels.  We also show
below the 21 sigma edges.

Sigma levels (20):0.9939, 0.9713, 0.9299, 0.8741, 0.8078, 0.7345,
0.6571, 0.5784, 0.5005, 0.4248, 0.3520, 0.2838, 0.2228, 0.1722, 0.1322,
0.1001, 0.0730, 0.0498, 0.0290, 0.0095

Sigma edges (21): 1.0000, 0.9879, 0.9547, 0.9051, 0.8432, 0.7725,
0.6964, 0.6178, 0.5390, 0.4620, 0.3875, 0.3165, 0.2510, 0.1945, 0.1498,
0.1146, 0.0855, 0.0605, 0.0390, 0.0190, 0.0000


1. Prognostic

e0054A.prg.tyymmdd 27,061,584 BYTES

Each prognostic field is an instantaneous (snap shot) quantity.
Each file contains one day: 00Z, 06Z, 12Z, 18Z.

NSFLD = 9 NUFLD = 6 NDFLD = 0 NDUM  = 2 NLEV = 20 PTOP = 10.000

1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 ALBD    SURFACE ALBEDO (0-1)
3 GWET    SURFACE GROUND WETNESS (0-1)
4 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
5 GTMP    SURFACE GROUND TEMPERATURE (K)
6 SLP     SEA LEVEL PRESSURE (MB)
7 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS2
8 UBAR    VERTICALLY INTEGRATED (BAROTROPIC) UWND
9 VBAR    VERTICALLY INTEGRATED (BAROTROPIC) VWND

1 UWND    U-WIND (M/S)
2 VWND    V-WIND (M/S)
3 HGHT    PERTURBATION GEOPOTENTIAL HEIGHT (M)
4 TMPU    TEMPERATURE (K)
```

```
5 SPHU    SPECIFIC HUMIDITY (G/KG)
6 QQ      TURBULENT KINETIC ENERGY (M**2/S**2)

UNDEFINED FILL VALUE (REAL)             1.0E+15
```

2. Primary diagnostic

e0054A.pri.tyymmdd 38,597,280 BYTES

Each diagnostic field is an average over six hours centered on the output time. The exceptions to this are the first 3 fields which are simply copies from the prognostic file (PS-PTOP is an instantaneous field, PHIS and currently LWI are specified).

Each file contains one day: 00Z, 06Z, 12Z, 18Z.

```
NSFLD = 3 NUFLD = 0 NDFLD = 10 NDUM = 2 NLEV = 20 PTOP = 10.000

1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
3 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS

1 TURBU   U-MOMENTUM  CHNGS DUE TO TURB (M/S/DAY)  20
2 TURBV   V-MOMENTUM  CHNGS DUE TO TURB (M/S/DAY)  20
3 TURBT   TEMPERATURE CHNGS DUE TO TURB (DEG/DAY)  20
4 TURBQ   MOISTURE    CHNGS DUE TO TURB (G/KG/DAY) 20
5 MOISTT  TEMPERATURE CHNGS DUE TO MOIST (DEG/DAY) 20
6 MOISTQ  MOISTURE CHANGES DUE TO MOIST (G/KG/DAY) 20
7 RADLW   TEMPERATURE CHNGS DUE TO LW RAD  (DEG/DAY) 20
8 RADSW   TEMPERATURE CHNGS DUE TO SW RAD  (DEG/DAY) 20
9 OMEGA   VERTICAL VELOCITY OMEGA   (MB/DAY)       20
10 PAVE    TIME-AVERAGED PS-PTOP (MB)               1

UNDEFINED FILL VALUE (REAL)             1.0E+15
```

3. Secondary diagnostic

e0054A.sec.tyymmdd 26,012,016 BYTES

Each diagnostic field is an average over six hours centered on the output time. The exceptions to this are the first 3 fields which are simply copies from the prognostic file (PS-PTOP is an instantaneous field, PHIS and currently LWI are specified).

Each file contains one day: 00Z, 06Z, 12Z, 18Z.

```
NSFLD = 3 NUFLD = 0 NDFLD = 7 NDUM = 2 NLEV = 20 PTOP = 10.000
```

```
1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
3 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS

1 LWCLR   CLEAR SKY LW HEATING RATES   (DEG/DAY)  20
2 CLMOLW  MAXIMUM OVERLAP CLOUD FRACTION FOR LW   20
3 CLROLW  RANDOM  OVERLAP CLOUD FRACTION FOR LW   20
4 SWCLR   CLEAR SKY SW HEATING RATES   (DEG/DAY)  20
5 CLDMAS  CLOUD MASS FLUX  (KG M/SEC**2)          20
6 PAVE    TIME-AVERAGED PS-PTOP (MB)              1
7 DTRAIN  CONVECTIVE CLOUD DETRAINMENT  (KG*M/S)  20

UNDEFINED FILL VALUE (REAL)              1.0E+15
```

B. Surface and vertically-integrated quantities

e0054A.sfc.tyymmdd 17,639,328 BYTES

Each file contains 1 day (Eight time periods:averaged over the previous 3 hours: 00Z, 03Z, 06Z, 09Z, 12Z, 15Z, 18Z, 21Z). The exceptions to this are the first 3 fields which are simply copies from the prognostic file (PS-PTOP is an instantaneous field, PHIS and currently LWI are specified).

NSFLD = 3 NUFLD = 0 NDFLD = 39 NDUM = 2 NLEV = 1 PTOP = 10.000

```
1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
3 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS

1 UFLUX   U-MOMENTUM SURFACE STRESS     (N/M**2)  1
2 VFLUX   V-MOMENTUM SURFACE STRESS     (N/M**2)  1
3 HFLUX   SRFCE FLX OF SENSIBLE    HEAT (W/M**2)  1
4 RADLWG  NET UPWARD   LW RAD. AT GRND  (W/M**2)  1
5 RADSWG  NET DOWNWARD SW RAD. AT GRND  (W/M**2)  1
6 CT      SURFACE DRAG COEF. FOR T AND Q (M/S)    1
7 CU      SURFACE DRAG COEF. FOR U AND V (M/S)    1
8 PREACC  TOTAL PRECIPITATION       (MM/DAY)      1
9 PRECON  CONVECTIVE PRECIPITATION  (MM/DAY)      1
10 WINDS   SURFACE WIND SPEED         (M/S)        1
11 TG      GROUND TEMPERATURE            (DEG)     1
12 TS      SURFACE AIR TEMPERATURE       (DEG)     1
13 QS      SATURATION SURFACE SPEC. HUM. (G/KG)    1
14 OLR     OUTGOING LONGWAVE RADIATION   (W/M**2)  1
15 OLRCLR  OUTGOING LONGWAVE RAD CLEAR SKY (W/M**2)1
16 LWGCLR  SURFACE LONGWAVE FLUX CLEAR SKY (W/M**2)1
```

```
17 RADSWT  INCIDENT SW RAD. AT TOP OF ATM. (W/M**2)1
18 EVAP    SURFACE EVAPORATION  (MM/DAY)          1
19 DPDT    SURFACE PRESSURE TENDENCY  (MB/DAY)    1
20 USTAR   USTAR    (M/SEC)                       1
21 Z0      SURFACE ROUGHNESS Z0   (M)             1
22 PBL     PBL DEPTH (MB)                        1
23 OSR     OUTGOING SHORTWAVE RADIATION  (W/M**2) 1
24 OSRCLR  OUTGOING SHORTWAVE RAD CLEAR  (W/M**2) 1
25 SWGCLR  SURF. SHORTWAVE FLUX CLEAR SKY (W/M**2) 1
26 VINTUQ  VERTICALLY INTEGRATED U*Q   (M/SEC G/KG)1
27 VINTVQ  VERTICALLY INTEGRATED V*Q   (M/SEC G/KG)1
28 VINTUT  VERTICALLY INTEGRATED U*T   (M/SEC  DEG)1
29 VINTVT  VERTICALLY INTEGRATED V*T   (M/SEC  DEG)1
30 CLDFRC  2-DIMENSIONAL TOTAL CLOUD FRACTION (0-1)1
31 QINT    PRECIPITABLE WATER   (GM/CM**2)        1
32 U2M     U AT  2 METERS (M/SEC)                 1
33 V2M     V AT  2 METERS (M/SEC)                 1
34 T2M     T AT  2 METERS (DEG)                   1
35 Q2M     Q AT  2 METERS (KG/KG)                 1
36 U10M    U AT 10 METERS (M/SEC)                 1
37 V10M    V AT 10 METERS (M/SEC)                 1
38 T10M    T AT 10 METERS (DEG)                   1
39 Q10M    Q AT 10 METERS (KG/KG)                 1

UNDEFINED FILL VALUE (REAL)            1.0E+15
```

C. Analysis Increments

e0054A.iau.tyymmdd 34045952 BYTES

These are the analysis fields minus the first guess fields divided by the interval between analysis times (six hours). The quantities are saved in the form required by the GCM prognostic equations on the c-grid in full precision Cray words. PI is the surface pressure minus the pressure at the top (currently 10 mb). Each file contains one day.

```
NSFLD = 1 NUFLD = 4 NDFLD = 0 NDUM = 2 NLEV = 20 PTOP = 10.000


1 ANALP   ANALYSIS INCREMENT OF PSURF    (MB/SEC)

1 ANALU   ANALYSIS INCREMENT OF U-WIND  (M/SEC/SEC)
2 ANALV   ANALYSIS INCREMENT OF V-WIND  (M/SEC/SEC)
3 ANALT   ANALYSIS INCREMENT OF PI*(T/pk) (MB*DEG/(MB)k/SEC)
4 ANALQ   ANALYSIS INCREMENT OF PI*SPHU  (MB*G/G/SEC)

UNDEFINED FILL VALUE (REAL)            1.0E+15
```

D. Pressure Level Data

1. Prognostic (standard format; all quantities together in one file)

e0054A.prog.prs.tyymmdd 49,089,248 BYTES

Each prognostic field is an instantaneous (snap shot) quantity.
Each file contains one day: 00Z, 06Z, 12Z, 18Z.

NSFLD = 9 NUFLD = 6 NDFLD = 0 NDUM = 2 NLEV = 18 PTOP = 10.000

ZLEV: 1000., 950., 900., 850., 800., 700., 600., 500., 400., 300., 250.,
200., 150., 100., 70., 50., 30., 20.

```
1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 ALBD    SURFACE ALBEDO (0-1)
3 GWET    SURFACE GROUND WETNESS (0-1)
4 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
5 GTMP    SURFACE GROUND TEMPERATURE (K)
6 SLP     SEA LEVEL PRESSURE (MB)
7 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS
8 UBAR    VERTICALLY INTEGRATED (BAROTROPIC) UWND
9 VBAR    VERTICALLY INTEGRATED (BAROTROPIC) VWND

1 UWND    U-WIND (M/S)
2 VWND    V-WIND (M/S)
3 HGHT    GEOPOTENTIAL HEIGHT (M)
4 TMPU    TEMPERATURE (K)
5 SPHU    SPECIFIC HUMIDITY (G/KG)
6 QQ      TURBULENT KINETIC ENERGY (M**2/S**2)

UNDEFINED FILL VALUE (REAL)              1.0E+15
```

2. Primary diagnostics (standard format)

e0054A.pri.prs.tyymmdd 69,643,712 BYTES

Each diagnostic field is an average over six hours centered on the
output time. The exceptions to this are the first 3 fields which are
simply copies from the prognostic file (PS-PTOP is an instantaneous
field, PHIS and currently LWI are specified).
Each file contains one day: 00Z, 06Z, 12Z, 18Z.

NSFLD = 3 NUFLD = 0 NDFLD = 10 NDUM = 2 NLEV = 18 PTOP = 10.000

```
ZLEV: 1000., 950., 900., 850., 800., 700., 600., 500., 400., 300., 250.,
200., 150., 100., 70., 50., 30., 20.

1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
3 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS

1 TURBU   U-MOMENTUM  CHNGS DUE TO TURB (M/S/DAY)  18
2 TURBV   V-MOMENTUM  CHNGS DUE TO TURB (M/S/DAY)  18
3 TURBT   TEMPERATURE CHNGS DUE TO TURB (DEG/DAY)  18
4 TURBQ   MOISTURE    CHNGS DUE TO TURB (G/KG/DAY) 18
5 MOISTT  TEMPERATURE CHNGS DUE TO MOIST (DEG/DAY) 18
6 MOISTQ  MOISTURE CHANGES DUE TO MOIST (G/KG/DAY) 18
7 RADLW   TEMPERATURE CHNGS DUE TO LW RAD  (DEG/DAY) 18
8 RADSW   TEMPERATURE CHNGS DUE TO SW RAD  (DEG/DAY) 18
9 OMEGA   VERTICAL VELOCITY OMEGA   (MB/DAY)       18
10 PAVE    TIME-AVERAGED PS-PTOP (MB)               1

UNDEFINED FILL VALUE (REAL)             1.0E+15
```

3. Time Series data

Same quantities as (B) and (D.1) above (excluding pave), except the data have been put in time series format, and the estimated height analysis errors (HGHTE from E below) have been included. This data set will likely be the preferred data set for many investigators, since the fields are stored in a time series format. That is, each upper level quantity is in a separate file (by month), making it easier to perform statistical calculations requiring long time histories of selected fields. These are the data provided to the DAAC at Goddard.

Analysis Output

E. Analysis and first guess fields

e0054A.prs.tyymmdd 61,673,600 BYTES

This file cantains two sets of the quantities listed below. The first guess fields are obtained by setting the record flag to -1. The analysis is obtained by setting the record flag to 0. (see sample read program)

```
NSFLD = 6 NUFLD = 6 NDFLD = 5 NDUM = 1 NLEV = 14 PTOP  = 10.000


ZLEV: 1000., 850., 700., 500., 400., 300., 250., 200., 150., 100., 70.,
```

```
50., 30., 20.

1 PHIS    surface geopotential heights (m/s)**2
2 TVSFC   surface virtual temperatures( K)
3 PSFC    surface pressure(mb)
4 SLP     sea--level pressure (mb)
5 SLU     sea--level u-wind   (m/s)
6 SLV     sea--level v-wind   (m/s)

1 UWND    u-wind (m/s)
2 VWND    v-wind (m/s)
3 HGHT    geopotential height (m)
4 TMPU    temperature (k)
5 MIXR    mixing ratio  (g/kg)
6 RH      relative humidity

1 PRSE    estimate of analysis error in SLP  (mb)             1
2 UWNDE   estimate of analysis error in u-wind  (m/s)        14
3 VWNDE   estimate of analysis error in v-wind  (m/s)        14
4 HGHTE   estimate of analysis error in height  (m)          14
5 MIXRE   estimate of analysis error in mixing ratio  (g/kg) 14
1(i) record type flag                    -1,  0.

Sample time periods (NTYPE =-1 indicates a first guess field)

CURRENT DATE: 850305      0  BEGINNING DATE: 850304 180000  NTYPE:  -1
CURRENT DATE: 850305      0  BEGINNING DATE: 850305      0  NTYPE:   0
CURRENT DATE: 850305  60000  BEGINNING DATE: 850305      0  NTYPE:  -1
CURRENT DATE: 850305  60000  BEGINNING DATE: 850305  60000  NTYPE:   0
CURRENT DATE: 850305 120000  BEGINNING DATE: 850305  60000  NTYPE:  -1
CURRENT DATE: 850305 120000  BEGINNING DATE: 850305 120000  NTYPE:   0
CURRENT DATE: 850305 180000  BEGINNING DATE: 850305 120000  NTYPE:  -1
CURRENT DATE: 850305 180000  BEGINNING DATE: 850305 180000  NTYPE:   0
```

F. First guess - observations

e0054A.del.tyymmdd about 10000000 BYTES

These files contain ungridded information about the data which was incorporated into the assimilation. This includes the first guess-observation at the observation location, the geographical location of the observations and quality control information. These data sets are written after the quality control for each of the three analysis segments (sea-level analysis, moisture analysis, and upper air analysis) is performed. The data sets are writtem using the format described below.

HEADER RECORD containing:

num_obs (integer) Number of observations for synoptic time
nymd (integer) Date in format yymmdd
nhms (integer) Time in format hhmmss
obs_type (char*3) Observation type (SLP,MIX, or ZUV)

NUM_OBS RECORDS containing

lat (real) Latitude of observation
lon (real) Longitude of observation
quality (real) Quality of observation, where 0 indicates a useless
report, 1 indicates only the mass, moisture, or
height data passed quality control (see value of
"obs_type"), 2 indicates only the wind data
passed, and 3 indicates all quantities passed. For
a "MIX" observation type, only values of 1 and 0
are possible.
level (real) Pressure level of observation (unused for "SLP"
observation type)
inst_type (real) Instrument type
u_wind (real) u-wind report (unused for "MIX" observation type)
v_wind (real) v-wind report (unused for "MIX" observation type)
quan_3 (real) Sea-level pressure report, mixing ratio report, or
height report (see value of "obs_type").

For the system used in e0054A, there are four synoptic times and three
analysis segments per data set. Please note that references to "reports"
and "observations" refer here to the "deltas", that is, the difference
between the measured quantity and the model first guess. The full
observed data can be reconstructed by interpolating the first guess
fields (see data sets e0054A.prs.tyymmdd in E above) to the observation
locations, and adding the "deltas".

G. Restarts

e0054A.rs.tyymmdd 29941760 BYTES

Each file contains a 03Z model restart. The restarts are saved every
second day. These are intended only for restarting the model or
assimilation.

Selected Monthly Means

The monthly mean files are organized the same way as the corresponding daily files, except for the addition of quadratic (second moment) quantities and omega in the prognostic files. The sigma level averages are done without mass weighting. The full path name for each set of monthly mean files is /u2/ctmss/e0054A/assim_monthly/monyy_mean/ followed by the file names given below (see section describing the unitree structure).

1. Prognostic fields averaged on sigma levels

e0054A.prog.sig.monyy.mean 23545684 BYTES/MONTH

1	PHIS	SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2	ALBD	SURFACE ALBEDO (0-1)
3	GWET	SURFACE GROUND WETNESS (0-1)
4	PS-PTOP	SURFACE PRESSURE - PTOP (MB)
5	GTMP	SURFACE GROUND TEMPERATURE (K)
6	SLP	SEA LEVEL PRESSURE (MB)
7	LWI	WATER(1) LAND(2) GLACIER(3) SEA-ICE(4)
8	UBAR	VERTICALLY INTEGRATED (BAROTROPIC) UWND
9	VBAR	VERTICALLY INTEGRATED (BAROTROPIC) VWND

1	UWND	U-WIND (M/S)
2	VWND	V-WIND (M/S)
3	HGHT	PERTURBATION GEOPOTENTIAL HEIGHT (M)
4	TMPU	TEMPERATURE (K)
5	SPHU	SPECIFIC HUMIDITY (G/KG)
6	QQ	TURBULENT KINETIC ENERGY (M**2/S**2)
7	OMEGA	OMEGA VERTICAL VELOCITY (MB/DAY)
8	UU	UWND * UWND
9	VV	VWND * VWND
10	TT	TMPU * TMPU
11	ZZ	HGHT * HGHT
12	MM	SPHU * SPHU
13	UV	UWND * VWND
14	UT	UWND * TMPU
15	VT	VWND * TMPU
16	UZ	UWND * HGHT
17	VZ	VWND * HGHT
18	UQ	UWND * SPHU
19	VQ	VWND * SPHU
20	WW	OMEGA*OMEGA
21	WT	OMEGA*TMPU
22	WQ	OMEGA*SPHU

```
2. Primary diagnostics averaged on sigma levels

e0054A.diag.pri.sig.monyy.mean 9649320 BYTES/MONTH

1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
3 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS

1 TURBU   U-MOMENTUM  CHNGS DUE TO TURB (M/S/DAY) 20
2 TURBV   V-MOMENTUM  CHNGS DUE TO TURB (M/S/DAY) 20
3 TURBT   TEMPERATURE CHNGS DUE TO TURB (DEG/DAY) 20
4 TURBQ   MOISTURE    CHNGS DUE TO TURB (G/KG/DAY)20
5 MOISTT  TEMPERATURE CHNGS DUE TO MOIST (DEG/DAY)20
6 MOISTQ  MOISTURE CHANGES DUE TO MOIST (G/KG/DAY)20
7 RADLW   TEMPERATURE CHNGS DUE TO LW RAD  DEG/DAY20
8 RADSW   TEMPERATURE CHNGS DUE TO SW RAD  DEG/DAY20
9 OMEGA   VERTICAL VELOCITY OMEGA   (MB/DAY)      20
10 PAVE    TIME-AVERAGED PS-PTOP (MB)              1

UNDEFINED FILL VALUE (REAL)             1.0E+15


3. Secondary diagnostics averaged on sigma levels

e0054A.diag.sec.sig.monyy.mean 6503004 BYTES/MONTH

1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
3 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS

1 LWCLR   CLEAR SKY LW HEATING RATES   (DEG/DAY)  20
2 CLMOLW  MAXIMUM OVERLAP CLOUD FRACTION FOR LW   20
3 CLROLW  RANDOM  OVERLAP CLOUD FRACTION FOR LW   20
4 SWCLR   CLEAR SKY SW HEATING RATES   (DEG/DAY)  20
5 CLDMAS  CLOUD MASS FLUX  (KG M/SEC**2)          20
6 PAVE    TIME-AVERAGED PS-PTOP (MB)              1
7 DTRAIN  CONVECTIVE CLOUD DETRAINMENT  (KG*M/S)  20

UNDEFINED FILL VALUE (REAL)             1.0E+15


4. Surface diagnostics

e0054A.sfc.monyy.mean 2204916 BYTES/MONTH

1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
```

```
2 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
3 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS

1 UFLUX   U-MOMENTUM SURFACE STRESS     (N/M**2)  1
2 VFLUX   V-MOMENTUM SURFACE STRESS     (N/M**2)  1
3 HFLUX   SRFCE FLX OF SENSIBLE    HEAT (W/M**2)  1
4 RADLWG  NET UPWARD   LW RAD. AT GRND  (W/M**2)  1
5 RADSWG  NET DOWNWARD SW RAD. AT GRND  (W/M**2)  1
6 CT      SURFACE DRAG COEF. FOR T AND Q (M/S)    1
7 CU      SURFACE DRAG COEF. FOR U AND V (M/S)    1
8 PREACC  TOTAL PRECIPITATION       (MM/DAY)      1
9 PRECON  CONVECTIVE PRECIPITATION  (MM/DAY)      1
10 WINDS   SURFACE WIND SPEED         (M/S)        1
11 TG      GROUND TEMPERATURE            (DEG)     1
12 TS      SURFACE AIR TEMPERATURE       (DEG)     1
13 QS      SATURATION SURFACE SPEC. HUM. (G/KG)    1
14 OLR     OUTGOING LONGWAVE RADIATION   (W/M**2)  1
15 OLRCLR  OUTGOING LONGWAVE RAD CLEAR SKY (W/M**2)1
16 LWGCLR  SURFACE LONGWAVE FLUX CLEAR SKY (W/M**2)1
17 RADSWT  INCIDENT SW RAD. AT TOP OF ATM. (W/M**2)1
18 EVAP    SURFACE EVAPORATION  (MM/DAY)           1
19 DPDT    SURFACE PRESSURE TENDENCY  (MB/DAY)     1
20 USTAR   USTAR    (M/SEC)                        1
21 ZO      SURFACE ROUGHNESS ZO   (M)             1
22 PBL     PBL DEPTH (MB)                         1
23 OSR     OUTGOING SHORTWAVE RADIATION  (W/M**2)  1
24 OSRCLR  OUTGOING SHORTWAVE RAD CLEAR  (W/M**2)  1
25 SWGCLR  SURF. SHORTWAVE FLUX CLEAR SKY (W/M**2) 1
26 VINTUQ  VERTICALLY INTEGRATED U*Q   (M/SEC G/KG)1
27 VINTVQ  VERTICALLY INTEGRATED V*Q   (M/SEC G/KG)1
28 VINTUT  VERTICALLY INTEGRATED U*T   (M/SEC  DEG)1
29 VINTVT  VERTICALLY INTEGRATED V*T   (M/SEC  DEG)1
30 CLDFRC  2-DIMENSIONAL TOTAL CLOUD FRACTION (0-1)1
31 QINT    PRECIPITABLE WATER   (GM/CM**2)         1
32 U2M     U AT  2 METERS (M/SEC)                 1
33 V2M     V AT  2 METERS (M/SEC)                 1
34 T2M     T AT  2 METERS (DEG)                   1
35 Q2M     Q AT  2 METERS (KG/KG)                 1
36 U10M    U AT 10 METERS (M/SEC)                 1
37 V10M    V AT 10 METERS (M/SEC)                 1
38 T10M    T AT 10 METERS (DEG)                   1
39 Q10M    Q AT 10 METERS (KG/KG)                 1

UNDEFINED FILL VALUE (REAL)              1.0E+15
```

5. IAU increments averaged on sigma levels

These are the analysis fields minus the first guess fields divided

by the interval between analysis times (six hours). The quantities are saved on the c-grid. The monthly means are converted to ieee format (this is different from the unaveraged IAU quantities which are saved in full-precision Cray words). PI is the surface pressure minus the pressure at the top (currently 10 mb).

e0054A.iau.monyy.mean 4247988 BYTES/MONTH

```
1 ANALP   ANALYSIS INCREMENT OF PSURF    (MB/SEC)

1 ANALU   ANALYSIS INCREMENT OF U-WIND  (M/SEC/SEC)
2 ANALV   ANALYSIS INCREMENT OF V-WIND  (M/SEC/SEC)
3 ANALT   ANALYSIS INCREMENT OF PI*(T/pk) (MB*DEG/(MB)k/SEC)
4 ANALQ   ANALYSIS INCREMENT OF PI*SPHU  (MB*G/G/SEC)

UNDEFINED FILL VALUE (REAL)             1.0E+15
```

6. Prognostic fields averaged on pressure levels

e0054A.prog.prs.monyy.mean 21238492 BYTES/MONTH

```
 ZLEV:  1000.00       950.000       900.000       850.000       800.000
        700.000       600.000       500.000       400.000       300.000
        250.000       200.000       150.000       100.000       70.0000
        50.0000       30.0000       20.0000

   1  PHIS      SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
   2  ALBD      SURFACE ALBEDO (0-1)
   3  GWET      SURFACE GROUND WETNESS (0-1)
   4  PS-PTOP   SURFACE PRESSURE - PTOP  (MB)
   5  GTMP      SURFACE GROUND TEMPERATURE (K)
   6  SLP       SEA LEVEL PRESSURE (MB)
   7  LWI       WATER(1) LAND(2) GLACIER(3) SEA-ICE(4)
   8  UBAR      VERTICALLY INTEGRATED (BAROTROPIC) UWND
   9  VBAR      VERTICALLY INTEGRATED (BAROTROPIC) VWND

   1  UWND      U-WIND (M/S)
   2  VWND      V-WIND (M/S)
   3  HGHT      GEOPOTENTIAL HEIGHT (M)
   4  TMPU      TEMPERATURE (K)
   5  SPHU      SPECIFIC HUMIDITY (G/KG)
   6  QQ        TURBULENT KINETIC ENERGY (M**2/S**2)
   7  OMEGA     OMEGA  VERTICAL VELOCITY  (MB/DAY)
   8  UU        UWND * UWND
   9  VV        VWND * VWND
  10  TT        TMPU * TMPU
  11  ZZ        HGHT * HGHT
```

```
 12  MM        SPHU * SPHU
 13  UV        UWND * VWND
 14  UT        UWND * TMPU
 15  VT        VWND * TMPU
 16  UZ        UWND * HGHT
 17  VZ        VWND * HGHT
 18  UQ        UWND * SPHU
 19  VQ        VWND * SPHU
 20  WW        OMEGA*OMEGA
 21  WT        OMEGA*TMPU
 22  WQ        OMEGA*SPHU
```

7. Primary diagnostics averaged on pressure levels

e0054A.diag.pri.prs.monyy.mean 8705464 BYTES/MONTH

```
 ZLEV:  1000.00       950.000       900.000       850.000       800.000
        700.000       600.000       500.000       400.000       300.000
        250.000       200.000       150.000       100.000       70.0000
        50.0000       30.0000       20.0000

1 PHIS    SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
2 PS-PTOP SURFACE PRESSURE - PTOP  (MB)
3 LWI     LAND_(2), WATER_(1), PERMANENT ICE_(3) SEA ICE_(4) FLAGS

1 TURBU   U-MOMENTUM  CHNGS DUE TO TURB (M/S/DAY) 18
2 TURBV   V-MOMENTUM  CHNGS DUE TO TURB (M/S/DAY) 18
3 TURBT   TEMPERATURE CHNGS DUE TO TURB (DEG/DAY) 18
4 TURBQ   MOISTURE    CHNGS DUE TO TURB (G/KG/DAY)18
5 MOISTT  TEMPERATURE CHNGS DUE TO MOIST (DEG/DAY)18
6 MOISTQ  MOISTURE CHANGES DUE TO MOIST (G/KG/DAY)18
7 RADLW   TEMPERATURE CHNGS DUE TO LW RAD  DEG/DAY18
8 RADSW   TEMPERATURE CHNGS DUE TO SW RAD  DEG/DAY18
9 OMEGA   VERTICAL VELOCITY OMEGA   (MB/DAY)      18
10 PAVE    TIME-AVERAGED PS-PTOP (MB)             1

UNDEFINED FILL VALUE (REAL)             1.0E+15
```

8. First guess and analysis averaged on pressure levels

e0054A.analysis.prs.monyy.mean 15418512 BYTES/MONTH

This file cantains two sets of the quantities listed below. The first guess fields are obtained by setting the record flag to -1. The analysis is obtained by setting the record flag to 0. (see sample read program)

```
ZLEV: 1000., 850., 700., 500., 400., 300., 250., 200., 150., 100., 70.,
50., 30., 20.

1 PHIS    surface geopotential heights (m/s)**2
2 TVSFC   surface virtual temperatures( K)
3 PSFC    surface pressure(mb)
4 SLP     sea--level pressure (mb)
5 SLU     sea--level u-wind   (m/s)
6 SLV     sea--level v-wind   (m/s)

1 UWND    u-wind (m/s)
2 VWND    v-wind (m/s)
3 HGHT    geopotential height (m)
4 TMPU    temperature (k)
5 MIXR    mixing ratio  (g/kg)
6 RH      relative humidity

1 PRSE    estimate of analysis error in SLP  (mb)             1
2 UWNDE   estimate of analysis error in u-wind  (m/s)        14
3 VWNDE   estimate of analysis error in v-wind  (m/s)        14
4 HGHTE   estimate of analysis error in height  (m)          14
5 MIXRE   estimate of analysis error in mixing ratio  (g/kg) 14
```

6.1.3 NCCS Unitree Directory Structure

/u2/ctmss/e0054A

/assim_sigma
/monyy
e0054A.prg.tyymmdd
e0054A.pri.tyymmdd
e0054A.sec.tyymmdd
/assim_sfc
/monyy
e0054A.sfc.tyymmdd
/assim_incr
/monyy
e0054A.iau.byymmdd.eyymmdd
/assim_press
/monyy
e0054A.prog.prs.byymmdd.eyymmdd
e0054A.pri.prs.byymmdd.eyymmdd
/monyy_ts
e0054A.prs.$$$$.byymmdd.eyymmdd
e0054A.tabl.$$$$.byymmdd.eyymmdd
/assim_monthly
/monyy_mean
e0054A.prog.sig.monyy.mean
e0054A.diag.pri.sig.monyy.mean
e0054A.diag.sec.sig.monyy.mean
e0054A.diag.sfc.monyy.mean
e0054A.iau.monyy.mean
e0054A.prog.prs.monyy.mean
e0054A.diag.pri.prs.monyy.mean
e0054A.analysis.prs.monyy.mean
/analysis
/monyy
e0054A.prs.tyymmdd
e0054A.del.tyymmdd
/restarts
/monyy
e0054A.rs.tyymmdd

6.1.4 Sample Read Program (Phoenix Format)

All data sets on unitree (accept for the time series files) are written in our local "Phoenix" format. Sample read programs are given below. The time series format is described in section 6.2.1.

```
Assign statements (for Cray)

1) cray data (iau and restart files)

assign -a $data1                          fort.21


2) ieee data (all other files)

assign -a $data1  -N ieee -F f77          fort.21


Fortran Program

SAMPLE READ PROGRAM
C************************************************************************ C
C PURPOSE
C Routine to read STANDARD Forecast/Analysis output
C
C READ DESCRIPTION
C JOB ..... CHARACTER*8 Experiment job identifier
C NYMD ..... INTEGER Current Date (YYMMDD) of data record
C NHMS ..... INTEGER Current Time (HHMMSS) of data record
C NYMDO .... INTEGER Beginning  Date (YYMMDD) of forecast
C NHMSO .... INTEGER Beginning  Time (HHMMSS) OF forecast
C
C XLABEL ... CHARACTER*80  Character description of experiment
C IM ....... INTEGER Longitudinal dimension of data
C JNP ...... INTEGER Latitudinal  dimension of data
C NSFLD .... INTEGER Number of Surface    Fields
C NUFLD .... INTEGER Number of Upper-air  Fields
C NDFLD .... INTEGER Number of Diagnostic Fields
C
C PTOP ..... REAL Model top pressure
C NULEV .... INTEGER Number of Upper-air  Levels
C ZLEV ..... REAL Upper-air Level Values (Pressure or Sigma Level)
C NLAY ..... INTEGER Number of Model Levels
C SIGE ..... REAL Model Sigma edge values
C
C NDUM ..... INTEGER Number of user-defined      Header variables
```

```
C DDUM ..... CHARACTER*40  User-defined Description of Header variable
C IDUM ..... INTEGER User-defined Integer        Header variable
C RDUM ..... REAL User-defined Real           Header variable
C CDUM ..... CHARACTER*8 User-defined Character      Header variable
C
C NAMES .... CHARACTER*8 Surface Field  Name
C DESCS .... CHARACTER*40  Surface Field  Description
C NAMEU .... CHARACTER*8 Upper-air  Field  Name
C DESCU .... CHARACTER*40  Upper-air  FIELD  Description
C NAMED .... CHARACTER*8 Diagnostic FIELD  Name
C DESCD .... CHARACTER*40  Diagnostic FIELD  Description
C NDLEV .... INTEGER Number of levels  defined for diagnostic field
C
C ZLEVID ... REAL Level Identifier (Pressure or Sigma Level)
C NAME ..... CHARACTER*8 Character Name of quantity being read
C Q ........ REAL Two-dimensional   quantity being read
C
C NOTE:
C For the Datasets produced by GEOS-DAS, the following User-Defined
C Header information has been used:
C
C NDUM    =  2
C DDUM(1) = 'OUTPUT RECORD TYPE (INTEGER)'
C IDUM(1) =  -1  FIRST GUESS
C             0  AFTER ANALYSIS (or STRAIGHT FORECAST)
C             1  INITIALIZED ANALYSIS
C RDUM(1) = 0.100000E+16
C CDUM(1) = 'LOG8R'
C
C DDUM(2) = 'UNDEFINED FILL VALUE (REAL)'
C IDUM(2) = 0
C RDUM(2) = 0.100000E+16
C CDUM(2) = 'UNDEF'
C
C*********************************************************************
C* GODDARD LABORATORY FOR ATMOSPHERES       *
C*********************************************************************
C
PARAMETER  ( IDIM =144 )
PARAMETER  ( JDIM = 91 )
DIMENSION Q( IDIM,JDIM )
CHARACTER*8  XLABEL(10) , JOB,  NAME
DIMENSION     SIGE (50) , NDLEV(100), ZLEV(50)
DIMENSION      IDUM(100),  RDUM(100)
CHARACTER*8   NAMES(100), NAMEU(100), NAMED(100), CDUM(100)
CHARACTER*40  DESCS(100), DESCU(100), DESCD(100), DDUM(100)
DATA KU  /21/
10 CONTINUE
```

```
C *********************************************************************
C ****                        READ HEADER                          ****
C *********************************************************************

READ(KU,END=500) JOB, NYMD, NHMS, NYMDO, NHMSO,
.                XLABEL, IM, JNP,
.                NSFLD, NUFLD, NDFLD,
.                PTOP, NULEV, (ZLEV(K),K=1,NULEV),
.                NLAY, (SIGE(K),K=1,NLAY+1),
.                NDUM, (DDUM(N),IDUM(N),RDUM(N),CDUM(N),N=1,NDUM)
if( ndum.eq.2 ) then
undef = rdum(2)
else
undef = -999.9
endif
READ(KU) ( NAMES(N), DESCS(N),           N=1,NSFLD ),
         ( NAMEU(N), DESCU(N),            N=1,NUFLD ),
           ( NAMED(N), DESCD(N), NDLEV(N), N=1,NDFLD )
C *********************************************************************
C ****           READ SURFACE FIELDS                          ****
C *********************************************************************
DO 20 N=1,NSFLD
READ(KU)  ZLEVID,NAME,Q
20 CONTINUE
C *********************************************************************
C ****           READ UPPER-AIR FIELDS                        ****
C *********************************************************************
DO 30 L=1,NULEV
DO 30 N=1,NUFLD
READ(KU)  ZLEVID,NAME,Q
30 CONTINUE
C *********************************************************************
C ****             READ DIAGNOSTICS                           ****
C *********************************************************************
DO 40 N=1,NDFLD
DO 40 L=1,NDLEV(N)
READ(KU)  ZLEVID,NAME,Q
40 CONTINUE
GOTO 10
500 CONTINUE
STOP
END
```

6.2 GDAAC Time Series Datasets

The Goddard DAAC (GDAAC) holds the time series subset of the full data set (see above; a copy of these data sets is also on the NCCS Unitree system). The files are organized to allow them to be easily read and displayed with the GrADS graphics and display program (available via anonymous ftp from grads.iges.org) developed at the Center for Ocean-Land-Atmosphere (COLA).

The GDAAC has an online Information Management System. This menu-driven interactive data system is open to the public, but access to certain restricted data products require special authorization. To connect to the system via Telnet follow the instuctions below.

Node name: daac.gsfc.nasa.gov

Node number: 192.107.190.139

Example: telnet daac.gsfc.nasa.gov

Username: daacims

Password: gsfcdaac

Upon connecting to the GSFC DAAC host computer you will be asked to register your name and address and will be offered a tutorial to help orient you to the system. For more information on access to the 4-D Assimilated Data Set, contact the GSFC DAAC User Services Office at (301) 286-3209, or via email at daacuso@daac.gsfc.nasa.gov

The DAAC also has a data guide for the time series available via MOSAIC

(http://daac.gsfc.nasa.gov/DATASET_DOCS/dao_data set.html)

Information about these data sets is found in the following four sections:

1. Description of the data sets in timeseries format.
2. Contents of the data sets.
3. Contents of the data descriptor (table) file.
4. Menu of available quantities.

6.2.1 Description of the Datasets

The data representation is IEEE 32 bit floating point and the record blocking is fortran block control words (in a Cray assign the data sets have ieee f77 format). The data sets are stored in one month segments. The directory structure of the timeseries data sets and the associated table files (which describe the contents of the timeseries data sets) is as follows:

```
/u2/ctmss/e0054A/assim_press/monyy_ts
e0054A.prs.$$$$.byymmdd.eyymmdd
e0054A.tabl.$$$$.byymmdd.eyymmdd
where:

monyy = month and year (e.g. mar85)
$$$$ = name of the desired upper level field (see menu in section 4) or name
given to the surface prognostic/diagnostic data set.  The surface prognostic and
diagnostic data sets have multiple fields.  The naming convention for the surface
prognostics is sfcprog.  The naming convention for the diagnostics is diag#
(where #=1,2,3 or 4; see the menu of available quantities in section 4)
byymmdd= beginning year, month and date (e.g. 850301)
eyymmdd= ending year, month and date

For example, the zonal wind for March of 1985 is in the following directory:

/u2/ctmss/e0054A/assim_press/mar85_ts/e0054A.prs.uwnd.b850301.e850331

The associated data descriptor (hereafter table) file (which describes the
contents of e0054A.prs.uwnd.b850301.e850331) is

/u2/ctmss/e0054A/assim_press/mar85_ts/e0054A.tabl.uwnd.b850301.e850331
```

Contents of the data sets

Each data set consists of a series of records. Each record contains one level of a field. The field at one level is read, then the next, and so on. Upper air fields (both prognostic and diagnostic) are available at 18 levels; see the table file (described in Section 3) for specific information about the pressure levels, the resolution, the date and time stamp, etc.

With the exception of the height error data sets (HGHTE), all data sets were obtained by interpolating assimilated data from 20 sigma levels to 18 pressure levels (14 mandatory levels given by 1000 mb, 850 mb, 700 mb, 500 mb, 400 mb, 300 mb, 250 mb, 200 mb, 150 mb, 100 mb, 70 mb, 50 mb, 30 mb, and 20 mb plus 4 additional levels given by 950 mb, 900

mb, 800 mb and 600 mb). In the case of HGHTE, the assimilated data was interpolated from 14 mandatory pressure levels to 18 pressure levels (as listed above).

The upper air fields on pressure levels are read with the bottom level (highest pressure) first, i.e.

```
u wind at 1000 mb
u wind at 950 mb
*
*
*
u wind at 30 mb
u wind at 20 mb
```

A sample READ program for data sets containing upper air quantities is:

```
PARAMETER (IM=144,JNP=91)
DIMENSION FIELD(IM,JNP)
DO 1 ITIMES=1,NTIMES
DO 2 ILEVS=1,NLEVS
READ(8) FIELD
     2     CONTINUE
     1     CONTINUE
```

where

NTIMES(=TDEF in table file) is the number of times data was written
NLEVS (=ZDEF in table file) is the number of pressure levels
IM(=XDEF in table file) is the number of gridpoints in longitude
JNP(=YDEF in table file) is the number of gridpoints in latitude

Some data sets contain a number of different fields at a single level; the table file is used to determine how many fields are in the data set. A sample READ program for these data sets is

```
PARAMETER (IM=144,JNP=91)
DIMENSION FIELD(IM,JNP)
DO 1 ITIMES=1,NTIMES
DO 2 IQTYS=1,NQTYS
READ(8) FIELD
     2     CONTINUE
     1     CONTINUE
```

where

NQTYS(=VARS in table file) is the number of different fields in the

data set. The order of the fields is given by the list of field descriptors located below the VARS entry in the table file. The other parameters in the READ program are as described above.

Contents of the Data Descriptor (Table) File

The table file contains all information necessary to READ a given data set. Each monthly data set has an associated table file. Thus, prior to accessing a given data set, the user should study the associated table file. The following is an example of a table file for a single upper air prognostic quantity:

```
DSET
FORMAT sequential
TITLE  OI1.3/vc5b20 Version 1.3.0
UNDEF  1.0E+15
XDEF  144 LINEAR  -180.0  2.500
YDEF  91 LINEAR  -90.0  2.000
ZDEF 18 LEVELS 1000.0  950.0  900.0  850.0  800.0  700.0  600.0  500.0
400.0  300.0 250.0  200.0  150.0  100.0  70.0  50.0  30.0  20.0
TDEF  112  LINEAR  00:00Z01MAR85  6hr
VARS  1
UWND  18  0  U-WIND (M/S)
ENDVARS
```

In this example, the first two lines are relevant only for users of the Grid Analysis and Display System (GrADS). Next, the title of the experiment is given followed by the undefined, or missing, value 1.0E+15. There are 144 grid points in the x-direction (longitude) with the first grid point at the dateline and with a grid spacing of 2.5 degrees. There are 91 grid points in the y-direction (latitude) with the first grid point at the south pole and with a grid spacing of 2.0 degrees. There are 18 pressure levels with the bottom level (highest pressure) first. Data was written 112 times starting at 00Z on March 1, 1985 and every six hours thereafter. There is 1 variable in this data set, namely the zonal wind UWND.

6.2.2 Menu of Available Quantities

The following is a list of quantities that are currently available in time series format.

```
================================================================
======================== PROGNOSTIC MENU =======================
================================================================


==================
===== sfcprog =====
==================


NAME  LEVELS DESCRIPTION
===== ====== ============
PHIS 1 SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
ALBD 1 SURFACE ALBEDO (0-1)
GWET 1 SURFACE GROUND WETNESS (0-1)
PS-PTOP 1 SURFACE PRESSURE - PTOP (MB)
GTMP 1 SURFACE GROUND TEMPERATURE (K)
SLP 1 SEA LEVEL PRESSURE (MB)
LWI 1 WATER_(1) LAND_(2) PERMANENT ICE_(3) SEA ICE_(4) FLAGS
UBAR 1 VERTICALLY INTEGRATED (BAROTROPIC) UWND (M/S)
VBAR 1 VERTICALLY INTEGRATED (BAROTROPIC) VWND (M/S)


==================
===== upper air  =====
==================


NAME LEVELS DESCRIPTION
===== ====== ===========
UWND 18 U-WIND (M/S)
VWND 18 V-WIND (M/S)
HGHT 18 GEOPOTENTIAL HEIGHT (M)
TMPU 18 TEMPERATURE (K)
SPHU 18 SPECIFIC HUMIDITY (G/KG)
QQ 18 TURBULENT KINETIC ENERGY (M**2/S**2)
HGHTE 18 STANDARD DEVIATION OF THE HEIGHT ERROR (M)
================================================================




================================================================
======================== DIAGNOSTIC MENU =======================
================================================================


==================
```

```
====== diag1 ======
==================


NAME  LEVELS DESCRIPTION
===== ====== ============
PS-PTOP 1 SURFACE PRESSURE - PTOP (MB)
PREACC 1 TOTAL PRECIPITATION (MM/DAY)
PRECON 1 CONVECTIVE PRECIPITATION (MM/DAY)
EVAP 1 SURFACE EVAPORATION (MM/DAY)
VINTUQ 1 VERTICALLY INTEGRATED U*Q (M/S G/KG)
VINTVQ 1 VERTICALLY INTEGRATED V*Q (M/S G/KG)
VINTUT 1 VERTICALLY INTEGRATED U*T (M/S DEG)
VINTVT 1 VERTICALLY INTEGRATED V*T (M/S DEG)
QINT 1 PRECIPITABLE WATER (GM/CM**2)


==================
====== diag2 ======
==================


NAME  LEVELS DESCRIPTION
===== ====== ============
PS-PTOP 1 SURFACE PRESSURE - PTOP (MB)
UFLUX 1 U-MOMENTUM SURFACE STRESS (N/M**2)
VFLUX 1 V-MOMENTUM SURFACE STRESS (N/M**2)
HFLUX 1 SURFACE FLUX OF SENSIBLE HEAT (W/M**2)
CT 1 SURFACE DRAG COEFFICIENT FOR T AND Q (M/S)
CU 1 SURFACE DRAG COEFFICIENT FOR U AND V (M/S)
WINDS 1 SURFACE WIND SPEED (M/S)
USTAR 1 USTAR (M/S)
ZO 1 SURFACE ROUGHNESS (M)
PBL 1 PBL DEPTH (MB)


==================
====== diag3 ======
==================


NAME  LEVELS DESCRIPTION
===== ====== ============
PS-PTOP 1 SURFACE PRESSURE - PTOP (MB)
RADLWG 1 NET UPWARD LW RADIATION AT GROUND (W/M**2)
RADSWG 1 NET DOWNWARD SW RADIATION AT GROUND (W/M**2)
OLR 1 OUTGOING LONGWAVE RADIATION (W/M**2)
OLRCLR 1 OUTGOING LONGWAVE RADIATION CLEAR SKY (W/M**2)
LWGCLR 1 SURFACE LONGWAVE FLUX CLEAR SKY (W/M**2)
```

```
RADSWT 1 INCIDENT SW RADIATION AT TOP OF ATMOS (W/M**2)
OSR 1 OUTGOING SW RADIATION (W/M**2)
OSRCLR 1 OUTGOING SHORTWAVE RADIATION CLEAR SKY (W/M**2)
SWGCLR 1 SURFACE SHORTWAVE FLUX CLEAR SKY (W/M**2)
CLDFRC 1 2-DIMENSIONAL TOTAL CLOUD FRACTION (0-1)


==================
====== diag4 ======
==================


NAME  LEVELS DESCRIPTION
===== ====== ============
PS-PTOP 1 SURFACE PRESSURE - PTOP (MB)
TG 1 GROUND TEMPERATURE (DEG)
TS 1 SURFACE AIR TEMPERATURE (DEG)
QS 1 SATURATION SURFACE SPECIFIC HUMIDITY (G/KG)
DPDT 1 SURFACE PRESSURE TENDENCY (MB/DAY)
U2M 1 U AT 2 METERS (M/S)
V2M 1 V AT 2 METERS (M/S)
T2M 1 T AT 2 METERS (DEG)
Q2M 1 Q AT 2 METERS (KG/KG)
U10M 1 U AT 10 METERS (M/S)
V10M 1 V AT 10 METERS (M/S)
T10M 1 T AT 10 METERS (DEG)
Q10M 1 Q AT 10 METERS (KG/KG)


==================
===== upper air =====
==================


NAME  LEVELS DESCRIPTION
===== ====== ============
TURBU 18 U-MOMENTUM CHANGES DUE TO TURBULENCE (M/S/DAY)
TURBV 18 V-MOMENTUM CHANGES DUE TO TURBULENCE (M/S/DAY)
TURBT 18 TEMPERATURE CHANGES DUE TO TURBULENCE (DEG/DAY)
TURBQ 18 MOISTURE CHANGES DUE TO TURBULENCE (G/KG/DAY)
MOISTT 18 TEMPERATURE CHANGES DUE TO MOIST PROC (DEG/DAY)
MOISTQ 18 MOISTURE CHANGES DUE TO MOIST PROC (G/KG/DAY)
RADLW 18 TEMPERATURE CHANGES DUE TO LW RAD (DEG/DAY)
RADSW 18 TEMPERATURE CHANGES DUE TO SW RAD (DEG/DAY)
OMEGA 18 VERTICAL VELOCITY (MB/DAY)
=========================================================================
```

6.3 Datasets Available from DAO Server

Several files reside on an anonymous ftp account on a local workstation in the DAO to allow outside users to obtain 1) the latest documentation of the data sets, 2) statistics about the observations going into the assimilation, 3) a detailed description of the diagnostics, 4) selected monthly mean and time series files, and 5) a subset of the data with surface marine fluxes. The anonymous ftp account is on hera.gscf.nasa.gov. Here you will find the following directories:

/pub/assimilation/e0054A/documentation/

assim_files.doc.ascii
: This is a plain text version of most of section 6.

assim_files.doc.ps
: This is a postscript version of most of section 6.

assim_files.doc.nofonts.ps
: This is a postscript version of most of section 6 without the embedded fonts.

data_problems
: Contains a summary of problems discovered with the data. This will be updated periodically (see Appendix).

suspic.tyymm (e.g. suspic.t8503)
: These are monthly summaries of outlier counts (and locations) based on the variance at each grid point for the assimilated prognostic fields. These are not necessarily bad data since extreme (many standard deviations) values do occur in nature, especially for specific humidity at upper levels. This is meant to be another source of information for judging the quality of the assimilated fields.

/pub/gcm/

geos1.0_gcm.doc.ps
: Details of the diagnostics (how they were computed, etc.) and further information about the GEOS-1 GCM used in the assimilation may be found in this postscript file.

/pub/assimilation/e0054A/observations

bar_obscountyymm.ps

These are postscript files (bar graphs of data counts). There are also text files summarizing the data going into the assimilation. (see section 5.2)

/pub/assimilation/e0054A/data/monthly_means

selected monthly mean files in GrADS format

/pub/assimilation/e0054A/data/time_series

selected single level time series in GrADS format

/pub/assimilation/e0054A/data/ocean_fluxes

The surface marine ocean flux data sets are individual files with fluxes of heat, momentum and fresh water over the ocean surface, along with basic quantities necessary for recalculation of these fluxes. The files are in NetCDF. There is very easy to use FORTRAN access software and a "point-and-click" IDL utility for browsing of the data.

Acknowledgements We would like to thank Arlindo da Silva for reviewing a draft version of this document. The assimilation experiment and research described in this report were supported by the EOS Interdisciplinary Science Program and by the NASA Research and Applications Program. Computer resources and funding were provided by the EOS Project through the Scientific Computational Facility of the Data Assimilation Office.

Appendix: History of Output Problems

The following list of output problems is updated on an irregular basis and may be obtained via anonymous ftp from the machine: hera.gsfc.nasa.gov (/pub/assimilation/e0054A/documentation). The list presented below was last updated on 1/10/95.

Specifics: (reporting date:problem)

1/10/95: All segments inadvertantly run with the wrong (the previous year's) SST and soil moisture boundary conditions have been replaced (see 9/15/94 entry below). For anyone that used the data with the incorrect boundary conditions it appears that the system adjusts very quickly (within a few days) so that error does not influence results much beyond the dates given below. We did, however, find an unexpected impact in the temperature (and probably other) fields at upper levels–this appears to be due to occasional differences in quality control decisions made (e.g. data rejection) with the incorrect bcs (upper levels are linked by the sea–level pressure analysis).

10/6/94: The diagnostic fields TURBU and TURBV have unrealistic values for 06Z, April 11,1986

10/6/94: The diagnostic field OMEGA is computed incorrectly at the international date-line (I=1).

9/15/94: The following segments were inadvertantly run with the wrong (the previous year's) SST and soil moisture boundary conditions . We are currently in the process of rerunning these segments. We will then replace the January time series data sets in the DAAC, and the monthly means on hera. This will probably take several weeks.

When this is completed we will update this file so please check back in a few weeks. We will also try to provide some indication of the impact on other fields. This problem, of course, impacts the near surface quantities and fluxes directly. We suspect that the impact on most upper level fields will be small.

860101 - 860104

870101 - 870104

880101 - 880104

890101 - 890102

900101 - 900103

9/15/94: There is also a minor problem with the interpolation used for determining the SST and soil moisture fields during the last half of every December, resulting in a slight discontinuity in these fields at the start of a new year.

9/15/94: Note: All clear sky radiative diagnostics are sample every 3 hours. This causes a signature of the diurnal cycle in the daily means purely due to the coarse sampling frequency. This is not true for the full (cloudy) radiative diagnostics, since they are accumulated at each model time step (2.5 minutes).

6/94: Semi–permanent unrealistic high pressure anomaly at (0E, 82S)

6/94: Semi–permanent temperature spike at (72N, 125W). Appears to be due to incorrect LWI flag for the years 1985–88 during which it is always considered a water point. (Should be ice during cold season).

2/94: Unrealistically cold surface and near surface temperatures over land during the cold season. This appears to occur somewhat randomly primarily at gridpoints in regions of steep topography.

8/17/93: Beginning 06Z Nov 16, 1985 the LWI output field was updated to include SEA_ICE (4).

6/18/93: March 15, 1985, 12Z and 15Z, unrealistic precip maximum ($> 200\ mm/day$) over gridpoint in Arctic.

References

Arakawa, A. and W. Schubert, 1974: Interaction of a cumulus ensemble with the large–scale environment. Part I, *J. Atmos. Sci.*, **31**, 674–701.

Arakawa, A. and M.J. Suarez, 1983: Vertical differencing of the primitive equations in sigma coordinates. *Mon. Wea. Rev.*, **111**, 34–45.

Arpe, K. 1990: Impact of changes in the ECMWF analysis–forecasting scheme on the systematic error of the model. Ten years of medium–range weather forecasting. Volume 1, 4–8 September 1989, June 1990, ECMWF, Shinfield Park, Reading RG2 9AX, U.K.

Asselin, R., 1972: Frequency filter for time integrations. *Mon. Wea. Rev.*, **100**, 487–490.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396–409.

Bengtsson, L. and J. Shukla, 1988: Integration of space and in situ observations to study global climate change. *Bull. Amer. Meteor. Soc.*, **69**, 1130–1143.

Bloom, S.C., L.L. Takacs, and E. Brin, 1991: A scheme to incorporate analysis increments gradually in the GLA assimilation system. *Ninth Conf. on Numerical Weather Prediction*, Denver, CO, Amer. Meteor. Soc., 110–112.

Bloom, S.C., L.L. Takacs, A.M. da Silva, and D. Ledvina, 1995: Data assimilation using incremental analysis updates. *To be submitted to Mon. Wea. Rev.*

Burridge, D.M. and J. Haseler, 1977: A model for medium range weather forecasting–adiabatic formulation, *Tech. Report. No. 4, European Center for Medium Range Weather Forecasts*, Brachnell, Berkshire, UK.

Cayan, D.R., C.F. Ropelewski, and T.R. Karl, 1986: An atlas of United States Monthly and Seasonal Temperature Anomalies, December 1930–November 1984. *NOAA US Climate Program Office*, 244pp.

Chou, M.–D., 1984: Broadband water vapor transmission functions for atmospheric IR flux computation. *J. Atmos. Sci.*, **41**, 1775–1778.

Chou, M.–D., and L. Peng, 1983: A parameterization of the absorption in 15–micron CO_2 spectral region with application to climate sensitivity studies. *J. Atmos. Sci.*, **40**, 2183–2192.

Cressman, G. P., 1959: An operational objective analysis scheme. *Mon. Wea. Rev.*, **87**, 329–340.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, Cambridge, 457pp.

Da Silva, A., C.C. Young and S. Levitus, 1995: Atlas of Surface Marine Data 1994 Vol. 1: Algorithms and Procedures, In preparation.

Davies, R., 1982: Documentation of the solar radiation parameterization in the GLAS climate model. NASA Tech. Memo. 83961, 57 pp. [Available from the U.S. Department of Commerce, National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161]

ECMWF, 1989: The description of the ECMWF/WCRP Level III–A global atmospheric data archive. European Center for Medium-Range Weather Forecasts, Shinfield Park, Reading/Berks., RG2 9AX, England. 5285 Port Royal Road, Springfield, VA 22161]

Gruber, A. and A.F. Krueger, 1984: The status of the NOAA outgoing longwave radiation data set. *Bull. Amer. Meteor. Soc.*, **65**, 958-962.

Hamming, R. W., 1989: *Digital Filters*, Prentice–Hall, Englewood Cliffs, N. J., 284pp.

Harshvardhan, R. Davies, D. A. Randall, and T. G. Corsetti, 1987: A fast radiation parameterization for atmospheric circulation models. *J. Geophys. Res.*, **92**, 1009–1016.

Helfand, H. M., and J. C. Labraga, 1988: Design of a non–singular level 2.5 second–order closure model for the prediction of atmospheric turbulence. *J. Atmos. Sci.*, **45**, 113–132.

Helfand, H. M., M. and S. D. Schubert, 1995: Climatology of the simulated Great Plains low–level jet and its contribution to the continental moisture budget of the United States. *J. Climate*, **April**.

Hurrel, J.W. and G.G. Campbell, 1992: Monthly mean global satellite data sets available in CCM history tape formats. NCAR Technical Note, NCAR/TN–371+STR, Boulder, CO.

Joseph, J.H., W.J. Wiscombe, and J.E. Weinman, 1976: The delta–Eddington approximation for radiative flux transfer. *J. Atmos. Sci.*, **33**, 2452–2459.

Kalnay, E. and R. Jenne, 1991: Summary of the NMC/NCAR reanalysis workshop of April 1991. *Bull. Amer. Meteor. Soc.*, **72**, 1897–1904.

King, M.D., and R. Harshvardhan,1986: Comparitive accuracy of selected multiple scatter-

ing approximations. *J. Atmos. Sci.*, **43**, 784–801.

Lacis, A. A., and J. E. Hansen, 1974: A parameterization for the absorption of solar radiation in the Earth's atmosphere. *J. Atmos. Sci.*, **31**, 118–133.

Lanczos, C., 1966: *Discourse on Fourier Series.* Hafner Publishing, 255 pp.

Large, W. G. and S. Pond, 1981: Open ocean momentum flux measurements in moderate to strong winds. *J. Phys. Oceanogr.*, **11**, 324–336.

Levitus, S., 1982: Climatological Atlas of the World Ocean, NOAA Prof. Paper No. 13, U. S. Government Printing Office, Washington DC, 17 fiches, 173 pp.

Molod, A, H.M. Helfand, and L.L. Takacs, 1995: The climate of the GEOS–1 GCM and its impact on the GEOS–1 data assimilation System. (submitted)

Moorthi, S., and M. J. Suarez, 1992: Relaxed Arakawa Schubert: A parameterization of moist convection for general circulation models. *Mon. Wea. Rev.*, **120**, 978–1002.

Pfaendtner, J., S. Bloom, D. Lamich, M. Seablom, M. Sienkiewicz, J. Stobie, A. da Silva, 1995: Documentation of the Goddard Earth Observing System (GEOS) Data Assimilation System–Version 1. NASA Tech. Memo. No. 104606, volume 4, Goddard Space Flight Center, Greenbelt, MD 20771.*

Rienecker, M. M., R. Atlas, S. D. Schubert, and C. A. Scholz, 1995: A comparison of surface wind products over the North Pacific Ocean. *Submitted to J. Geophysical Research–Oceans.*

Rodgers, C. D., 1968: Some extension and applications of the new random model for molecular band transmission. *Quart. J. Roy. Meteor. Soc.*, **94**, 99–102.

Rosenfield, J. E., M. R. Schoeberl, and M. A. Geller, 1987: A computation of the stratospheric diabatic circulation using an accurate radiative transfer model. *J. Atmos. Sci.*, **44**, 859–876.

Sadourny, R., 1975: The dynamics of finite difference models of the shallow water equations, *J. Atmos. Sci.*, **32**, 680–689.

Schemm, J., S. Schubert, J. Terry, and S. Bloom, 1992: Estimates of monthly mean soil moisture for 1979–1989. NASA Tech. Memo. No. 104571, Goddard Space Flight Center, Greenbelt, MD 20771.

Schubert, S. D., C.–Y. Wu, J. Zero and J.–K. Schemm, 1991: Quality control of the consolidated ECMWF/TOGA analyses. Goddard Space Flight Center, Greenbelt, MD. [Report avaliable from S. Schubert, e–mail schubert@dao.gsfc.nasa.gov]

Schubert, S. D., J. Pfaendtner and R. Rood, 1993: An assimilated data set for Earth Science applications, *Bull. Am Met. Soc.*, **74**, 2331–2342.*

Schubert, S. D. and Y. Chang, 1995: An objective method for assessing sources of model error, *To be submitted to Mon. Wea. Rev.*.

Shapiro, R., 1970: Smoothing, filtering and boundary effects. *Rev. Geophys. Space Phys.*, **8**, 359–387.

Spangler, W.M.L. and R. L. Jenne, 1990: World monthly surface station climatology. National Center for Atmospheric Research, Boulder, CO.

Suarez, M.J. and Takacs, L.L., 1995: Documentation of the Aries-GEOS Dynamical Core: Version 2. NASA Tech. Memo. No. 104606, volume 5, Goddard Space Flight Center, Greenbelt, MD 20771.*

Sud, Y. C., and A. Molod, 1988: The roles of dry convection, cloud–radiation feedback processes and the influence of recent improvements in the parameterization of convection in the GLA GCM. *Mon. Wea. Rev.*, **116**, 2366–2387.

Takacs, L.L., A. Molod, and T. Wang, 1994: Documentation of the Goddard Earth Observing System (GEOS) General Circulation Model–Version 1. NASA Tech. Memo. No. 104606, volume 1, Goddard Space Flight Center, Greenbelt, MD 20771.*

Takacs, L.L., and, M.J. Suarez, 1995: Evaluation of the climate of the GEOS–1 GCM: Dynamical processes. (in progress)

Trenberth, K.E., and J.G. Olson, 1988: ECMWF global analyses 1979–1986: Circulation statistics and data evaluation. NCAR Technical Note, NCAR/TN–300+STR, Boulder, CO.

Wentz, F.J., 1992: Revision–1 Update for SSM/I Geophysical Tapes User's Manual. Remote Sensing Systems Techical Memorandum 040792.

* Available from DAO MOSAIC home page (http://hera.gsfc.nasa.gov/dao.home_page.html)

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**

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1995	Technical Memorandum

4. TITLE AND SUBTITLE
Technical Report Series on Global Modeling and Data Assimilation
Volume 6 - A Multiyear Assimilation With the GEOS-1 System: Overview and Results

5. FUNDING NUMBERS
C - NAS5-32332
Code 910.3

6. AUTHOR(S)
Siegfried Schubert, Chung-Kyu Park, Chung-Yu Wu, Wayne Higgins, Yelena Kondratyeva, Andrea Molod, Lawrence Takacs, Michael Seablom, and Richard Rood

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS (ES)
Laboratory for Atmospheres
Data Assimilation Office
Goddard Space Flight Center
Greenbelt, Maryland 20771

8. PEFORMING ORGANIZATION REPORT NUMBER
95B00079

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

10. SPONSORING / MONITORING ADGENCY REPORT NUMBER
TM-104606, Vol. 6

11. SUPPLEMENTARY NOTES
Schubert and Rood: Goddard Space Flight Ctr., Greenbelt, MD; Wu, Kondratyeve, Molod, Takacs, and Seablom: General Sciences Corp., Laurel, MD; Park: Joint Center for Earth System Science, Univ. of Maryland, College Park, MD; Higgins: Climate Analysis Ctr, NMC/NWS/NOAA, Washington, D.C.

12a. DISTRIBUTION / AVAILABILITY STATMENT
Unclassified - Unlimited
Subject Category 46
This publication is available from the NASA Center for AeroSpace Information, 800 Elkridge Landing Road, Linthicum Heights, MD 21090-2934, (301)621-0390.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The Data Assimilation Office (DAO) at Goddard Space Flight Center has produced a multiyear global assimilated data set with version 1 of the Goddard Earth Observing System Data Assimilation System (GEOS-1 DAS). One of the main goals of this project, in addition to benchmarking the GEOS-1 system, was to produce a research quality data set suitable for the study of short-term climate variability. The output, which is global and gridded, includes all prognostic fields and a large number of diagnostic quantities such as precipitation, latent heating, and surface fluxes. Output is provided four times daily with selected quantities available eight times per day. Information about the observations input to the GEOS-1 DAS is provided in terms of maps of spatial coverage, bar graphs of data counts, and tables of all time periods with significant data gaps.

The purpose of this document is to serve as a users' guide to NASA's first multiyear assimilated data set and to provide an early look at the quality of the output. Documentation is provided on all the data archives, including sample read programs and methods of data access. Extensive comparisons are made with the corresponding operational European Center for Medium-Range Weather Forecasts analyses, as well as various *in situ* and satellite observations. This document is also intended to alert users of the data about potential limitations of assimilated data, in general, and the GEOS-1 data, in particular. Results are presented for the period March 1985-February 1990.

14. SUBJECT TERMS
Data assimilation, GEOS-1, reanalysis, multiyear assimilated data set

15. NUMBER OF PAGES
207

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