The NMC/NCAR 40-Year Reanalysis Project

E. Kalnay, M. Kanamitsu, R. Kistler,
W. Collins, D. Deaven, L. Gandin, M. Iredell,
S. Saha, G. White, J. Woollen, Y. Zhu
(NMC Development Division)

M. Chelliah, W. Ebisuzaki, W.Higgins, J. Janowiak, K. C. Mo,
C. Ropelewski, J. Wang,
(NMC Climate Analysis Center)

A. Leetmaa, R. Reynolds
(NMC Coupled Model Project)

Roy Jenne, Dennis Joseph
(NCAR)

Accepted for publication in
Bulletin of the American Meteorological Society

DRAFT
(minor typos possible)

April 1995

ABSTRACT

NMC and NCAR are cooperating in a project (denoted "Reanalysis") to produce a 40-year record of global analyses of atmospheric fields in support of the needs of t research and climate monitoring communities. This effort involves the recovery of l surface, ship, rawinsonde, pibal, aircraft, satellite and other data, quality contro assimilating these data with a data assimilation system which is kept unchanged over reanalysis period 1957 through 1996. This eliminates perceived climate jumps associ with changes in the data assimilation system.

The NMC/NCAR 40-year reanalysis uses a frozen state-of-the-art global data assimilation system, and a data base as complete as possible. The data assimilation the model used are identical to the global system implemented operationally at NMC o January 1995, except that the horizontal resolution is T62 (about 210 km). The data has been enhanced with many sources of observations not available in real time for operations, provided by different countries and organizations. The system has been designed with advanced quality control and monitoring components, and can produce on month of reanalysis per day on a Cray YMP/8 supercomputer. Different types of outpu archives are being created to satisfy different user needs, including a "quick look" ROM (one per year) with twice-daily six tropospheric and stratospheric fields, as we

surface, top of the atmosphere and isentropic fields. Output variables are classifi four classes, depending on the degree to which they are influenced by the observatio and/or the model. For example "C" variables (such as precipitation and surface flux are completely determined by the model during the data assimilation, and should be u with caution. Nevertheless, a comparison of these variables with observations and w several climatologies shows that they generally contain considerable useful informat Eight day forecasts, produced every five days, should be useful for predictability s and for monitoring the quality of the observing systems.

The 40 years of reanalysis (1957-1996) should be completed in early 1997. A continuation into the future through an identical Climate Data Assimilation System (will allow researchers to reliably compare recent anomalies with those in earlier de Since changes in the observing systems will inevitably produce perceived changes in climate, parallel reanalyses (at least one-year long) will be generated for the peri immediately after the introduction of new observing systems, such as new types of sa data.

NMC plans currently call for an updated reanalysis using a state-of-the-art sys every five years or so. The successive reanalyses will be greatly facilitated by th generation of the comprehensive data base in the present reanalysis.

1. Introduction

The NMC/NCAR Reanalysis Project began in 1991 as an outgrowth of the NMC Climate Data Assimilation System (CDAS) project. The motivation for the CDAS projec was the apparent "climate changes" that resulted from many changes introduced in the NMC operational Global Data Assimilation System (GDAS) over the last decade in orde to improve the forecasts. These jumps in the perceived climate parameters obscure, some extent, the signal of true short-term climate changes or interannual climate variability. An obvious example is presented in Fig. 1 which shows large jumps in t analyzed virtual temperature at 1000 hPa in the Pacific Ocean when the model was changed. The impact of system changes on other parameters, such as estimated precipitation and its distribution, is more subtle and therefore harder to separate true climate anomaly signals.

The basic idea of the Reanalysis Project is to use a frozen state-of-the-art analysis/forecast system and perform data assimilation using past data, from 1957 to present (reanalysis). Moreover, the same frozen analysis/forecast system will be us continue to perform data assimilation into the future (CDAS), so that climate resear can assess whether current climate anomalies are significant when compared to a long reanalysis without changes in the data assimilation system. In addition, there will one-way coupled ocean reanalysis, in which the surface fluxes from the atmospheric model will be used for the ocean data assimilation. The NMC/NCAR 40 years reanalysi should be a research-quality data set suitable for many uses, including weather and term climate research.

The project development has been supported by the NOAA Office of Global Programs. An Advisory Panel chaired by Julia Nogu s-Paegle guided it throughout the developmental period (1989-1993). After the execution phase started in 1994, the Advisory Panel was replaced by a Users' Advisory Committee, chaired by Dr. Abraham Oort. The reanalysis system was designed at NMC, with participation of over 25 scie from NMC's Development Division, Climate Analysis Center, the Coupled Model Project, and Automation Division. Scientists at NCAR performed most of the data collection, obtained many special data sets from international sources which were not available operationally through the Global Telecommunications System (GTS). Prof. E. Kung (University of Missouri) acquired early data from China. We also had the collaborat NOAA/NESDIS, who provided the TOVS data, the Geophysical Fluid Dynamics Laboratory (GFDL) for the ocean reanalysis, the United Kingdom Meteorological Office (UKMO) who will supply their global ice and SST reanalyses (GISST) for the earlier periods, the Japanese Meteorological Agency (JMA), who provided cloud-track winds and special rawinsonde data not available on GTS, and European Centre for Medium-range Weather

Forecasts (ECMWF) who filled some data gaps, and provided a sea-ice data base. NASA/GLA has provided retrievals missing from the NMC archives, and offered to perfo TOVS retrievals for several months of missing data in 1986. The NOAA/ERL/Climate Diagnostic Center has provided funding for archival and tape handling development an support.

The early design of the Project was discussed in an NMC/NCAR Reanalysis Workshop that took place at NMC in April 1991 (Kalnay and Jenne, 1991). The Worksho had the participation of representatives of all the groups planning to perform reana that time (Center for Ocean, Land and Atmosphere interactions (COLA), ECMWF and NASA/GLA), as well as of the major types of users (e.g., for short and long term dyn and diagnostics, transport of trace gases, climate change, predictability, angular momentum and length-of-day, coupled models, etc.). The near-final design was review in October 1993 by the Advisory Committee, who suggested several additional tests an modifications before the start of the operational phase (started in May 1994). Representatives of the major agencies interested in the project (NOAA, NSF, NASA, DO and of the other groups performing reanalyses also participated in the review of the NMC/NCAR Project. The other plans presented in the October 1993 review were those o ECMWF, which will perform a 15-year reanalysis for 1979-1994, NASA/GLA (Schubert et al, 1993) which performed a reanalysis for 1985-1990, the US Navy (1985 to the prese and COLA which performed an 18-month reanalysis for the 1982-83 El Ni¤o. Such multiple reanalysis projects offer a great opportunity for cooperation and intercomp which should enhance each of the projects. In particular, NMC has benefitted from t COLA project through the transfer of the Gridded Analysis and Display System (GrADS) which has greatly enhanced the NMC developmental graphical display system, and from the close interaction with NASA scientists performing a similar reanalysis.

The purpose of this paper is to update the documentation of the NMC/NCAR system design, output and plans for distribution. The basic characteristics of the are summarized in Section 2, and the data to be used in Section 3. The three module the reanalysis system (data quality control preprocessor, analysis module with autom monitoring system, and output module) are described in sections 4, 5, and 6 respect The CDAS, which uses the same frozen system but continues the analysis into the futu is discussed in Section 7. Section 8 summarizes the coupling with the ocean reanaly Section 9 contains an assessment of the reliability of the reanalysis output, and th of changes in the observing system. Section 10 summarizes the project. More detail documentation is available from NMC (Office Note 401).

2. Overview of the NMC/NCAR Reanalysis System and Execution Plan

The NMC/NCAR Reanalysis Project has two unique characteristics: the length of the period covered, and the assembly of a very comprehensive observational data base The reanalysis will cover the 40-year period 1957-1996, and will continue into the f with the CDAS. The observations will be saved in the WMO binary universal format representation (BUFR), with additional information, such as the first guess and qual control decisions incorporated into the report. We are also considering the feasibi extending the reanalysis back to 1946, when the Northern Hemisphere (NH) upper air network was established, as suggested by several researchers. The length of the reanalysis, and the desire to carry it out as fast as possible to increase its usefu us to design a system able to perform one month of reanalysis per day. Such a fast of execution required the development of a reanalysis system much more robust and automated than the analysis-forecast systems used for operational numerical weather prediction. As a result, the NMC/NCAR Reanalysis System has many novel features not yet present in operational or research numerical weather forecasting systems.

As shown in Fig. 2, the NMC/NCAR Reanalysis System has three major modules: data decoder and quality control (QC) preprocessor, data assimilation module with an automatic monitoring system, and archive module. The central module is the data assimilation, which has the following characteristics:

T62 model (equivalent to a horizontal resolution of about 210 Km) with 28 verti levels. The model is identical to the NMC global model operational implemented 10 January 1995, except for the horizontal resolution, which is T126 (105 Km) f the operational model (Kanamitsu, 1989, Kanamitsu et al 1991);
Spectral Statistical Interpolation (SSI or 3-D variational) analysis, with no n nonlinear normal mode initialization (Parrish and Derber, 1992, Derber et al, 1 improved error statistics, and the balance constraint on the time derivative of divergence equation implemented at NMC in January 1995 are also included (Derber et al 1994);
Complex QC of rawinsonde data, including time interpolation checks, with confid corrections of heights and temperatures (Collins and Gandin, 1990, 1992); OI-based complex QC of all other data (Woollen, 1991, Woollen et al, 1994);
Optimal Averaging of several parameters over a number of areas, providing more accurate averages, and estimates of the error of the average (Gandin, 1993);
Optimal Interpolation SST reanalysis (Reynolds and Smith, 1994) starting from 1982; UKMO global ice and SST reanalysis (GISST) for earlier periods;
One-way coupled ocean model 4-D assimilation for 1982-onwards (Ji et al, 1994);
The same Climate Data Assimilation System (CDAS) will be used into the future, from January 1995 onwards.

In order to support a rate of reanalysis of about one month per day, it is nece to ensure that the data input is generally free of major data problems such as wrong dates, wrong locations, garbled information, etc, for both conventional and remotely sensed data. This is particularly important for old data which have not been previo used at NMC. Similarly, the rate of one month of reanalysis per day does not allow detailed human scrutiny that operational output normally receives. For this reason, created a data quality control preprocessor and an analysis output QC monitoring mod (Fig. 2):

The data input is pre-processed, and all the analysis output fields are monitor with a "complex QC" monitoring system, in which the statistics of the data, tim tendencies, etc, are compared to climatological statistics in order to detect e These statistics include tendency checks. (These monitoring systems will also implemented operationally and their development constitutes a major spin-off from this project for NMC).

It was decided early in the project that one type of output could not satisfy needs of the many different types of users. For this reason the output module allow several different archives:

Level-2 observational data in Binary Universal Format Representation (BUFR) including QC, climatological, analysis and 6-hour forecast information.
Comprehensive analysis, first guess and diagnostic fields presented in "synopti form (all fields every 6 hours) in the model sigma coordinates, as well as in pressure and isentropic coordinates, in GRIdded Binary (GRIB) format. A restar file is included once a month to allow rerunning shorter periods with enhanced diagnostics.
A time series archive in which each field is available for all times, including standard pressure level fields, precipitation, surface fluxes and other widely diagnostic fields. This format will be the most useful for many users.
A "quick look" archive on CD-ROMs, one per year, including most widely used fields: daily values of variables at selected tropospheric and stratospheric pr levels, surface and top of the atmosphere fluxes, precipitation, monthly and zo averages of most fields, covariances, isentropic level variables, etc.
8-day forecasts performed every 5 days, which should allow predictability studi and estimates of the impact of inhomogeneities in the observing systems coverag with anomaly correlation scores.

A subset of the output is posted in the NMC public server, and is available through anonymous FTP.

NCAR, NCDC and CDC will distribute the bulk of the reanalysis data

An important question that has repeatedly arisen is how to handle the inevitabl changes in the observing system, especially the availability of new satellite data, undoubtedly have an impact on the perceived climate of the reanalysis. Basically, t choices are a) to select a subset of the observations that remains stable throughout year period of the reanalysis, or b) to use all the available data at a given time. (a) would lead to a reanalysis which has the most stable climate, and choice (b) to analysis which is as accurate as possible throughout the 40 years. With the guidanc the Advisory panel, we have chosen (b), i.e., make use of the most data at availabl any time. However, in order to assess the impact of the introduction of new observi systems on the perceived climate of the reanalysis, we have decided to

Produce a parallel reanalysis, at least one-year long, without using a large ne observing system. This will allow the users to assess the extent to which the observing system influences the perceived climate and the annual cycle.

The execution phase started in May 1994, on the Cray YMP 8 supercomputer provided by NMC for this project. About 24 hours of the CRAY YMP (2-7 processors) ar needed in order to perform one month of reanalysis and forecasts per day. By Septem 1995 thirteen years (1982-1994) should be completed (in addition to several years of reruns performed to assess the impact of changes in observing systems, to correct problems discovered in the data base, etc.). Next, the period 1979-1982 will be reanalyzed, and completed around the end of 1995, followed by the 1957-1978 decades. We expect to complete the 40 years of reanalysis (1957-1996) by early 1997. The extension into 1948-1957, if feasible, would be done during 1997.

This first phase of reanalysis will be followed by a second phase in which a 19 state-of-the art system will be used for a second reanalysis. NMC plans currently c an updated reanalysis every five years or so. The successive reanalyses will be mad easier by the availability of the comprehensive data base in BUFR generated by the present reanalysis.

3. The Preparation of Data for Reanalysis

The data collection is a major task that has been performed mostly at NCAR. Surface and upper-air observations are being prepared for the reanalysis. The plan use the data available for the original operational NMC analyses (available from Mar 1962-on), and to add other datasets to capture the older data from about 1948-on. Additional data inputs for 1962-on will provide much more data than was first availa operationally, and will be merged and formatted in BUFR at NMC. The component datasets are listed below. For further details consult the NMC Office Note 401, and extended texts provided by NCAR (see list of texts below).

a) Global Rawinsonde Data: NCAR has tapes of the NMC GTS data with upper-air observations from March 1962-on, which will be the main data source for reanalysis. plan to provide both the GTS data (which also has pibals and aircraft) and also raob national archives in various countries. NCAR has raobs received directly from some countries such as South Africa, Australia, Canada, Argentina, Brazil, UK, France and the U.S. (NCDC). The USAF prepared a global collection of data (TD54) that is mostl the period 1948-70, which will be included. GFDL is helping with processing and che this set, which will all be ready for the first reanalysis. The University of Misso Kung) is collaborating with some of the checking between different sources of the sa data, and has obtained daily upper-air data of 30 stations over China from the Chine State Meteorological Administration for the period of 1954-1962. Under the US-Russi bilateral exchange effort led by Jenne (NCAR) and Shumbera (NCDC), the US has received 20 magnetic tapes with upper air data for 57 USSR stations for 1961-1978. Japanese Meteorological Agency has provided NMC with additional data not available o GTS.

During the reanalysis, it was found that the count of significant level winds w from August 1989 to September 1991 in the NMC tapes, but not over the US and China.

ECMWF supplied their data to fill the gap. Interestingly, the ECMWF had a similar b complementary low count over the US and China.

NCAR, NCDC, Russia, Europe, and other organizations including WMO, have interests in improving the global archive of rawinsonde data. We anticipate various collaborations to improve the basic input sets and to accomplish merges. However, t results will be available for later reanalyses, not for the first one.

b) COADS Surface Marine Data. The Comprehensive Ocean Atmosphere Data Set (COADS) data set, first released in 1983 and recently updated, includes ships, fixed buoys, drifting buoys, pack ice buoys, near-surface data from ocean station reports etc.), and some other data. An update for 1980-93 has been completed, and work is progressing on all the surface marine data for 1947-79.

c) Aircraft Data. Available from the NMC GTS source starting in March 1962. Additional data have been gathered from several sources, including data from New Zealand for February 1984 to June 1988, some of which did not get onto GTS. Aircraf data from experiments such as GATE (summer 1974) and FGGE (1979) will be used. Selected Air Force reconnaissance data is available starting 1947. Data from TWERLE constant- pressure balloons for the S. Hemisphere (Jul 1975 - Aug 1976) will be in t data set. These balloons provide data similar to a single level rawinsonde near 150

d) Surface Land Synoptic Data. Global GTS data (usually every 3 hours) are availab starting 1967 from Air Force or NMC sources. Earlier years are available from the A Force Tape Deck 13, and from U.S. hourly data (from NCDC). The data coverage is fai good from 1949-on.

e) Satellite Sounder Data. The basic radiances are available for the following peri

SIRS IR sounders	Apr 1969 - Apr 1971
SIRS on early NMC tapes (not radiances)	Nov 1969 - Sep 1992
VTPR IR sounders	Nov 1972 - Feb 1979
TOVS sounders (HIRS, MSU, SSU)	Nov 1978 - present
HIRS Data Test System	Aug 1975 - Mar 1976

In the first phase of the reanalysis we plan to use the original operational TO retrievals of NESDIS (2.5o space resolution). A system based on the variational assimilation of variances will be used in the second phase of the reanalysis (to sta 1997/1998). It should be noted that the pilot experiments comparing reanalysis with without the use of satellite data, to be discussed in Section 9, have provided usefu information regarding the uncertainties of the analysis without satellite data. Thi important for the period before 1979 when no TOVS satellite soundings were available We hope to assimilate VTPR and HIRS data available before 1979 for the Southern Hemisphere, although we have no recent experience with that data and there may be unforeseen problems.

f) SSM/I surface wind speeds. SSM/I data became available in July 1987, and at NMC the assimilation of surface winds became operational on 10 July 1993. We adopted th neural network algorithm of Krasnopolsky et al (1994), which results in wind speeds significantly closer to buoys wind speeds, and with better coverage under cloudy conditions than the present operational algorithm used at NMC. We initially used a of the high resolution SSM/I radiance data archived by NESDIS for climate purposes. However, after over 4 years were reanalyzed, several problems were found which indicated that it would be necessary to use the original data set. The high volume these data (much larger than all other data together) also resulted in a significant down of the reanalysis. For this reason it was decided that the first phase of the reanalysis will not include wind speeds from SSM/I (except for limited data impact s We plan to use the SSM/I wind speed, as well as total precipitable water and other parameters in the second phase of reanalysis.

g) Satellite Cloud Drift Winds. From the original NMC tapes, and from the GMS cloud drift winds received from the JMA for the period 1978-1991.

A text entitled "Data for Reanalysis: Inventories" has various maps and display that illustrate the typical coverage of surface and upper air data that are already Most of this information covers the period from about 1948-on. The coverage of data rather encouraging, even for the earlier years. We note, however, that rawinsonde observing networks for Antarctica and the west coast of South America did not start July 1957. Many other reports have been prepared that give more information about t attributes of different datasets and the status of projects to prepare the data. Pa been prepared that focus on different issues; a selection of these papers is given b

Selection of Texts about Data for Reanalysis. Contact NCAR for further information:

Text	
Data for Reanalysis: Inventories	Nov 1992
Sea Surface Temperature data	1 Feb 1993
Sea ice Data	2 Apr 1993
Rawinsonde Data for Reanalysis	24 Oct 1994
Data set of Tropical Storm Locations	26 Jan 1993
NMC Upper Air Data, 1962-72	29 Mar 1994
Global Satellite Sounder Data	12 Aug 1994
Surface Land Synoptic Data	May 1994
Ice Cap Buoy Update	5 Apr 1994
Inventories of Data for Reanalysis	Mar 1995
Analyses for the SH 1951-on	18 Mar 1993
Status of Reanalysis Data	1 Apr 1993
Other (ask for list)	

4. Data Preprocessor

The purpose of the preprocessing Reanalysis module (Fig. 2) is to reformat the data coming from many different sources (Fig. 3) into a uniform BUFR format, and to preprocess one or more years at a time, before the actual Reanalysis module is execu at the rate of one month per day. This allows detection of major data problems with sufficient lead time (a few days before the execution of the Reanalysis), so that hu monitors can try to take corrective action. The preprocessor thus minimizes the nee reanalysis reruns due to the many data problems that frequently appear, such as data wrong dates, satellite data with wrong longitudes, etc. The preprocessor also inclu preparation of the surface boundary conditions (SST, sea ice, etc.).

a) Satellite data

A special satellite TOVS soundings data monitoring system has been developed. is aimed at a quality control of the data in the NESDIS archive tapes that can suffe errors in dates and orbits not likely to occur in the daily operational products. S data in grid boxes of 10o by 10o as well as single satellite observations, are quali controlled. The average in the box, the variance in the box, and the absolute value tendency of the box average are compared with a climatology to flag suspicious group satellite data.

b) CQC with temporal check

The Complex Quality Control system (described in the next section) is included the Preprocessor, but without the use of the first guess of the model. The baseline in the preprocessor (see next section), allows the detection of changes in the stati locations, an important problem that interferes with the accurate detection of clima change.

c) Climatological QC test of data

The automatic monitoring system developed for the reanalysis output (Section 8) based on climatological tests with 3-dimensional (grid point) statistics computed fo month. The space-time character of the statistics proved to be very successful in f problems in the pilot experiments for the reanalysis (Section 8), which were then re unusual data errors, leading also to corrections of several errors present in the op system. This led us to check the data directly within the preprocessor, by expressi observation anomaly in units of standard deviations with respect to climatology, a n which can be generated from the BUFR "events" archive (see Section 5f). Such a chec allows human monitors to check for unusual data present in unusual amounts, before t execution of the monthly reanalysis, and provides the OIQC with additional informati can be used by its decision making algorithm as input to the reanalysis.

d) Boundary fields

The following analyses and climatologies are used for the boundary fields: 1) S Reynolds reanalysis for 1982-on, when AVHRR data became available; and the UKMO Global Ice and SST analysis (GISST) for earlier periods, (David Parker, pers. comm.) Snow cover: NESDIS weekly analyses and climatology, updated weekly (Don Garrett, pers. comm.). 3) Sea-ice: The ice field derived from SSMR/SSMI data, and quality controlled by Nomura for the ECMWF reanalysis, has also been adopted at NMC for the period 1979 through 1993. Beyond 1993 a similar algorithm developed by R. Grumbine being used. For earlier periods we plan to use the analysis from Joint Ice Center a when available, John Walsh and GISST, analyses otherwise. These have been incorporated into the SST analysis, so that all values below -1.8C are considered se R. Grumbine has inserted a more realistic glacial coverage for the Ross Ice Shelf an other regions of the Antarctic. A simple check (comparison with monthly climatologi standard deviations) should help to ensure that no major errors are present in the d made inadvertently during their use. 4) Albedo: Matthews (1985) 5) Soil wetness: Upd during the analysis cycle. The model uses the Pan-Mahrt,1987, Mahrt and Pan, 1984) model. There is no nudging of the soil moisture using concurrent data, and a very s coefficient (0.05) is used to nudge towards climatology. Soil moisture fields show interannual variability but no long-term drift (Fig. 7). 6) Roughness length: fro 7) Vegetation resistance: from SiB (Dorman and Sellers, 1989). However, preliminary reanalyses showed that the original resistance over regions deemed to be covered by winter wheat, had excessively high plant resistance in the summer and fall, resultin high temperatures and low precipitation in the eastern North America summer (S. Saha l. Pan, pers. comm.) For this reason, we are using the minimum monthly resistance v for each grid point. Monthly climatologies are the backup of the analyzed fields wh these updated fields are not available.

5. Data assimilation module

a) System Configuration

The CDAS/Reanalysis is executed at the NOAA Central Computer Facility in Suitland, Maryland. Unlike the operational NMC system, which currently is based on IBM-MVS type and Cray-UNIX computers, in the CDAS/Reanalysis system all processing is done in the Cray-UNIX environment. Observations will be encoded in BUFR and gridded data in GRIB, the standard WMO formats. This system will soon be also adopt at NMC for its normal operations.

The Reanalysis will be performed using the present Cray YMP 8 processors, 128 MW supercomputer and the smaller Cray EL2. Other hardware includes a Robotic Silo, upgraded in August 1994 with 4490 STK drives, with storage capacity of 0.6 GB per ta Over 2000 tape slots have been reserved for this project. Since the Cray YMP was saturated, the start of the Reanalysis had to wait for the new Cray C90 acquired by to be installed (early 1994) and the operational systems migrated out of the Cray YM (April 1994). Software used includes the Unicos 7 operating system, NFS mount of Cr complex files, Bourne shell UNIX scripts, Fortran, some C, some X Windows, the Data Migration Facility, the Cray Reel Librarian, and the graphics system GrADS (COLA).

Recent changes include the installation of UNICOS 8, and the replacement of the Bour
shell with the Korn (POSIX) shell. In addition, we expect that the Cray YMP and Cra
will be replaced in 1995 by two Cray J916.

b) Analysis scheme

The Spectral Statistical Interpolation, a 3-dimensional variational analysis sc (Parrish and Derber, 1992, Derber et al., 1991) is used as the analysis module. Its implementation in 1991 replacing an Optimal Interpolation analysis led to major anal and forecast improvements, especially in the tropics, and a major reduction in the precipitation spinup. An important advantage of the SSI is that the balance imposed the analysis is valid throughout the globe, thus making unnecessary the use of nonli normal mode initialization. Recent enhancements, such as improved error statistics, the use of the full tendency of the divergence equation in the cost function (replac original linear balance of the increments constraint), have also been included (Derb 1994, Parrish et al, 1995) . The SSI used in the reanalysis is the same as the syst implemented in the operational system in January 1995, which was tested in parallel over 10 months and resulted in significantly improved forecasts.

c) Model

The T62/28 level NMC global spectral model is used in the assimilation system, implemented in the NMC operational system in December 1994. The vertical structure the model is shown in Table 1. The model has 5 levels in the boundary layer and abo levels above 100 hPa. The lowest model level is about 5 hPa from the surface, and t top level is at about 3 hPa. This vertical structure was chosen so that the boundar is reasonably well resolved and the stratospheric analysis at 10 hPa is not much aff by the top boundary conditions. The details of the model dynamics and physics are described in Development Division (1988), Kanamitsu (1989), and Kanamitsu et al (199 The model includes parameterizations of all major physical processes, i.e., convecti large scale precipitation, shallow convection, gravity wave drag, radiation with diu and interaction with clouds, boundary layer physics, an interactive surface hydrolog vertical and horizontal diffusion processes. A major difference with the model as described by Kanamitsu et al (1991) is the use of a simplified Arakawa-Schubert convective parameterization scheme developed by Pan and Wu (1994) based on Grell (1993). Pre-implementation experiments showed that the simplified Arakawa-Schubert scheme results in much better prediction of precipitation than the previous Kuo sche over the continental US, as measured by equitable threat scores over North America. addition, the precipitation patterns over the tropics are more realistic, with a smo distribution, and less concentration over tropical orographic features. Two other r improvements were also implemented into the reanalysis model. The first is a better diagnostic cloud scheme (Campana et al, 1994) which has resulted in model generated outgoing long-wave radiation (OLR) in much better agreement with observations. The second is a new soil model, based on Pan and Mahrt (1987), which has also resulted i much more realistic surface temperature, and more skillful predictions of precipitat North America in the summer. These changes to the model were systematically tested running two months of assimilations in the summer and in the winter, and 25 forecast from each assimilation. Some tuning of the cloudiness and cloud optical properties performed to correct systematic temperature and cloudiness errors. The final versio the model also produced good 5-day forecast scores.

d) Complex Quality Control of Rawinsonde data

The method of Complex Quality Control (CQC) (Gandin, 1988), is used to quality control the rawinsonde heights and temperatures. CQC first computes residuals from several independent checks (i.e., it computes the difference between an observation the expected value for that observation from each check). It then uses these residua together with an advanced decision making algorithm (DMA) to accept, reject, or corr data (Collins and Gandin, 1992). The checks included in the CQC code for rawinsond heights and temperatures used for the reanalysis include: hydrostatic check, increme check with respect to the 6-hour forecast, horizontal interpolation check, and verti interpolation check. In addition, there is a baseline check based upon the differen between the station elevation and the elevation that is consistent with the reported pressure and the lowest two reported heights, using a standard lapse rate and the hydrostatic equation. Using the same information and assumptions, a mean-sea-level pressure may be obtained and compared with a forecast mean-sea-level pressure. In t way, both an increment and horizontal residual of mean-sea-level pressure are comput The baseline check may allow the determination of errors in the location of stations as changes in their locations.

In addition to these checks, used operationally at NMC, the reanalysis affords possibility of performing also a temporal interpolation check, which cannot be done NWP system. The value of the heights and temperatures at observation time may be compared with those for 12 or 24 hours earlier and later. The temporal residual is difference between the reported height or temperature and the value interpolated fro value before and one after, when they are available. Statistics show this check to comparable value to the incremental check. It is used along with other available ch and is particularly useful in the data preprocessor, where the first guess, and henc incremental check are not available.

The CQC for rawinsonde heights and temperatures performs quite well. The code has been running operationally at NMC for several years and has undergone steady improvement. At present, about 7% of the rawinsonde observations are found to have least one error. Of the hydrostatically detectable errors in mandatory level height temperatures, 75% are confidently corrected, and 60% of the errors detected by use o baseline check are also corrected. The absolute number of corrections for the early of reanalysis may be anticipated to be smaller, depending upon data density, but the may be a higher percentage of data that need to be corrected.

The CQC methodology is also used to quality control the mandatory and significa level rawinsonde winds. We expect to develop a limited capability to correct some errors, e.g. winds manually entered off by multiples of 100 degrees in direction or of 100 knots in speed.

e) Optimal Interpolation Quality Control of all data

The Optimal Interpolation Quality Control (OIQC, Woollen, 1991, Woollen et al, 1994) was developed as the final screening for observations to be used in the data assimilation. The goal of OIQC is to detect and withhold from the assimilation, dat containing gross errors generated by instrumental, human, or communication related mistakes which may occur during the process of making or transmitting observations. also withholds observations with large errors of representativeness, which are accur but whose measurements represent spatial and temporal scales impossible to resolve properly in the analysis-forecast system. The OIQC system uses the same statistical representativeness error model as the objective analysis system it precedes and, therefore, will detect observations unrepresentative for that system.

Three principles guide the OIQC algorithm: 1) use of multivariate three-dimensi statistical interpolation for obtaining comparison values for each observation from neighbors; 2) a complex of independent quality control components consisting of interpolation and other types of checks which when evaluated collectively suggest wh errors exist in an observation; and, 3) "non-hierarchical" decision making algorithm which no final accept/reject decision is made for any datum until all checks which m affect that decision are completed.

The OIQC components are interpolation checks; an OI of the appropriate variable made to each datum, from a group of observations nearby, forms a comparison value. Univariate and geostrophic horizontal checks are performed for each datum checked, a well as a univariate vertical (profile) check for sounding data. Temperature data i converted to units of equivalent thickness difference from a background (usually a s forecast) for the checks, while wind data is checked in terms of vector wind deviati combination of individual check outcomes determines whether a datum is accepted by t system (see NMC Office Note 401, and Woollen, 1991, Woollen et al., 1994, for furthe details).

For the reanalysis we plan to add to the complex of checks performed by the OIQC, two more quantitative checks: One is produced by the time interpolation check the CQC (see previous section), and the second is the deviation with respect to climatology, measured in units of the local analysis standard deviation climatology discussion in Section 4c). Both of these should be powerful additions to the QC, an made possible by the use of the BUFR with QC system for storing data (see next subsection).

f) BUFR Observation "Events" Files

The final step in observation preprocessing, described in Section 3, consolidat incoming data from all sources into BUFR files, an internationally accepted standard format for level 2 data. Provisions have been made to archive in the reanalysis BUFR files, along with each original observation (in a "push-down" fashion), a spectrum o processing information, collectively known as "events". At present, these include an indication of the observation's source, all quality control decisions and their sour history of all modifications made to the observation prior to the analysis (QC corre radiation corrections, virtual temperature conversions, etc.), and various backgroun quantities relevant to the analysis process (i.e. interpolated first guess values, i analyzed values, interpolated climatological values, interpolated climatological var and observation error estimates). As a result, the observational data base produced reanalysis system contains a fairly complete processing history of each observation, can be useful in the evaluation of the performance of the analysis procedures themse as well as to other reanalysis projects carried out at NMC, or at other research cen The BUFR database archive is a major enhancement of prior NMC observational formats, and as such, has been implemented as an operational product in the NMC Global Data Assimilation System.

g) Optimal Averaging

The NMC/NCAR Reanalysis system will include not only the computation of grid-point (and spherical harmonics) values, but also temporal and spatial averages over prescribed areas. A new method, known as Optimal Averaging (Gandin, 1993), will be used in the course of Reanalysis. This method assures minimum (in statistical sense root-mean-square averaging errors and, particularly important for the Reanalysis pur it provides this estimated rms error as a by-product. The incorporation of Optimal Averaging will result in an increased ability for detection of climate change, becau averaged values are less subject to the small-scale everyday variability that acts a complicating the climate change signal detection.

In the Reanalysis System, optimal averages over prespecified areas are computed for temperature, specific humidity, u and v components of the wind, and wind speed a seven (1000, 850, 700, 500, 300, 200, 100 hPa) pressures levels. The horizontal are currently include nine 20o latitude bands from the South to the North Pole. The wei are computed by normalizing the optimal weights so that the sum of the weights over area is equal to one. We also include the geographical regions chosen by the Intergovernmental Panel for Climate Change for climate change monitoring (IPCC 1990, p157), plus two regions covering South America: Tropical South America, 10N-20S, 40W 80W; and Extratropical South America, 20S-50S, 50W-70W.

Two additional computations are included in the optimal averaging component. Optimal averages of the data increments (observation value minus first guess value) calculated by the same module that currently averages the actual observations. In addition, averages of the first guess over the same areas are computed directly from spectral coefficients of the assimilation model.

h) Periodic Forecasts from the Reanalysis

Prof. J. M. Wallace suggested to perform global forecasts during the reanalysis order to keep the computational requirement at a feasible level, we perform one 8-da forecast every 5 days. Such forecasts will be useful for predictability studies, in estimates of the accuracy of the analysis, and estimates of the impact of changes in observing systems. They will also support the development of adaptive Model Output Statistics from the long reanalysis (Paul Dallavalle, pers. comm.)

6. Reanalysis output

The design of the Reanalysis Output has been a major component of the project development (see also Schubert et al, 1995, for a discussion of the NASA reanalysis output). During the April 1991 Reanalysis Workshop, it became clear that there are different types of possible applications for the reanalysis output, and that some of (e.g., transport of greenhouse gases, which needs in principle all turbulent transpo between any two layers) have storage requirements that far exceed what can be handle by the project. For this reason, it was decided that each unit of reanalysis output month) will include restart files, so that special purpose shorter reanalyses with e output can be performed a posteriori.

The reanalysis gridded fields have been classified into four classes, depending the relative influence of the observational data and the model on the gridded variab A indicates that the analysis variable is strongly influenced by observed data, and is in the most reliable class (e.g., upper air temperature and wind). The designatio indicates that, although there are observational data that directly affect the value variable, the model also has a very strong influence on the analysis value (e.g., hu and surface temperature). The letter C indicates that there are no observations dir affecting the variable, so that it is derived solely from the model fields forced by assimilation to remain close to the atmosphere (e.g., clouds, precipitation, and sur fluxes). Finally, the letter D represents a field that is obtained from climatologi and does not depend on the model (e.g., plant resistance, land-sea mask). Appendix contains the complete classification of variables. Although this classification is necessarily somewhat subjective, the user should exercise caution in interpreting th results of the reanalysis, especially for variables classified in categories B and C addition to this simple guidance, the user should keep in mind that quadratic variab (e.g., kinetic energy, transport of water vapor) are in general less reliable than t components from which they were computed. Appendix B contains the list of mandatory pressure levels, sigma levels and isentropic levels of the output.

The reanalysis archive has been designed to satisfy two major requirements: 1) output should be comprehensive, allowing, for example, the performance of detailed budget studies, and 2) it should be easily accessible to the users interested in lon series of data. It became clear that it was not possible to satisfy both requiremen single archival format. For this reason the output module includes several differen archives. In this section we describe four types of archives, and the automatic mon system that was designed to quality control the output.

a) BUFR Observational archive

Reanalysis observational data undergo multiple processing stages, any of which may influence the quality of subsequent analysis and forecast products (see Sections and 5). For purposes of monitoring and review, and for research based on the Reanaly it is useful to be able to trace the progression of QC and related processing to whi particular observation, or group of observations, have been subjected prior to their non-use) in the actual data assimilation. The BUFR observation event archive format (described in Section 5f) has been designed to provide researchers with this capabil

Although the details of the BUFR format, and the BUFR structures devised to support the observation events archive, are rather complicated, a FORTRAN programmer interface package has been developed to simplify a user's interaction with these fil enable fairly straightforward access to all of the archive information, without the great deal of technical expertise in BUFR. These "user-friendly" FORTRAN interface routines, along with appropriate documentation and instructions for their use, will available to Reanalysis investigators.

b) Main synoptic archive

This is the most comprehensive archive of the reanalysis, and will contain a la number of analysis and first guess "pressure" fields at 00, 06, 12, and 18 GMT on a latitude-longitude grid; "flux", "diagnostic" and "sigma" files on the model gaussia 192*94 points, in order to maintain maximum accuracy, and restart files at full reso order to ensure reproducibility. The complete list of output fields with their clas given in the Appendix A. Table 2 summarizes the type and volume of the BUFR data archive and the gridded synoptic archives:

REANALYSIS ARCHIVE VOLUME ESTIMATES

Table 2: Comprehensive "synoptic format" archives. See Appendix B of the NMC Offi Note 401 for a detailed list of fields, units, etc., contained in each of the files

(GB=Giga Bytes)

FILE	ANAL	GUESS	TOTAL	Mb/File	Mb/day	GB/Mon	GB/40y
Restart files (non-grib)							
Sigma	4	4	8	1.9			

.9

9

15.4

5.4

5.4

0.47

.47

47

225.6

FILE	ANAL	GUESS	TOTAL	Mb/File	Mb/day	GB/Mon	GB/40y
Surface	4	4	8	1.2			

.2

2

9.6

.6

.6

0.29

.29

9

139.2

Grib SST/SNOW/SEA-ICE

(GRIB)T/SNOW/SEA-ICE

SST	1	0	1				
	NA	0.09	0.00	1 1.24			
Snow	1/week	0	1/week				
	NA	0.00	0.00	1/week 0.00			
Sea-ice	1	0	1				
	NA	0.00	0.00	1 0.0101			

BUFR Observation

Prepbufr	0	4	5	7.44	29.76	0.89	428.54
Fnlbufr	4	0	10	8.60	34.40	1.03	495.36

Grib files (Grid point)

Grib files (Grid point)

Pressure	4	4	8	1.734	13.84		

3.84

.84

0.42 199.30

Sigma	4	4	8	3.312.9	26.48	0.79	381.31
Grb2d	0	4	4	1.35			

.35

35

5.40

5.40

 5.40
  0.16   77.76296.4
Grb3d   0   4   4   6.310.1   25.24

25.24

                                                              25.24
  0.76   363.465.1
Isen   0   4   4   0.99

.99

                                             99
    3.96   0.12

.12

                                                                         12
    57.0227.6
Non-grib file

Zonal	4	4	8	0.05	0.42	0.01	6.11
Optavrg	4	0	4	0.03.0	0.12		

.12

                                                                2
    0.001   1.723.1
Total (GB)   4.942.7   2376.60.5111

STK cartridge (.6 GB each)   9   3960

c) Reduced "Time Series" archive

This archive contains basic upper air parameters on standard pressure levels (T 3), selected surface flux fields (Table 4), and diabatic heating and radiation terms analysis cycle for the entire Reanalysis period. Most of the data will be saved in format. The pressure level data will be saved on a 2.5 degree lat/lon grid, while t surface flux fields and radiation/diabatic heating data will be saved on a T62 Gauss (192 x 94). In addition, monthly means of vorticity, divergence, virtual temperatur specific humidity, and surface pressure are saved in spherical harmonic form on sigm levels. Monthly means of the flux terms are stored on the gaussian grid.

The radiation/diabatic heating data will be composed of two radiative terms (sho long wave) and four diabatic heating terms (large-scale condensation, deep convectio shallow convection, and vertical diffusion). Monthly means of these data are stored each sigma level of the 00Z, 06Z, 12Z, and 18Z cycles separately, so that the monthl mean diurnal cycle in these fields is preserved.

The data storage order will be markedly different from the manner in which mode have traditionally been stored at NMC. Since the climate research community general use individual parameters at a single atmospheric level but for many time periods (r that all parameters for a single time), much of the data are stored in chronological synoptic order. That is, individual fields for a single atmospheric level are avail time" from a single data structure. The basic pressure level data and surface flux be stored in this order, which is referred to as "time series" in the rest of the pa

Table 3. Data fields on standard pressure levels to be saved on a 2.5 degree lat/lon grid (144 x 73).

A	Zonal wind
A	Meridional wind
A	Geopotential Height
A	Virtual Temperature
A	Absolute Vorticity
B	Vertical Velocity (1000 to 100 hPa only)
B	Specific Humidity (1000 to 300 hPa only)

17 levels: 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 7 30, 20, 10 hPa

Table 4. Surface flux data to be saved on the T62 Gaussian grid (192 x 94):

B	Surface Temperature	B	Precipitable Water
C	Skin Temperature	C	Snow Depth
B	2 m Temperature	B	Snow Cover
B	Surface Pressure	C	Precipitation (total & convective)
D	Albedo	B	Mean Relative Humidity (multiple layers)
C	Surf. Sens. & Lat. Fluxes	C	Soil Wetness, Temperature
C	Top of atmos. Fluxes	C	Surface Runoff
B	Zonal wind at 10 meters	C	Cloud fraction (hi, mid, low)
B	Meridional " " "	C	Cloud forcing, clear sky fluxes
C	Surface Wind Stress	C	Gravity Wave Drag
A	Mean Sea Level Press.	B	Max and Min temperature

d) Quick-look CD-ROM output

Following the suggestion of the Advisory Committee, we are creating a "Quick-lo data base that can fit into a relatively small (1 per year) number of CD-ROMS. The complete content of the CD-ROMs is described in Appendix C. It includes twice-daily values of u, v, Z, T at 3 tropospheric and 3 stratospheric pressure levels (1000, 50 100, 50, 20 hPa). In addition, the CD-ROM contains daily values of total precipita water, surface stress, latent and sensible heat flux, net long and short wave flux a surface, precipitation, and surface pressure, SST, air surface temperature, soil temperature and moisture (2 levels). It also includes isentropic potential vorticit p at 3 isentropic surfaces; monthly averages and zonal cross-sections of many fields their covariances. See Appendix C for the contents of the CD-ROMs.

e) Automatic monitoring system for the Reanalysis Output

As previously noted, the NMC Reanalysis system was designed to perform one month of analyses (analyzed and archived every 6 hours) every day, 5 days a week. Since the volume of reanalysis output is very large, it is not possible for a human to review and check all the reanalysis products, detect major errors, drifts, etc. fulfill this requirement, we developed an automatic monitoring system (Saha and Chel 1993, Kistler et al, 1994). At the end of each month of reanalysis we check the tim series of geopotential height, zonal wind, meridional wind, temperature and humidity standard pressure levels generated for every 6-hour period (00, 06, 12 and 18 UTC).

To monitor the pressure time-series, we use a preliminary climatology based on NMC's daily global data assimilation system (GDAS) over a 7-year period from 1 July 1986 to 30 June 1993. From these daily values we computed monthly averages, standard deviations from the monthly means, standard deviations of the tendency (difference between successive analyses at 00Z) and standard deviations of the interpolation che (difference between the analysis and the interpolation from analyses made 24 hours before and after). These statistics were computed for geopotential height, zonal wi meridional wind, and temperature at each grid point at 12 mandatory pressure surface and for humidity at 6 pressure surfaces. For monitoring the surface flux quantities year preliminary climatology (1 Feb 1992 - 31 Jan 1993) of daily surface flux files T62 model-based operational GDAS system at NMC has also been created, since no long-term archive was available. These short term preliminary climatologies will be replaced by longer ones derived from the reanalysis itself.

We use the monthly statistics of climatological means, daily standard deviation and daily standard deviations of the time interpolation check, at each 3-D grid poin for each month of the year, to perform several statistical checks: We check for "fi outliers" by computing for each variable the percentage of points whose distance to climatological monthly mean is larger than 2 standard deviations. If this percentag larger than the largest value observed in the 7-year climatology, we identify the fi "outlier", and proceed with further checks and diagnostics. These percentages are a graphically presented to the human monitor so that trends or jumps are immediately apparent. In addition, we also identify "single grid point outliers" that differ fr climate mean at each point by more than a specified number of standard deviations. check has proven to be very important in identifying bad input data, such as radioso data or satellite wind data. A "time interpolation" check, is performed by computi similar statistics for the percentage of the points whose difference with a time int field (from plus and minus 24 hours values) is larger than expected.

Both of the above checks have proven to be very effective, and have been consequently adapted for use in the data preprocessor (Section 4). A number of subt errors due to format changes, unusual data types, or quality control decisions, were discovered not only in the Reanalysis system, but also in the operational system. T usefulness of the check is due to the fact that the climatological standard deviatio computed for every month and for each grid point in 3-D, and were therefore much mor sensitive than any other "gross check" previously used at NMC. Further refinements the automatic monitoring system will be developed using the first 5 year reanalysis "climatology" now available.

With respect to the level 2 data, the human monitor will also have available th following information: output of the climatological check of the observations (avail the preprocessor before the monthly reanalysis), the normal operational output of th OIQC and the CQC, and several additional plots. These include a plot of the mean an rms data fits to the first guess and to the analysis, classified by region; a "curta plot of the daily normalized rms fits of the data at all levels; and a plot of the d out by the OIQC. These plots should allow a human monitor to check large amounts of data in order to detect serious problems.

7. The Climate Data Assimilation System (CDAS)

The Reanalysis Project originated with the idea of performing a "post analysis" a Climate Data Assimilation System (CDAS), which would remain frozen into the future 1990, Profs. Mark Cane and Julia Nogu s-Paegle of the Advisory Committee suggested that a very long reanalysis would be more useful than the CDAS alone. The developme of the Reanalysis system was then started and became the largest component of the project. It is clear that the combination of Reanalysis for the past, and the CDAS future, both using the same frozen system, will be much more helpful to researchers either component alone.

The CDAS analysis will be performed within 3 days of the end of the month, with the same software as the reanalysis. This will allow time to capture the bulk of an delayed data, and serve as the basis for the generation of the monthly Climate Diagnostics Bulletin of CAC.

As noted before our plans include a second phase of the CDAS/Reanalysis to star sometime in 1997, after the first phase is completed. In the second phase, the Reanalysis-2 will be performed with a 1998 state-of-the-art system, coupled with a corresponding CDAS-2 into the future. Such reanalyses would then be repeated every five years or so using the most advanced systems and the additional recovered data f the past. The CDAS-1, however, will be continued into the foreseeable future in ord maintain the longest homogeneous data assimilation product possible. Given that the CDAS-1 will become less expensive with time, it may be feasible to consider running fixed observation system ("choice a" at the end of section 2) for comparison with th current reanalysis, which has considerable variations in the observing systems.

8. Coupling with the ocean

In the first phase of Reanalysis we will couple the atmospheric analysis with t optimal interpolation reanalysis of SST for 1982 onwards. For the earlier periods w use the Global sea-Ice and SST Analysis (GISST) that the UKMO has offered to make available (Parker et al, 1993), or possibly a reconstructed SST under development at NMC. In addition a one-way coupled ocean reanalysis will be also performed.

a) The NMC SST Analysis

The NMC routinely produces a 1o gridded SST analysis using optimum interpolatio (OI). The analysis is produced both daily and weekly, using seven days of in situ d (ship and buoy) and bias corrected satellite SST. The first guess for the SST analy the preceding analysis. Because time scales of SST anomalies are of the order of months, the analysis from the previous week is a much better estimate of the current than climatology. There is a large-scale bias correction for satellite data, found from past experience and because the OI method assumes that the data are unbiased (Reynolds, 1988 and Reynolds and Marsico, 1993). The present version of the OI with the bias correction is a significant improvement over the earlier NMC analysis and o any other analysis that uses uncorrected satellite data. In the tropics, the equato eastern Pacific and Atlantic cold tongues are more realistically shown in the OI. A latitudes, the OI shows tighter gradients in the Gulf Stream, the Kuroshio and the Falklands/Malvinas current regions. The statistics estimated in the process of deve the SST OI analysis, show that ship SST observations have larger errors (1.3oC) compared to the errors of buoy and satellite SSTs (0.3-0.5oC). In addition the e-fo correlation scales have been found to range between 500 and 1200 km (Reynolds and Smith, 1994).

The weekly version of the OI SST reanalysis has been computed for the period November 1981 to the present. It is not practical to extend the period prior to Nov 1981, because the present operational satellite instrument (AVHRR) first became operational at that time. To develop a method to produce reliable SST analyses befo November 1981, empirical orthogonal functions were computed from the monthly OI analyses for the 12 year period from January 1982 trough December 1993. A reduced s of spatial EOFs were then used as basis functions that were fitted to the in situ to determine the correct temporal weighting of each function. Monthly SST anomalies we reconstructed from the spatial EOFs and the temporal weights for the period 1950-198 and a 2o grid from 45S to 69N. These fields capture most of the variance shown by i analyses, while eliminating much of the noise due to sparse in situ data sampling. collaborative effort, the UKMO is testing a modification of the original GISST inclu EOF expansion in order to enhance the signal-to-noise ratio for the periods before 1 (N. Rayner, pers. comm., 1995).

b) The ocean reanalysis system

The Coupled Model Project at NMC has been performing 4-D ocean analyses for the last 7 years in order to document more thoroughly current and past climate varia These analyses also serve as the initial conditions and verification fields for the ocean-atmosphere model used for multi season forecasting (Ji et al., 1994). Since m of the potential extended predictability is thought to be the result of coupled inte the tropics, the focus at NMC has been the development of the ocean analysis in this region, but we plan to extend the ocean analysis domain to the entire globe.

The ocean model used was developed at the Geophysical Fluid Dynamics Laboratory (GFDL). The Pacific model has a domain that extends from 45S to 55N and 120E to 70W. The Atlantic domain extends from 100W to 20E and 50S to 65N. The bottom topography is variable and there are 28 levels in the vertical. The zonal gr spacing is 1.5o in the Pacific and 1o in the Atlantic. The meridional grid spacing within 10o of the equator and gradually increases outside this zone to 1o poleward o Within 10o of the northern and southern boundaries the model fields are relaxed to climatological estimates. A Richardson number formulation for the vertical mixing i in the upper ocean. Lateral mixing is formulated as proportional to the square of t equivalent horizontal wave number.

The data assimilation system (Derber and Rosatti, 1989) is a 3-D variational technique applied continuously in time. Presently, only thermal data is used in the analyses. All available temperature data from ships, satellite estimates, drifting moored buoys, and expendable bathythermographs (XBTs) are used. Extensive quality control procedures have been developed to screen the data before they are used. Corrections are made to the model thermal fields in the upper ocean down to 720 mete This depth range contains the maximum depth of the bulk of the available subsurface thermal data from the T4 and T7 type XBT probes. Surface observations are kept in th analyses for two weeks and subsurface observations are kept in for four weeks with weights varying linearly in time. Maximum weight is given at the observation time wi minimum weights at the beginning and end of the time interval the data is being used

Routine weekly ocean model based analyses are performed for the Pacific and Atlantic Basins, with a two week delay in order to allow be able to use the XBT data assimilation for four weeks. The models are forced with a weekly averaged stress fie derived from the 4 times per day near surface winds produced by NMC's global atmospheric analyses. These winds are converted to stress using a constant drag coefficient of 1.3 E-3. The net heat flux used to force the model is set to zero in facilitate the evaluation of the heat fluxes in the analysis system. The net fresh is also set to zero; once a year the salinity field is restored to the mean climatol of Levitus (1982).

Using this system, ocean reanalyses have been performed for the Pacific and Atlantic basins. For this purpose, all available historical subsurface data was obt from the archives and edited for the time period June 1982 to the end of 1992. These were merged with the data that were available in real time. The Pacific reanalysis June 1982 to December 1992 has been completed, and a reanalysis for the Atlantic for the same period is underway.

Work on implementing a global model-based ocean analysis system will start in 1995. The routine weekly analysis capability will be implemented first; the global reanalysis for the period 1982 to the present will be started in 1995, using the atm reanalysis, and should take a few months for completion. With respect to earlier pe the scarcity of subsurface ocean data implies that a meaningful reanalysis can only done for the Northern Hemisphere, and only for the period starting in the late 1960' Nevertheless, once a consistent set of forcing fields is available from the atmosphe reanalysis, we plan to perform the ocean reanalysis for the whole 40 years period (1 1996).

9. Preliminary results and reliability of the atmospheric reanalysis

The NMC/NCAR reanalysis will produce 40 years of daily atmospheric and surface fields, which, for some variables, are close to a best estimate of the evolving stat atmosphere. The analysis cycle, with the use of the 6-hour forecast as a first gues able to transport information from data rich to data poor regions, so that even in r data void areas the reanalysis can estimate the evolution of the atmosphere over bot synoptic and climatological time scales.

A researcher using the reanalysis should be aware, however, that the different outputs are not uniformly reliable. As indicated in section 6, fields derived from dimensional analysis are not equally influenced by observations. Some, such as uppe mass and temperature fields, (classified as A in the Appendix A) are generally well by the observations, and, given the statistical interpolation of observations and fi provide an estimate of the state of the atmosphere better than would be obtained usi observations alone. Others (classified as B) are partially defined by the observati also strongly influenced by the model characteristics. For example, the amount of moisture that the tropical model atmosphere can hold depends on its parameterization cumulus convection, since some convection schemes tend to dry out the atmosphere more than others. Therefore, even if the analysis incorporates rawinsonde and satel moisture data, the overall humidity will be influenced by the climatology of the mod is even more true for quantities which are not directly observed, or whose observati are not currently assimilated into the present analysis systems. Examples of these quantities (classified as C) are precipitation and surface fluxes. To the extent th model and its physical parameterizations are realistic, these fields can be reliable provide estimates as accurate as any other available, even on a daily time scale. However, they will have regional biases if the model tends to be biased. For exampl over southeastern US the model tends to be colder and drier than the atmosphere duri the summer months. As a result, during the 6-hour forecast in the analysis cycle, t model tends to precipitate the increment of moisture added by the rawinsonde observations during the analysis. This process of permanent "spin-down" within the analysis cycle leads to excessive reanalysis precipitation in this area.

In this section we present a few results from the first 5 years of Reanalysis (1989), comparisons with the then operational Global Data Assimilation System (GDAS), and several diagnostic studies. The impact that unavoidable changes in the observin systems (especially the introduction of new satellite data) will have on the reanaly also assessed. These results should provide an indication of the reliability of dif reanalysis fields. Further results are presented in the proceedings of the Climate Diagnostics Workshop (Chelliah, 1994, Saha et al, 1994, Janowiak, 1994, Smith, 1994, White, 1994).

Global energy and water balance

Output from the NMC/NCAR reanalysis includes many diagnostics of the physical forcing of the atmospheric flow, including complete surface energy and hydrological budgets, top-of-the-atmosphere radiation budget, angular momentum budgets and monthl mean diabatic heating (White, 1994). Fig. 4a compares the global mean radiation bud at the top of the atmosphere and 4b the surface energy budget from the reanalysis fo 1985-91 with climatological estimates from Ramanathan et al. (1989) and Morel (1994) For most of these fields, classified as "C", the reanalysis agrees with the climatol well as the different climatologies agree with each other. At the top of the atmosp upward short wave radiation from reanalysis appears to be 11 W/m2 stronger than the climatological estimates (which are forced to be in balance) and the atmosphere lose W/m2 to space. There is some evidence that the ocean surface albedo in the NMC mod is too high, and this may increase the upward solar radiation. At the surface the radiation is 5-8 W/m2 less than the climatological estimates and the atmosphere lose W/m2 to the surface. Consistent with the loss of energy to space and the surface, t NMC model cools slightly during the six-hour first guess forecast. The zonal mean a regional distributions of surface fluxes in reanalysis also appear to be consistent climatological estimates.

Fig. 5 displays monthly means of the global mean hydrological budget and 12-month running means of net atmospheric fluxes from the reanalysis for 1985-91. Over entire period, evaporation exceeded precipitation by 0.04 mm/day. An annual cycle c be seen, with maximum values in July. Global mean precipitation is within the range climatological estimates. There is little evidence of any long-term drift in globa in reanalysis.

The effect of SSM/I wind speeds

SSM/I data became available from the Defense Meteorological Satellite System (DMSP) in July 1987, and NMC started to use operationally ocean surface wind speeds derived with the algorithm of Goodberlet et al (1989) in July of 1993, after paralle showed a positive impact of this data. We originally reprocessed a data base of SSM radiances archived for climate purposes provided by NESDIS (N. Grody and R. Ferraro, pers. comm.) to derive estimates of wind speeds. We used a neural-network algorithm developed by Krasnopolsky et al (1995). The neural network algorithm which is nonli results in significantly closer collocations with buoys than the previous operationa algorithm of Goodberlet et al. (1989), and is less sensitive to clouds and moisture, much larger coverage.

It was discovered in March 1995 that the SSM/I wind speeds assimilated for the period July 1987 - Dec 1991, computed from the climate SSM/I data base, did not cont a transformation from "antenna" temperature to "brightness" temperature. A prelimin evaluation estimated that this error created 10m wind speeds with a positive bias of 2 m/sec. This bias resulted in an increase in surface fluxes of 5-10%. With the c brightness temperature SSM/I data the jump is much smaller. However, the very larg volume of the original SSM/I radiance data, and even of the reduced SSM/I radiance d archived and quality controlled by F. Wentz (NASA, pers. comm.), results in a very significant slow down of the speed of the reanalysis processing. For this reason we decided not to use SSM/I winds n the first phase of the reanalysis. The second phas include the use of all SSM/I derived products.

Sensitivity of monthly means

Before the reanalysis began, the impact on monthly mean fields of changes in th model used for the first guess was examined. These changes included the effect of horizontal and vertical resolution (T62 and T126) and different convection schemes. Salstein (1992) also examined the effect of horizontal resolution during May 1992. results indicated that upper level divergent flow, precipitation and stratospheric w most sensitive to changes in the NMC analysis/forecast system. The large-scale patt of upper-level divergent flow (the scales represented by the velocity potential) app be fairly robust in the tropics; however, the magnitude of the upper-level divergen the tropics and the smaller-scale features are still poorly defined by a modern stat art analysis system.

Precipitation and soil moisture

Precipitation and soil moisture have "C" classifications, which means that data these types are not assimilated, but rather are derived completely from the model 6- forecasts. Figures 6a and 6b depict precipitation from both Reanalysis and from a d set containing satellite-derived rainfall estimates over the oceans (Spencer 1993, S 1994) from the Microwave Sounding Unit (MSU) and raingauge data over land. Precipitation maps are presented for the July-August means of the years 1987 and 198 and soil moisture maps for August 1987 and 1988 are included.

We chose to compare the results between 1987 and 1988 because of the large precipitation shifts that were observed in many important regions of the tropics, as with the transition between "warm episode" El Ni¤o/Southern Oscillation (ENSO) conditions (Rasmusson and Carpenter 1983) and "cold episode" conditions (Shukla and Paolino 1983) in the tropical Pacific during the 1986-1988 period. SSTs were more t 3oC higher over much of the tropical Pacific during the warm event compared to the c event, which had large impacts on the pattern of tropical convection and subsequent heat release. Ropelewski and Halpert (1987, 1989) have shown that precipitation ten less (more) than normal over India and the surrounding ocean during warm (cold) episodes, and provide evidence that the Pacific ITCZ is displaced southward during w episode conditions relative to cold episode conditions. The precipitation patterns India and the Pacific ITCZ region were subsequently documented for these specific ev by Janowiak and Arkin (1991). The rainfall patterns between the Pacific warm and col episodes that are represented in the MSU/raingauge data set described above are also consistent with the studies mentioned above.

The difference in the pattern of Reanalysis precipitation between the northern summers of 1987 and 1988 compare well with those observed by the MSU/raingauge data set and with the studies mentioned above, over both the India region and the Pacific While the amplitude of the differences are considerably less than those of the MSU/raingauge estimates in the tropical Pacific, there are also large differences in magnitude among independent satellite estimates of rainfall such as those based on M infrared, and SSM/I data. The MSU estimates, like infrared algorithms, tend to overestimate rain rates and their geographical extent (P. Arkin, pers. comm.). The reanalysis contains smaller scales than the MSU, which is probably too smooth.

The soil moisture changes (Fig. 6c, cover color picture) show that in the reana India was wetter and North America generally drier in 1988 than in 1987, as observed The MSU/gauge estimates (Fig. 6b) suggest that Central America had more rain in 1988 than in 1987, but that most of the rest of South America was drier in 1988. This is true in the reanalysis, but with weaker amplitudes. The reanalysis underestimates t intensity of the drought in the south and east of the U.S during 1988.

Overall, the soil moisture generally appears to be reasonable, and does not sho long term tendency to drift into excessively dry or wet regimes (Fig. 7a), even wit of surface data but with a small nudging towards climatology. Figs 7b and 7c show t maximum and minimum monthly soil moisture content as a percentage of the field capac (2.00 m), and the month of occurrence (only every third grid point arrow is plotted) expected, the maximum soil moisture generally occurs at the end of the winter in mid latitudes, and at the end of the monsoonal regime in the tropics. The minimum occurs generally after the summer in mid-latitudes, and before the monsoon in the tropics.

A quantitative comparison of the reanalysis precipitation anomalies over the U. (with respect to the 5 year mean) with the monthly precipitation anomalies estimated the NCDC Climate Divisions, shows a pattern correlation of the anomalies of about 40 60%, somewhat higher in the winter than in the summer.
Fig. 8 shows daily precipitation rates for May 1985-89 over the U.S. in (a) the NMC/ reanalysis and in (b) the observations. The corresponding standard deviations of th mean precipitation rates within May 1985-89 are also shown for the reanalysis (c) an the observations (d). The observations were obtained from the hourly precipitation base compiled by the Techniques Development Laboratory of the NWS, and contain about 300 NWS sites and 2500 cooperative stations. This data were gridded on a 2.5o grid (Ying Li, pers. comm.). A comparison of Figs. 8a and 8b shows that in the southeast U.S., the reanalysis precipitation is larger than observed by a factor of almost two previously mentioned, this is due to a regional spin-down of the model, which, being slightly drier and colder than atmosphere, tends to rain out increments of moisture introduced by the analysis. However, the daily variability of the precipitation ana compares quite well with the station variability (Figs. 8c and d).

Quasibiennial oscillation and the stratospheric analysis

The operational NMC global data assimilation system had poor resolution in the stratosphere until July 1993, when the vertical resolution was increased from 18 to levels, and the top model levels was moved up to 2.7hPa, changes which were also incorporated into the reanalysis. Fig. 9a shows a 50 hPa Hovmoeller diagram (longitude/time) of the zonal velocity at 5N to 5S for the operational GDAS for the years 1985-1989 (denoted climate diagnostics data base or CDDB). This was the highe mandatory pressure level available in the GDAS at the time. Fig. 9b shows the same year plot for the reanalysis. The quasi- biennial oscillation is very clear in the and essentially absent in the operational analyses. Fig. 10 shows a log pressure/ti cross-section from 100 hPa to 10 hPa for a point in the reanalysis near Canton Islan The characteristics of the cross-section are similar to those shown by Reed and Roge (1962) for the Canton Island data. Their hand analysis of the station rawinsonde da has a downward propagation of the phase, with a faster change from easterlies to westerlies than the reverse. Comparisons of the reanalysis near Singapore also show there is good agreement with the rawinsonde data, indicating that the analysis syste able to assimilate well the data even in the upper equatorial stratosphere.

More generally, comparisons of the stratospheric reanalysis with the off line stratospheric analysis performed by NMC (Finger et al, 1993) shows very good agreeme (S-K. Yang, pers. comm.).

Impact of the FGGE observing system

A study was made to assess the impact that the introduction of the full satelli observing system will have on the reanalysis (Mo et al, 1995). Two sets of analyses forecasts were made with and without the use of satellite data (SAT and NOSAT) withi the data assimilation. The resulting impact is smaller than that obtained in previo satellite impact studies made using data from the FGGE (1979) experiment, reflecting effect of improvements that have taken place in the global analysis scheme and the model. Overall, the results are very encouraging, indicating that a long reanalysis be useful even before 1979, when the FGGE satellite observing system was established In the Northern Hemisphere (NH) the analyses of both primary variables and eddy flux are basically unaffected by the satellite data, and even in the Southern Hemisphere large component of both the monthly and the daily anomalies can be captured in the absence of the satellite data.

Fig. 5a of Mo et al (1995) showed the zonal average of the square of the correlation between the NOSAT and SAT daily analyses. It indicated that the NOSAT analysis explains close to 100% of the daily variance of the SAT geopotential height analysis in the NH extratropics, between 70% to 90% in the tropics, more than 90% in mid latitudes of the SH, and between 40% and 80% in the Antarctic region. Fig. 5b showed the square correlation of the zonally asymmetric stationary (monthly averaged eddies defined by the two analyses. The comparison suggested that the NOSAT capture over 90% of the zonal variance of monthly mean stationary waves of the SAT analyses most of the tropics and SH down to 60S, whereas in the NH extratropics the agreement once again close to 100%. With respect to the bias of the zonally averaged values, agreement between SAT and NOSAT is generally good, except above 200 hPa and in the polar regions. Obviously the differences increase for more sensitive quantities, su quadratic fluxes and their divergence. Typically, the relative differences between meridional fluxes of zonal momentum or heat estimated by the SAT and the NOSAT are less than 10% in the NH extratropics and less than 20% in the SH midlatitudes, but t can be as large as of order one in the tropics, stratosphere and south of 60S. Sate data did not impact substantially the estimated precipitation fields.

Comparisons with other operational analyses

Finally, we compare the NMC operational Global Data Assimilation System in use during 1992 (T126/18 levels and Kuo convection) with parallel runs using the new Simplified Arakawa-Schubert scheme (Pan and Wu, 1994) and the T62/28 level system adopted in the reanalysis. We also compared the NMC and several other operational analyses. These differences are probably the best way to estimate the precision of resulting analyses given similar observational data bases, and therefore, are representative of the robustness of the NMC reanalysis fields.

We define "internal analysis differences" as the rms difference between monthly means computed with NMC systems using different models. "External analysis differences" are the rms differences between NMC's monthly mean analysis and those o other operational systems. The internal differences reflect the sensitivity to the used in the analysis, and are an estimate of the uncertainty in the monthly mean ana of the NMC system. For the Northern Hemisphere (20N-80N), the internal differences about 3 m at 850 hPa, and 6 m at 500-200 hPa. In the Southern Hemisphere the intern differences are 5m, 8m, 15m, 30m at 850, 500, 300 and 200 respectively, reflecting t much higher uncertainty introduced by the lack of rawinsonde data. The external differences between the NMC analysis and the UKMO analysis are about 12m, 7m, 9m, and 12m at 850, 500, 300, and 200 hPa respectively. The larger values at 850 hPa re the uncertainty introduced by different terrains and extrapolations below the surfac the Southern Hemisphere (20S-80S) the differences are about 20m, 12m, 15m, and 25m respectively. Comparisons with other operational systems were similar.

For the monthly mean wind analysis, in the NH the internal rms differences in b the zonal and meridional component are about 0.4 m/s at 850 hPa, and 0.7 m/s at 200 hPa. The external rms differences are about 1 m/sec at 850 hPa and 1.2 m/s at 200 h In the tropics (20S-20N), the internal rms differences for the zonal wind analysis a 0.7m/s at 850 hPa and 2m/s at 200 hPa. The external rms difference between the NMC and the UKMO operational analysis in 1992 was 2 m/s at 850 hPa and 2.5 at 200 hPa. For the meridional component, the rms differences were about 30% smaller. In the Southern Hemisphere, the internal differences at 850 hPa were about 0.8m/s and the external differences 1.3m/s, and the meridional wind rms differences about 20% small At 200 hPa both the internal and external rms differences were about 1.8m/s for U an m/s for V.

These figures indicate the precision with which modern analysis systems can determine monthly mean meteorological fields and can be regarded as a lower estimate the accuracy with which such fields can be determined. Since all the analyses used employ very similar data bases, it is likely that the true error is larger than the between the different analyses, since errors due to data gaps and measurement errors would be similar in the different analysis systems.

10. Summary

NCAR and NMC have collaborated to create a very long reanalysis using a frozen state-of-the-art global data assimilation system, and a data base as complete as pos Changes in the observing systems can still produce perceived changes in the analyzed climate, but this problem is approached by producing parallel reanalyses (at least o long) with and without using the new observing system for the period immediately aft introduction.

The system has been designed with advanced quality control and monitoring systems, and can produce one month of reanalysis per clock day on a CRAY YMP/8 supercomputer. Different types of output archives are being created for different u needs, including a "quick look" CD-ROM (one per year) archive with the most frequent used atmospheric fields, as well as surface, top of the atmosphere and isentropic fi

The output variables have been classified into four classes, depending on the degree to which they are influenced by the observations and/or the model. Users are cautioned that "C" variables (such as precipitation and surface fluxes) are complete determined by the model, forced by the data assimilation to remain close to the atmosphere. Nevertheless, a comparison of these variables with different types of observations and climatologies show generally useful information on time scales from few days to interannual variability.

11. Acknowledgments

We are very grateful to the Director of NMC, Ron McPherson, for his continued enthusiastic support for this project. Without the generous and significant help pr John Derber, David Parrish, Ken Campana, Robert Grumbine, Hua-lu Pan, Bert Katz, Jordan Alpert, Alan Basist, and Don Garrett, from NMC, Steve Worley and Chi-Fan Shih from NCAR, this project could not have been carried out. George Murphy and Sarah Ro have supported our computer resources needs. The consultation with the Cray Researc analysts Ron Bagby, George Vanderberg, Mary McCann and S. Wang, has also been very helpful. David Parker, (UKMO) has generously offered the sea surface temperature analysis pre-1982. Rex Gibson and Per Kallberg, (ECMWF), and Siegfried Schubert and David Lamich (NASA/GLA) have provided us with observations to fill data gaps. Sever countries, including Japan, China, Argentina, Brazil and others have contributed dat available in real time on GTS especially for this project. Prof. E. Kung (U. of Mis collaborated with the Chinese Meteorological Agency in recovering early chinese rawinsonde data. The COADS (surface marine) data set has many participants, includi Scott Woodruff (NOAA/ERL), Steve Worley (NCAR), Joe Elms (NCDC), Bob Keeley (Canada/MEDS). Other key NCAR participants in data preparation are Bob Dattore (raobs, aircraft), Wilbur Spangler (upper air), Gregg Walters (upper air), Ilana Ste (surface), Roy Barnes (surface, TD13), and Joey Comeau (TOVS). Dick Davis (NCDC) has helped with surface data and consulting on old rawinsonde data, and John Lanzant (GFDL) with the preparation of old raob dataset TD54. People in many countries and laboratories also deserve credit for data preparation. Key participants are the obs around the world who work day and night to take the observations that make these analyses possible.

The project has been supported since its inception by the NOAA Office for Globa Programs and by the National Weather Service (NMC and GFDL), and by the National Science Foundation (NCAR). Without this generous support we would not have been abl to develop this project. We are particularly grateful to Michael Coughlan and Ken M for their support and guidance.

The Advisory Committee, chaired by Julia Nogu s-Paegle from 1989 to 1993 and by Abraham Oort from 1993 to the present, has been a continuous source of advice and comfort when problems arose. Mark Cane, Julia Nogu s-Paegle, and Milt Halem origina suggested to perform a very long reanalysis, and J. Shukla spearheaded such a projec throughout the research community. We are grateful to them and to the other members the Panel, Maurice Blackmon, Donald Johnson, Per Kallberg, David Salstein, Siegfried Schubert, John Lanzante and James Hurrell.

Prof. I. M. Navon and two anonymous reviewers made very helpful suggestions and corrected many errors in the original manuscript.

Model level	mid-level sigma	delta sigma	thick- ness(m)	mand. press. level (hPa)
28	2.73	6.57	"ì"	3.0
27	10.06	7.29	5599.	10.0
26	18.34	9.23	3828.	20.0
25	28.75	11.60	3053.	30.0
24	41.79	14.51	2621.	
23	58.05	18.03	2342.	50.0
22	78.15	22.22	2142.	70.0
21	102.78	27.09	1984.	100.0
20	132.61	32.62	1851.	
19	168.23	38.67	1729.	150.0
18	210.06	45.03	1612.	200.0
17	258.23	51.35	1495.	250.0
16	312.48	57.16	1376.	300.0
15	372.05	61.97	1260.	400.0
14	435.68	65.26	1139.	
13	501.68	66.69	1017.	500.0
12	568.09	66.06	895.	
11	632.90	63.47	776.	
10	694.26	59.19	664.	700.0
9	750.76	53.72	560.	
8	801.42	47.54	466.	
7	845.79	41.15	384.	850.0

6	883.84	34.93	313.	
5	915.92	29.19	253.	925.0
4	942.55	24.05	203.	
3	964.37	19.59	162.	
2	982.08	15.82	129.	
1	995.00	10.00	80.	1000.0

Table 1: Model levels, mid-level sigma value, sigma thickness, geopotential thickne and approximate location of the mandatory pressure levels (hPa). Figure Legends

Fig. 1: Trace of the 1000 hPa virtual temperature averaged for the tropical Pacific the NMC operational Global Data Assimilation System (solid line), showing the impact changes in the model, and in the reanalysis (dotted line).

Fig. 2: Schematic of the main components of the NMC/NCAR Reanalysis System, and their state of readiness on January 1995. "Underway" means that the component is working but it still being improved.

Fig. 3: Schematic of the data input archive structure for the reanalysis. The data classified into eight basic types (rawinsondes, aircraft, land surface, marine, surf bogus, satellite temperature soundings, cloud-tracked winds, and SSM/I). The PREPBUFR archive adds "events" on each datum as it flows through the reanalysis.

Fig. 4: Global mean (top) radiative balance at the top of the atmosphere and (botto surface energy budget for 1985-91 from reanalysis. SH is sensible heat, LH latent h SW shortwave radiation, LW longwave radiation, OLR outgoing longwave radiation, D downward, U upward, N net, HF heat flux, RF radiative flux and Rad radiation. Values from reanalysis are compared to two climatological estimates: Ramanathan et al. (19 and Morel (1994).

Fig. 5: (top) Monthly mean globally averaged precipitation (P) and evaporation (E) (bottom)12 month running means of the net surface heat flux (SFC), net radiative fl the top of the atmosphere (TOA) and the net flux out of the atmosphere (TOTAL) for 1985-91 from reanalysis.

Fig. 6: a) Estimation of average precipitation during July and August 1988, July and August 1987, and their difference, accumulated in the 6 hour forecasts of the Reanal Contour lines at 2, 4, 8, 12, and 16 mm/day. b) As in a) but estimated from MSU an rain gauges. c) Soil moisture (percentage of total field capacity, 200 cm) and sea s temperatures averaged for August 1988, August 1987 and their difference.

Fig. 7: a) Evolution of the soil moisture content averaged over several regions of world during the first 5 years of reanalysis. The units are percentages of the tota capacity (200 cm). b) Average maximum soil moisture content estimated from seven yea of reanalysis. The arrows indicate the month at which the soil moisture is, on the maximum. c) As in b) but for the minimum soil moisture content. d) Difference betwe maximum and minimum soil moisture.

Fig. 8: Daily mean precipitation rates (mm/day) for May 1985-89 over the United Stat (a) the NMC/NCAR reanalysis and in (b) the observations. Standard deviation of the mean precipitation rates (mm/day) within May 1985-89 in (c) the NMC/NCAR reanalysis and (d) the observations. Contour interval, 1 mm/day, greater than 1 mm/day shaded.

Fig. 9: a) Hovmoeller diagram (longitude time) for the zonal wind component at 50 hP the Equator (5N-5S), in the operational NMC Global Data Assimilation System (from th Climate Diagnostics Data Base). b) As in a) but for the Reanalysis.

Fig. 10: Log pressure/time cross-section from the 1985-89 Reanalysis for a point nea Canton Island, from 100 to 10 hPa.

Table 1: Model levels, mid-level sigma value, sigma thickness, geopotential thickne and approximate location of the mandatory pressure levels (hPa).

Table 2: Comprehensive "synoptic format" archives. See Appendix B of the NMC Offi Note 401 for a detailed list of fields, units, etc., contained in each of the files

Table 3. Data fields on standard pressure levels to be saved on a 2.5 degree lat/lo (144 x 73).

Table 4. Surface flux data to be saved on the T62 Gaussian grid (192 x 94 grid poin

12. References

Campana K.A., Y-T Hou, K.E. Mitchell, S-K Yang, and R. Cullather, 1994: Improved diagnostic cloud parameterization in NMC's global model, 10th AMS Conferenc on Numerical Weather Prediction, Portland OR, 324-325.

Chelliah, M., 1994: An assessment of the monthly mean atmospheric fields from the NMC/NCAR Reanalysis Project. Proc. 19th Annual Climate Diagnostics Workshop, College Park, MD, Nov. 14-18, 1994. Available from NOAA/NMC CAC, Washington DC 20233.

Collins, W.G. and L.S. Gandin, 1992: Complex quality control of rawinsonde heights and temperatures (CQCHT) at the National at the National Meteorological Cent NMC Office Note 390. (Available from NOAA/NMC, Washington, D.C., 20233).

Collins, W.G. and L.S. Gandin, 1990: Comprehensive hydrostatic quality control the National Meteorological Center. Mon. Wea. Rev., 18, 2754-2767.

Derber, J.D., and A. Rosatti, 1989. A global oceanic data assimilation system. Phys. Oceanogr., 19, 1333-1347.

Derber, J. C., D. F., Parrish and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. Wea. and Forecasting, 6, 538

Derber, J., D. Parrish, W.-S. Wu, Z.-X. Pu and S. Rizvi, 1994: Improvements to operational SSI global analysis system. Preprints of the 10th AMS Conf. on Num. Weather Pred., July 18-22, 1994, Portland OR.

Development Division staff, 1988: Documentation of the NMC global model. Available from NOAA/NMC Development Division, Washington DC 20233.

Dorman, J.L., and P.J. Sellers, 1989: A global climatology of albedo, roughnes and stomatal resistance for atmospheric general circulation models as represented by Simple Biosphere Model (SB), J. Appl. Meteor., 28, 833-855.

Gandin, L.S., 1993: Optimal Averaging of Meteorological Fields. NMC Office Note 397. Available from NOAA/NMC Development Division, Washington DC 20233.

Gandin, L.S.: 1988: Complex quality control of meteorological observations. Mon Wea. Rev., 116, 1138-1156.

Finger, F.G., M. E. Gelman, J. D. Wild, M. L. Chanin, A. Hauchecorne, and A. J. Miller, 1993: Evaluation of NMC Upper-stratospheric temperature analyses using rocketsondes and lidar data. Bull. Amer. Meteor. Soc., 74, 789-799.

Goodberlet, M. A., C. T Swift and J.C. Wilkerson, 1989: Remote sensing of ocean surface winds with the special sensor microwave/imager. J. Geophys. Res. 94, 14547-14555.

Grell, G.A., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations. Mon. Wea. Rev., 121, 764-787.

IPCC (1990): Climate Change: The IPCC Scientific Assessment. WMO/UNEP International Panel on Climate Change.

Janowiak, J.E. and P. A. Arkin, 1991: Rainfall variations in the tropical Pacif during 1986-1989, as estimated from observations of cloud-top temperature. J. Geoph Res., 96 (Supplement), 3359-3373.

Janowiak, J., 1994: An evaluation of NMC/NCAR Reanalysis tropical rainfall duri 1986-88. Proc. 19th Annual Climate Diagnostics Workshop, College Park, MD, Nov. 14-1

1994. Available from NOAA/NMC CAC, Washington DC 20233.
Jenne, R. 1993: Various documents describing the data base (see section III). Available from R. Jenne, NCAR, PO Box 3000, Boulder CO.
Ji, M., A. Kumar, and A. Leetmaa, 1994: A multi season climate forecast system the National Meteorological Center. Bull. Am. Met. Soc., 75, 557-569.
Kalnay E., and R. Jenne, 1991: Summary of the NMC/NCAR Reanalysis Workshop of April 1991. Bulletin American Meteorological Society, 72, 1897-1904.
Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Ire S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, J. Janowiak, K. C C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, E. Kung and D. Salstein, 1993: The NMC/NCAR CDAS/Reanalysis Project. NMC Office Note 401. Available from NOAA/NMC Development Division, Washington DC 20233.
Kanamitsu, M., 1989: Description of the NMC global data assimilation and forec system. Wea. and Forecasting, 4, 334-342.
Kanamitsu, M., J.C. Alpert, K.A. Campana, P.M. Caplan, D.G. Deaven, M. Iredell, B. Katz, H.L. Pan, J. Sela and G.H. White, 1991: Recent changes implemented into th global forecast system at NMC. Weather and Forecasting, 6, 0001-0012.
Kistler, R., S. Saha and J. Woollen, 1994: The NMC/NCAR Reanalysis monitoring system. Proc. 19th Annual Climate Diagnostics Workshop, College Park, MD, Nov. 14-18 1994. Available from NOAA/NMC CAC, Washington DC 20233.
Krasnopolsky, V.M., L.C. Breaker, and W.H. Gemmill, 1995. A neural network as nonlinear transfer function model for retrieving surface wind speeds from the SSM/I. in press.
Legates, D.R., 1993: Global Observations, Analyses and Simulation of Precipitation, WGNE/GEWEX Workshop, NMC, Camp Springs, MD, USA, 27-30 Oct. 1992. World Meteorological Organization, Geneva, Switzerland, WMO/TD No. S44, 23-29
Levitus, S. 1982. Climatological Atlas of the World Ocean. DOC, NOAA Prof. Paper 13, U. S. Government Printing Office, Washington DC, 17 fiches, 173 pp.
Lorenc, A. C., 1981: A global three-dimensional multivariate statistical inter scheme. Mon. Wea. Rev., 109, 701-721.
Mahrt, L., and H.-L. Pan, 1984: A two-layer model of soil hydrology. Boundary Layer Meteor. 29,1-20.
Matthews, E., 1985: Atlas of archived vegetation, land-use and seasonal albedo data sets. NASA Tech. Memo. 86199. 53pp.
Morel, P., 1994: Scientific issues underlying the global energy and water cycl European Conference on the Global Energy and Water Cycle, London, England, 18-22 July 1994.
Mo, K.C., X. L. Wang, R. Kistler, M. Kanamitsu and E. Kalnay, 1995: Impact of satellite data on the CDAS-reanalysis system. Mon. Wea. Rev., 123, 124-139.
Pan, Hua-lu and Wan-Shu Wu,1994: Implementing a mass flux convective parameterization package for the NMC Medium-range forecast model. Preprints of the 10th AMS Conf. on Num. Weather Pred., July 18-22, 1994, Portland OR.
Pan, H.-L., and L. Mahrt, 1987: Interaction between soil hydrology and boundar layer development. Boundary Layer Meteor., 38, 185-220.
Parker, D. E. , C. K. Folland, M.N. Ward, K. Maskell, A. Bevan, and M. Jackson, 1993: Ocean-atmosphere climatic fluctuations on interannual to century time scales. of Hadley Centre for Climate Prediction and Research, UKMO, Bracknell, Berks., England.
Parrish, D. F. and J.C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. Mon. Wea.Rev.,120, 1747-1763.
Parrish, D. , J. Derber, J. Purser, W. Wu and Z. Pu, 1995: The NMC Global Analysis System: recent improvements and future plans. Journal of the Meteorologica Society of Japan.
Ramanathan, V., B.R. Bartstrom and E.F. Harrison, 1989: Climate and Earth's radiation budget. Physics Today, 42, 22-32.
Rasmusson, E. M. and T. H. Carpenter, 1983: The relationship between eastern equatorial Pacific sea surface temperatures and rainfall over India and Sri Lanka. Wea. Rev., 110, 517-528.
Reed, R. and D.G. Rogers, 1962: Circulation of the tropical atmosphere in the years 1954-1960. J. Atmos. Sciences, 19,127-135.
Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. J

Climate, 1, 75-86.
Reynolds, R. W., 1993: Impact of Mount Pinatubo Aerosols on Satellite-Derived Sea Surface Temperatures. J. Climate, 6, 768-774.
Reynolds, R. W. and D. C. Marsico, 1993: An improved real-time global sea surface temperature analysis. J. Climate, 6, 768-774.
Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. J. Climate, 7,929-948.
Ropelewski, C. F. and M. S. Halpert, 1987: Global and regional scale precipitat patterns associated with the El Ni¤o/Southern Oscillation cycle. Mon. Wea. Rev., 11 1606-1626.
Ropelewski, C. F. and M. S. Halpert, 1989: Precipitation patterns associated wi the high index phase of the Southern Oscillation. J. Climate, 2, 268-284.
Saha, S., and M. Chelliah, 1993: Automatic Monitoring system for the atmospheri Reanalysis project at NMC. Proc. of Clim. Diagn. Workshop, Boulder, CO 1-5 Nov 1993.Available from NOAA/NMC CAC, Washington DC 20233.
Saha, S., M. Kanamitsu, M. Chelliah, and C. Ropelewski, 1994: Surface precipitation over the US from the NMC/NCAR Reanalysis. Proc. 19th Annual Climate Diagnostics Workshop, College Park, MD, Nov. 14-18, 1994.
Salstein, D., 1993: Temporal and spatial resolution impact on some climate anal diagnostics. Proc. Clim. Diagn. Workshop, Norman, OK, 18-23 Oct. 1993. Available f NOAA/NMC CAC, Washington DC 20233.
Schubert, S., R. Rood and J. Pfaendtner, 1993: An assimilated data set for Eart Sciences applications. Bull. Am. Met. Soc., 74, 2331-2342.
Shukla, J. and D. A. Paolino, 1983: The Southern Oscillation and long-range forecasting of the summer monsoon rainfall over India. Mon. Wea. Rev., 111, 1830-18
Smith, T. M., 1994: Preliminary analysis of the annual cycle of vertical heat f the Pacific from the NMC/NCAR Reanalysis. Proc. 19th Annual Climate Diagnostics Workshop, College Park, MD, Nov. 14-18, 1994. Available from NOAA/NMC CAC, Washington DC 20233.
Spencer, R. W., 1993: Global oceanic precipitation from the MSU during 1979-199 and comparisons to other climatologies. J. Clim. 6, 1301-1326
White, G., 1994: Diagnostics of atmospheric forcing from reanalysis. Proceedi of 19th Annual Climate Diagnostics Workshop, College Park, MD. Available from NOAA/NMC CAC, Washington DC 20233.
Woollen, J. S., 1991: New NMC operational OI quality control. Ninth Conferenc On Numerical Weather Prediction, American Meteorological Society, 24-27.
Woollen, J. S., E. Kalnay, L. Gandin, W. Collins, S. Saha, R. Kistler, M. Kanam and M. Chelliah, 1994: Quality Control in the Reanalysis System. Tenth Conference O Numerical Weather Prediction, American Meteorological Society, 13-14.

Appendix A: NMC/NCAR Reanalysis Comprehensive Output Variables

The output variables are classified into four categories, depending on the rela influence of the observational data and the model on the gridded variable. An A ind that the analysis variable is strongly influenced by observed data, and hence it is most reliable class (e.g., upper air temperature and wind). The designation B indica that, although there are observational data that directly affects the value of the v the model also has a very strong influence on the analysis value (e.g., humidity, an surface temperature). The letter C indicates that there are no observations directl affecting the variable, so that it is derived solely from the model fields forced by assimilation to remain close to the atmosphere (e.g., clouds and precipitation). Fi letter D represents a field that is fixed from climatological values, and does not d the model (e.g., plant resistance, land-sea mask). Appendix A contains the complete classification of variables. Although this classification is necessarily somewhat the user should exercise caution in interpreting the results of the reanalysis, espe variables classified in categories B and C. In addition to this rule of thumb, the should keep in mind that quadratic variables (e.g., kinetic energy, transport of wat are in general less reliable than the components from which they were computed.

1. Standard GRIB output

1.1. Pressure: Pressure coordinate output

... Regular latitude-longitude grid (2.5o x 2.5o)
... All fields are instantaneous values at a given time

A	Geopotential height (gpm) at 17 levels
A	u-wind (m/s) 17 levels
A	v-wind (m/s) 17 levels
A	Temperature (K) 17 levels
B	Pressure vertical velocity (Pa/s) 12 levels
B	Relative humidity (%) 8 levels
A	Absolute vorticity (/s) 17 levels
A	u-wind of the lowest 30 hPa layer (m/s)
A	v-wind of the lowest 30 hPa layer (m/s)
B	Temperature of the lowest 30 hPa layer (K)
B	Relative humidity of the lowest 30 hPa (%)
B	Pressure at the surface (Pa)
B	Precipitable water (kg/m2)
B	Relative humidity of the total atmospheric column (%)
A	Temperature at the tropopause (K)
A	Pressure at the tropopause (Pa)
A	u-wind at the tropopause (m/s)
A	v-wind at the tropopause (m/s)
A	Vertical speed shear at the tropopause (1/s)
B	Surface lifted index (K)
B	"Best" (4-layer) lifted index (K)
A	Temperature at the maximum wind level (K)
A	Pressure at the maximum wind level (Pa)
A	u-wind at the maximum wind level (m/s)
A	v-wind at the maximum wind level (m/s)
D	Geopotential height at the surface (gpm)
A	Pressure reduced to MSL (Pa)
B	Relative humidity in 3 sigma layers:0.44-0.72,0.72-0.94,0.44-1.0 (%)
B	Potential temperature at the lowest sigma level (K)
B	Temperature at the lowest sigma level (K)
B	Pressure vertical velocity at the lowest sigma level (Pa/s)
B	Relative humidity at the lowest sigma level (%)
B	u-wind at the lowest sigma level (m/s)
B	v-wind at the lowest sigma level (m/s)

1.2. Grb2d...2-dimensional diagnostic file

C	Cloud forcing net longwave flux at the top of atmosphere (W/m2)
C	Cloud forcing net longwave flux at the surface (W/m2)
C	Cloud forcing net longwave flux for total atmospheric column (W/m2)
C	Cloud forcing net solar flux at the top of the atmosphere (W/m2)
C	Cloud forcing net solar flux at the surface (W/m2)
C	Cloud forcing net solar flux for total atmospheric column (W/m2)
C	Convective precipitation rate (kg/m2/s)
C	Clear sky downward longwave flux at the surface (W/m2)
C	Clear sky downward solar flux at the surface (W/m2)
C	Clear sky upward longwave flux at the top of the atmosphere (W/m2)
C	Clear sky upward solar flux at the top of atmosphere (W/m2)
C	Clear sky upward solar flux at the surface (W/m2)
C	Cloud work function (J/Kg)
C	Downward longwave radiation flux at the surface (W/m2)
C	Downward solar radiation flux at the top of the atmosphere (W/m2)
C	Downward solar radiation flux at the surface (W/m2)
C	Ground heat flux (W/m2)
D	Ice concentration (ice=1;no ice=0) (1/0)
D	Land-sea mask (1=land;0=sea) (integer)

```
C    Latent heat flux (W/m2)
C    Near IR beam downward solar flux at the surface (W/m2)
C    Near IR diffuse downward solar flux at the surface (W/m2)
C    Potential evaporation rate (w/m2)
C    Precipitation rate (kg/m2/s)
C    Pressure at high cloud top (Pa)
C    Pressure at high cloud base (Pa)
C    Pressure at middle cloud top (Pa)
C    Pressure at middle cloud base (Pa)
C    Pressure at low cloud top (Pa)
C    Pressure at low cloud base (Pa)
C    Pressure at the surface (Pa)
C    Run off (kg/m2 per 6 hour interval)
D    Surface roughness (m)
C    Nearby model level of high cloud top (integer)
C    Nearby model level of high cloud base (integer)
C    Nearby model level of middle cloud top (integer)
C    Nearby model level of middle cloud base (integer)
C    Nearby model level of low cloud top (integer)
C    Nearby model level of low cloud base (integer)
C    Sensible heat flux (W/m**2)
C    Volumetric soil moisture content (fraction) (2 layers)
B    Specific humidity at 2m (kg/kg)
C    Total cloud cover of high cloud layer (%)
C    Total cloud cover of middle cloud layer (%)
C    Total cloud cover of low cloud layer (%)
B    Maximum temperature at 2m (K)
B    Minimum temperature at 2m (K)
AB   Temperature at the surface (skin temperature) (K)
C    Temperature of the soil layer (3 layers) (K)
B    Temperature at 2m (K)
C    Temperature of high cloud top (K)
C    Temperature of low cloud top (K)
C    Temperature of middle cloud top (K)
C    Zonal gravity wave stress (N/m2)
B    Zonal component of momentum flux (N/m2)
B    u-wind at 10m (m/s)
C    Upward longwave radiation flux at the top of the atmosphere (W/m2)
C    Upward longwave radiation flux at the surface (W/m2)
C    Upward solar radiation flux at the top of the atmosphere (W/m2)
C    Upward solar radiation flux at the surface (W/m2)
C    Meridional gravity wave stress (N/m2)
C    Visible beam downward solar flux at the surface (W/m2)
C    Visible diffuse downward solar flux at the surface (W/m2)
C    Meridional component of momentum flux (N/m2)
B    v-wind at 10m (m/s)
C    Water equivalent of accum. snow depth (kg/m2)
```

1.3. Grb3d...3-dimensional diagnostic file

```
... Gaussian grid (192 x 94) on 28 model levels
... All fields are average of 6 hour integration starting from a given time

C    Deep convective heating rate (K/s)
C    Deep convective moistening rate (kg/kg/s)
C    Large scale condensation heating rate (K/s)
C    Longwave radiative heating rate (K/s)
C    Shallow convective heating rate (K/s)
C    Shallow convective moistening rate (kg/kg/s)
C    Solar radiative heating rate (K/s)
```

C   Vertical diffusion heating rate (K/s)
C   Vertical diffusion moistening rate (kg/kg/s)
C   Vertical diffusion zonal accel. (m/s/s)
C   Vertical diffusion meridional accel. (m/s/s)

1.4. Sigma

... Gaussian grid (192 x 94) on 28 model levels or surface
... All fields are instantaneous values at a specified time

A   Relative vorticity (28 levels) (/s)
B   Divergence (28 levels) (/s)
A   Temperature (28 levels) (K)
B   Specific humidity (28 levels) (kg/kg)
A   x-gradient of log pressure (surface) (1/m)
A   y-gradient of log pressure (surface) (1/m)
A   u-wind (28 levels) (m/s)
A   v-wind (28 levels) (m/s)
A   Pressure (surface) (Pa)
A   Geopotential height (surface) (gpm)
A   x-gradient of height (surface) (m/m)
A   y-gradient of height (surface) (m/m)

1.5. Isen..Isentropic coordinate output

... Gaussian grid (192 x 94) most on 10 isentropic levels
... All fields are instantaneous values at a specified time

A   Potential temperature (surface) (K)
A   Temperature (K)
A   u-wind (m/s)
A   v-wind (m/s)
B   Pressure vertical velocity (Pa/s)
B   Relative humidity (%)
A   Montgomery stream function (m2/s2)
B   Brunt-Vaisala frequency squared (1/s2)
B   Potential vorticity (m2/s/kg)

2. Other non-GRIB output files

2.1 Zonal diagnostic file (binary)

... Average over 90S-60S, 60S-30S, 30S-30N, 30N-60N, 60N-90N and global
... Unmarked fields are instantaneous values at a given time
... (Av) indicates average during the 6 hour integration

A   u component of wind (m/s) at 28 model levels
A   v component of wind (m/s) at 28 model levels
A   virtual temperature (K) at 28 model levels
B   specific humidity (g/g) at 28 model levels
B   squared vorticity (1/s2) at 28 model levels
C   squared divergence (1/s2) at 28 model levels
B   pressure vertical velocity (Pa/s) at 28 model levels
A   temperature (K) at 28 model levels
B   relative humidity (%) at 28 model levels
B   kinetic energy (m2/s2) at 28 model levels
C   convective heating (K/s) at 28 model levels (Av)
C   large scale heating (K/s) at 28 model levels (Av)

```
C    shallow convection heating (K/s) at 28 model levels (Av)
C    vertical diffusion heating (K/s) at 28 model levels (Av)
C    convective moistening (g/g/s) at 28 model levels (Av)
C    shallow convection moistening (g/g/s) at 28 model levels (Av)
C    vertical diffusion moistening (g/g/s) at 28 model levels (Av)
C    zonal accel by vertical diffusion (m/s2) at 28 model levels (Av)
C    meridional accel by vertical diffusion (m/s2) at 28 model levels
C    short wave radiation heating (K/s) at 28 model levels (Av)
C    long wave radiation heating (K/s) at 28 model levels (Av)
C    total precipitation (Kg/m2) (Av)
C    convective precipitation (Kg/m2) (Av)
C    sensible heat flux (w/m2) (Av)
C    latent heat flux (w/m2) (Av)
B    zonal stress (dyn/m2) (Av)
B    meridional stress (dyn/m2) (Av)
C    rain area coverage (%)
C    convective rain area coverage (%)
B    surface pressure (hPa)
C    surface skin temperature (K)
C    soil wetness (cm)
C    snow depth (m)
C    10 cm deep soil temperature (K)
C    50 cm deep soil temperature (K)
D    500 cm deep soil temperature (K)
C    surface net short wave flux (W/m2) (Av)
C    surface net long wave flux (W/m2) (Av)
B    relative humidity at the lowest model level (%)
B    virtual temp at the lowest model level (K)
B    temperature at the lowest model level (K)
B    specific humidity at the lowest model level (K)
D    surface roughness (m)
D    land sea sea-ice mask (int)
C    zonal accel by gravity wave drag (m/s2) (Av)
C    meridional accel by gravity wave (m/s2)  (Av)
B    surface torque (g/m2/s2) (Av)
C    gravity wave drag torque (g/m2/s2)  (Av)
B    mountain torque (g/m2/s2) (Av)
B    total angular momentum (m2/s)
B    planetary angular momentum (m2/s)
```

3. RESTART FILES (binary)

... Spectral (28 model levels) or Gaussian grid (192 x 94)
... All fields are instantaneous values at a specified time

3.1. Sigma spectral coefficient file

```
D    Surface geopotential
B    Natural log of surface pressure
A    Virtual temperature
B    Divergence
A    Vorticity
B    Specific humidity
```

3.2. Surface file (on Gaussian grid)

```
C    earth surface temperature (K)
     C    soil moisture level 1 (% volume)
C    soil moisture level 2 (% volume)
C    snow depth (m)
```

C	soil temperature level 1 (K)
C	soil temperature level 2 (K)
C	soil temperature level 3 (K)
D	surface roughness length (m)
C	convective cloud cover (%)
C	convective cloud bottom height (sigma)
C	convective cloud top height (sigma)
C	albedo (fraction)
C	snow/ice/land mask
D	minimum stomatal resistance (s/m)
C	canopy water content (m)
C	ratio of 10m and lowest sigma level winds (fraction)

Appendix B: Output levels

Standard Pressure levels (hPa):

1000	925	850	700	600	500	400	300	250	100	150
70	50	30	20	10						

Isentropic surfaces (oK):

650	550	450	400	350	330	315	300	290	280	270 K

Sigma levels:

0.9950	0. 9821	0.9644	0.9425	0.9159	0.8838	0.8458	0.8014	0.7508
0.6943	0.6329	0.5681	0.5017	0.4357	0.3720	0.3125	0.2582	0.2101
0.1682	0.1326	0.1028	0.0782	0.0580	0.0418	0.0288	0.0183	0.0101
0.0027								

Appendix C: NMC/NCAR Reanalysis output on CD-ROMs

Some of the reanalysis products will be distributed on CD-ROMs. Currently two types of CD-ROMs are being planned. The first would contain reanalysis products for single year (1 CD-ROM per year). The second type would be produced after about 10 years and would contain time series of relatively few variables. We believe this is efficient way to satisfy the requirements of most members of the meteorological community, many of which were consulted in the preparation of the output list. The following is the plan for the first type of CD-ROMs.

Note that the output variables should be classified into four categories, depen on the relative influence of the observational data and the model on the gridded var (see Appendix A for a complete classification). The user should exercise caution interpreting the results of the reanalysis, especially for variables classified into B and C.

00Z and 12Z analyses:	estimated size
U, V, temp. at 850, 500, 200 hPa	114 MB
geopotential height 1000, 850, 700, 500, 300, 200 hPa	29 MB
(sea level pressure can derived from above fields)	
surface pressure	21 MB

omega at 500 mb	10 MB
precipitable water	10 MB
temperature at 2 m	18 MB
specific humidity at 2 m	14 MB
U, V at 10 m	29 MB
RH at 500, 200 hPa	13 MB
Total for 00 and 12Z analyses	258 MB/Yr

Daily averaged analyses:

zonal, meridional wind stress	19 MB
net short/long wave flux at surface	14 MB
precipitation	8 MB
latent/sensible heat flux	16 MB
Model OLR	7 MB
downward short wave flux at surface	7 MB
outgoing short wave flux at top	7 MB
Tmin, Tmax (24 hour period)	17 MB
skin temperature (includes SST)	9 MB
snow (liquid water equivalent)	13 MB
Total for daily averaged fields	117 MB/Yr

OOZ (isentropic) and 12Z (stratosphere) analyses:

height, temperature at 100, 50 and 20 hPa	29 MB
U, V at 100, 50 and 20 hPa	28 MB
potential vorticity on 3 é surfaces (315, 330, 450 K)	15 MB
U, V on 3 theta surfaces (315, 330, 450 K)	30 MB
temperature on 3 theta surfaces (315, 330, 450 K)	14 MB
Total for isentropic and stratospheric analyses	117 MB/Yr

All cross-sections (monthly averaged)	2 MB/Yr
Monthly means, variances and covariances	138 MB/Yr
Observed OLR	5 MB/Yr
GrADS control and index files	24 MB/Yr

Documentation:
BAMS paper (Office Note 401 with updates)
Office Note 388, (GRIB table of local definitions, documentation)
Miscellaneous

Total documentation volume	3 MB
Software to read grib (PC-GrADS, wgrib)	6 MB
ESTIMATED TOTAL VOLUME	670 MB/Yr

Appendix D: List of Acronyms

AVHRR: Advanced Very High Resolution Radiometer
BUFR: Binary Universal Format Representation

CDAS: Climate Data Assimilation System
CDC: Climate Diagnostics Center
CD-ROM: Compact Disc - Read Only Memory
COADS: Comprehensive Ocean Atmosphere Data Set
COLA: Center for Ocean, Land and Atmosphere
CQC: Complex Quality Control
DOE: Department of Energy
ECMWF: European Centre for Medium-range Weather Forecasts
ENSO: El Ni¤o-Southern Oscillation
ERL: Environmental Research Laboratories
FGGE: First GARP Global Experiment (1979)
FTP: File Transfer Protocol
GARP: Global Atmospheric Research Program
GATE: GARP Atmospheric Tropical Experiment
GB: Gigabytes (1012 bytes)
GDAS: Global Data Assimilation System
GFDL: Geophysical Fluid Dynamics Laboratory
GISST: Global Ice and Sea Surface Temperature data set
GLA: Goddard Laboratory for Atmospheres
GrADS: Grid Analysis and Display System
GRIB: GRidded Binary format
GTS: Global Telecommunications System
HIRS: High Resolution Infrared Sounder
hPa: hecto Pascals (also pronounced "milli bars")
IPCC: Intergovernmental Panel on Climate Change
IR: Infrared
ITCZ: InterTropical Convergence Zone
JMA: Japan Meteorological Agency
MEDS (Canada)
MSU: Microwave Sounding Unit
NCAR: National Center for Atmospheric Research
NCDC: National Climate Data Center
NCEP: National Centers for Environmental Modeling (formerly NMC)
NESDIS: National Environmental Satellite, Data and Information Service
NH: Northern Hemisphere
NMC: National Meteorological Center (now NCEP)
NOAA: National Oceanic and Atmospheric Administration
NSF: National Science Foundation
OA: Optimal Averaging
OGP: Office of Global Programs
OI: Optimal Interpolation
OLR: Outgoing Long-wave Radiation
OIQC: OI-based Quality Control
QBO: Quasi-biennial Oscillation
QC: Quality Control
SiB: Simple Biosphere Model
SH: Southern Hemisphere
SIRS: Satellite Infrared Spectrometer
SSI: Spectral Statistical Interpolation (also known as a 3-D VAR scheme)
SSM/I: Special Sounding Microwave/Imager
SMMR: Scanning Multichannel Microwave Radiometer
SST: Sea Surface Temperature
SSU: Stratospheric Sounding Unit
T62: Triangular 62 waves truncation
TB: Tera byte (1015 bytes)
TOVS: TIROS-N Operational Vertical Sounder
TWERLE: Tropospheric Wind Earth Radio Location Experiment
UKMO: United Kingdom Meteorological Office
USAF: United States Air Force
VTPR: Vertical Temperature and Pressure Radiometer
WMO: World Meteorological Organization

XBT: Bathythermograph