THE NCEP CLIMATE FORECAST SYSTEM REANALYSIS

BY SURANJANA SAHA, SHRINIVAS MOORTHI, HUA-LU PAN, XINGREN WU, JIANDE WANG, SUDHIR NADIGA, PATRICK TRIPP, ROBERT KISTLER, JOHN WOOLLEN, DAVID BEHRINGER, HAIXIA LIU, DIANE STOKES, ROBERT GRUMBINE, GEORGE GAYNO, JUN WANG, YU-TAI HOU, HUI-YA CHUANG, HANN-MING H. JUANG, JOE SELA, MARK IREDELL, RUSS TREADON, DARYL KLEIST, PAUL VAN DELST, DENNIS KEYSER, JOHN DERBER, MICHAEL EK, JESSE MENG, HELIN WEI, RONGQIAN YANG, STEPHEN LORD, HUUG VAN DEN DOOL, ARUN KUMAR, WANQIU WANG, CRAIG LONG, MUTHUVEL CHELLIAH, YAN XUE, BOYIN HUANG, JAE-KYUNG SCHEMM, WESLEY EBISUZAKI, ROGER LIN, PINGPING XIE, MINGYUE CHEN, SHUNTAI ZHOU, WAYNE HIGGINS, CHENG-ZHI ZOU, QUANHUA LIU, YONG CHEN, YONG HAN, LIDIA CUCURULL, RICHARD W. REYNOLDS, GLENN RUTLEDGE, AND MITCH GOLDBERG

A new coupled global NCEP Reanalysis for the period 1979–present is now available, at much higher temporal and spatial resolution, for climate studies.

AFFILIATIONS: SAHA, MOORTHI, PAN, BEHRINGER, STOKES, GRUMBINE, HOU, CHUANG, JUANG, SELA, IREDELL, TREADON, KEYSER, DERBER, EK, AND LORD—Environmental Modeling Center, NCEP/NWS/NOAA, Camp Springs, Maryland; VAN DEN DOOL, KUMAR, W. WANG, LONG, CHELLIAH, XUE, SCHEMM, EBISUZAKI, XIE, CHEN, AND HIGGINS—Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, Maryland; WU, JI. WANG, NADIGA, KISTLER, WOOLLEN, H. LIU, GAYNO, JU. WANG, KLEIST, VAN DELST, MENG, WEI, AND YANG—Science Applications International Corporation (SAIC), McLean, Virginia; ZOU, CHEN, HAN, CUCURULL, AND GOLDBERG—Center for Satellite Applications and Research, NESDIS/NOAA, Camp Springs, Maryland; Q. LIU—Joint Center for Satellite Data Assimilation, NASA–NOAA, Camp Springs, Maryland; RUTLEDGE—NOAA/NESDIS/National Climatic Data Center, Asheville, North Carolina; TRIPP—I.M. Systems Group, Inc., Rockville, Maryland; REYNOLDS—Cooperative Institute for Climate & Satellites, NOAA, and North Carolina State University, Raleigh, North Carolina; HUANG, LIN, AND ZHOU—Wyle Information System, McLean, Virginia
CORRESPONDING AUTHOR: Dr. Suranjana Saha, NOAA Science Center, 5200 Auth Road, Camp Springs, MD 20746
E-mail: suranjana.saha@noaa.gov

DOI: 10.1175/2010BAMS3001.1

A supplement to this article is available online (10.1175/2010BAMS3001.2)
In final form 12 April 2010

The first reanalysis at NCEP (all acronyms are defined in the appendix), conducted in the 1990s, resulted in the NCEP–NCAR reanalysis (Kalnay et al. 1996), or R1 for brevity, and ultimately covered many years, from 1948 to the present (Kistler et al. 2001). It is still being executed at NCEP, to the benefit of countless users for monthly, and even daily, updates of the current state of the atmosphere. At the same time, other reanalyses were being conducted, namely, ERA-15 (Gibson et al. 1997) was executed for a more limited period (1979–93) at the ECMWF, COLA conducted a short reanalysis covering the May 1982–November 1983 period (Paolino et al. 1995), and NASA GSFC conducted a reanalysis covering the 1980–94 period (Schubert et al. 1997). The general purpose of conducting reanalyses is to produce multiyear global state-of-the-art gridded representations of atmospheric states, generated by a constant model and a constant data assimilation system. To use the same model and data assimilation over a very long period was the great advance during the 1990s, because gridded datasets available before 1995 had been created in real time by ever-changing models and analysis methods, even by hand analyses prior to about 1965. The hope was that a reanalysis,

made after real time, would help in advancing climate studies by eliminating fictitious trends caused by model and data assimilation changes in real time. Although "climate" was the consideration of overriding importance in justifying the enormous effort involved in conducting a multiyear reanalysis, the techniques used were those of NWP. This climate–weather linkage lives on to this day. In particular, the analysis at any given time (t) is the result of a short forecast (the guess field), initialized from a previous analysis (valid at time $t - \Delta t$), modified by assimilating new observations available in a narrow window centered at t. The increment Δt is 6 h by convention, but it could, in principle, be either shorter or longer.

One very time consuming, but rewarding, part of the first reanalyses was to assemble and convert older observational datasets into BUFR, an internationally agreed-upon format to store and exchange observational data. The data mining and preparation of these datasets is a lasting legacy, and they have been exchanged among national and international partners and used in several more reanalyses (Uppala et al. 2005; Onogi et al. 2007; Bosilovich 2008). Every new reanalysis benefits from all previous reanalyses, and it is likely that mistakes are discovered and corrected to the benefit of the next user. This truly has been an ongoing activity in the international arena.

Errors, both big and small, were made in R1. Some were discovered and documented (available online at www.cpc.noaa.gov/products/wesley/reanalysis.html#problem).

Kanamitsu et al. (2002) executed a corrected version of R1, often called R2, covering only the satellite era from 1979 to the present. Importantly, in spite of correcting many errors and updating some components of the system, "only minor differences are found between R1 and R2 in the primary analysis variables, such as free atmospheric geopotential height and winds in the Northern Hemisphere extratropics" (Kanatmitsu et al. 2002). Both R1 and R2 are being continued at NCEP, an activity sometimes referred to as the CDAS, which respectively is noted as CDAS1 for R1 and CDAS2 for R2. Running a CDAS implies that certain technologies and computer algorithms are frozen in time, which in the face of ever-changing data ingest and computer configurations can be a challenge, and is ultimately impossible.

ECMWF has conducted two subsequent reanalyses: ERA-40 (Uppala et al. 2005) and ERA-Interim. The JMA conducted JRA-25 (Onogi et al. 2007), while GSFC is conducting MERRA (Bosilovich 2008; Schubert et al. 2008). A 100-yr reanalysis using only surface pressure data is being conducted at NOAA/ESRL (Compo et al. 2006). It is worth noting that the analysis system used in CFSR for the atmosphere, the GSI scheme, is nearly the same as the one used by MERRA at the NASA GSFC. The MERRA atmosphere-only reanalysis is being conducted over the same years with nearly the same input data. Obviously, the near-parallel development and execution of the reanalyses (CFSR and MERRA) can provide value-added results. The new reanalysis (ERA-Interim) executed presently at the ECMWF will be kept up to date (which was not the case for ERA-15 and its successor, ERA-40), thereby increasing its utility for real-time applications.

In this paper we only discuss global reanalyses. There has been one regional reanalysis conducted at NCEP, namely, NARR (see Mesinger et al. 2006), but this type of activity is outside the scope of this paper. However, it is worth noting that the CFSR has a global horizontal resolution of ~38 km, which nearly matches the 32-km resolution of NARR over the limited area of North America.

It has become increasingly clear over the last two decades that a single reanalysis will not suffice. Not only is the data ingest subject to continual improvements, but the models used to create the guess field improve continually, and so too do the data assimilation systems. Bengtsson et al. (2007) advocate the idea of a permanent, or ongoing, analysis of the Earth system. These efforts would not only utilize state-of-the-art forecast and data assimilation methods, but, in some configurations, the observations may even be reduced to whatever was available continuously over a century. Such kinds of reanalyses would be homogeneous over time (Compo et al. 2006) and could be used for climate studies at longer time scales.

New data assimilation techniques have been introduced since the mid-1990s, including 3DVAR, 4DVAR, and ensembles of analyses (e.g., EnKF), which produce not only an ensemble mean analysis but also a measure of the uncertainty (or spread; see Compo et al. 2006; Anderson et al. 2009). Although reanalyses may be primarily intended for climate studies, they also generate initial states that can be used to start integrations by a numerical prediction model, either for weather or climate, coupled to the ocean or not, in real time or from some historical state. Hindcasts (sometimes called reforecasts or retrospective forecasts) that are used to calibrate subsequent real-time forecasts, generated by the

same model, have been under considerable development since the 1990s, and the availability of a reanalysis is necessary for these efforts to succeed (Hamill et al. 2006). For instance, when NCEP operationally implemented its first global coupled seasonal forecast system, the CFS version 1 (Saha et al. 2006) in August 2004, a complete reforecast dataset was created over the period from 1981 to the present to calibrate the real-time operational forecasts. The initial states for this reforecast effort were taken from R2 for the atmosphere and land, and from the GODAS (Behringer and Xue 2004; Behringer 2007), which is forced by fluxes from R2, for the ocean. This relatively quick effort would have been inconceivable without R2 being available.

Since the CFS version 1 implementation in 2004 (Saha et al. 2006), CPC and many other users have utilized the 4-times-daily seasonal integrations for their monthly and seasonal forecast products. However, the system has many internal inconsistencies. For instance, the R2 atmospheric initial states are made with technology from the 1990s, while the atmospheric model component of CFS is from a decade later. Thus, the initial states and forecast model are inconsistent, which is a situation that leads to loss of skill during the early part of the integrations. One of the major goals of executing the new CFSR was to create initial states for the atmosphere, ocean, land, and sea ice that are both state of the art and as consistent as possible with the next version of the CFS version 2, which is to be implemented operationally at NCEP in 2010. Given the pace of model and data assimilation development, such a reanalysis will be needed roughly every 5–10 yr.

We have now completed the CFSR for the 31-yr period of 1979–2009. It took almost 2 yr to accomplish this feat. The primary novelties of this latest reanalysis are i) the coupling to the ocean during the generation of the 6-h guess field, ii) an interactive sea ice model, and iii) the assimilation of satellite radiances for the entire period. In addition, the much higher horizontal and vertical resolution (T382L64) of the atmosphere, model, and assimilation improvements over the last 10–15 yr, and the use of prescribed CO_2 concentrations as a function of time, should make for substantial improvements over R1 and R2 (which were at T62L28 resolution). Another major advance was the real-time monitoring that took place during the execution of the CFSR. Thousands of graphical plots were generated automatically at the end of each reanalyzed month and were displayed on the CFSR Web site in real time. Many scientists from both CPC and EMC monitored different aspects of the reanalysis during this 2-yr process. There were many times that the reanalysis was halted to address concerns that something may have gone wrong, and many corrections, backups, and restarts were made to ensure that the process was done correctly and homogeneously. This extremely large "atlas" of plots depicting nearly all aspects of the CFSR is open to the public (available online at http://cfs.ncep.noaa.gov/cfsr).

The layout of the paper is as follows. There are sections devoted to an overview, the observations, the atmosphere, the ocean, the sea ice, the coupler, and the land. Two final sections address a *preliminary* analysis of the CFSR, which resulted from the ongoing monitoring. Far more *substantial* analyses of CFSR will be forthcoming in a dozen or so journal articles that are currently under consideration. In the concluding section, we look ahead to the completion of the CFSRR project, namely, the reforecasts over the 1982–present period that will be initialized by the CFSR. This project is currently in progress. We will also allude to the possibility of going back to 1948, or further, with a reduced resolution of the CFSR system, such that both R1 and R2 have a single modern successor, and the old technologies can be retired gracefully. A description of some of the data that are available for distribution is given in the online supplemental version of this paper (online at http://dx.doi.org/10.1175/2010BAMS3001.2; hereafter referred to as "the supplement"). Information about data access is also given in the supplement. It must be noted here that data from this reanalysis have been archived at unprecedented spatial (0.5° × 0.5°), vertical (37 pressure levels for the atmosphere and 40 levels for the ocean), and temporal (hourly) resolution, exceeding over 250 TB.

OVERVIEW. Before presenting details about all subcomponents in CFSR in the following sections, the reader needs some overview of the plan and progress of the CFSR project. In the "Execution" section, we describe the execution of the CFSR with parallel streams. In the "Peculiarities of some of the data ingest" section, we confront some aspects of the manner in which some of the data were assimilated, which may not seem logical to the uninitiated. In the "R1 versus CFSR" section, we present a short comparison between R1 and CFSR.

Execution. To ensure the project would be completed in a 2-yr period, the CFSR was produced by running

six simultaneous streams of analyses, covering the following periods:

Stream 1: 1 December 1978 to 31 December 1986
Stream 2: 1 November 1985 to 31 December 1989
Stream 5: 1 January 1989 to 31 December 1994
Stream 6: 1 January 1994 to 31 March 1999
Stream 3: 1 April 1998 to 31 March 2005
Stream 4: 1 April 2004 to 31 December 2009

As can be seen, there is a full 1-yr overlap between the streams to address spinup issues concerning the deep ocean, the upper stratosphere, and the deep soil. Thus, the entire CFSR covers 31 yr (1979–2009), plus five overlap years.

Figure 1 (upper half) shows the CFSR execution of one day of reanalysis, which can be itemized as follows:

- Atmospheric T382L64 (GSI) analysis is made at 0000, 0600, 1200, and 1800 UTC, using a coupled 9-h guess forecast.
- Ocean and sea ice analysis (GODAS with MOM4) is made at 0000, 0600, 1200, and 1800 UTC, using the same 9-h coupled guess forecast.
- From each of the four cycles, a 9-h coupled guess forecast (GFS at T382L64) is made with 30-min coupling to the ocean (MOM version 4).
- Land (GLDAS) analysis, using observed precipitation with the Noah land model, is made only at 0000 UTC.

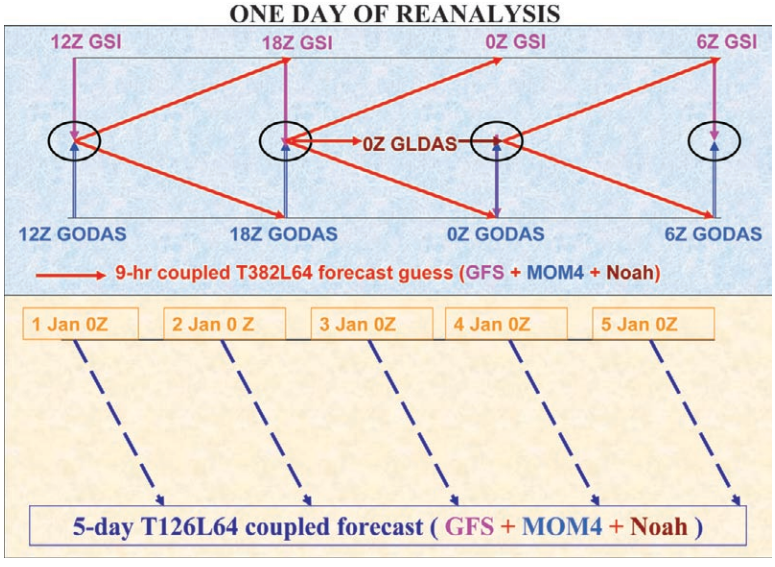

Fig. 1. Schematic of the execution of 1 day of the (top) CFS reanalysis and (bottom) layout of the 5-day forecasts to monitor the quality of the analysis.

- The lower half of Fig. 1 shows the layout of the coupled 5-day forecast, from every 0000 UTC initial condition, which is made with an identical but reduced horizontal resolution (T126L64) version of the atmosphere, for a sanity check.

Although the analysis is carried out every 6 h, 9-h forecast guess fields are required to accommodate both the data window and to handle information about the time derivative.

Before the actual production phase of the CFSR, a "light" version (CFSR-Lite) of the analysis was carried out to sweep through the entire data inventory. This was done with the uncoupled atmospheric model of the CFSR at T62L64 resolution. Each year was a single stream.

Peculiarities of some of the data ingest. A purist may expect that a coupled data assimilation system ingests observations and observations only. However, there are still some exceptions, primarily at the interface of the atmosphere with the land, and the ocean and the sea ice. Specifically, analyses of snow cover, SST, precipitation, and sea ice are used as input by the CFSR analysis subcomponents in gridded form. This has historical and/or practical reasons. Some of these variables, in the form of the original observations, are hard to assimilate by present state-of-the-art analysis systems. To the extent that these gridded fields are viewed by the scientific community as the gold standard of verification, we do not want to depart too much from them, if at all. In all cases, an analysis (which is often univariate) of these variables has already taken place (preceding any CFSR activity) to provide SST, snow cover, etc., in gridded form to the CFSR. This step may change the distribution of the original observations to something completely different. For this reason, the term "bogus" observations is sometimes used in some cases, such as gridded SLP data in the Southern Hemisphere during years when very few true observations were available. Some of these analyses are considered much better than the original scarce input data because of an ongoing manual analysis aspect. Ultimately, the use of bogus observations may no longer be needed, and we mention the following two area of significant progress:

1) In R1, satellite retrievals were assimilated, that is, the radiance observations were converted into something akin to a radiosonde profile. This is no longer necessary because the CFSR makes use of the raw observed radiance measurements. This is a major step forward.
2) Over land, it was hoped that the CFSR would obtain values of T2m that are close to those based on the univariate analyses of T2m that are used by the climate change community to estimate global warming trends. It is extremely difficult to assimilate T2m over land in systems like the CFSR. For this reason, ERA-40 post-processed observed T2m into their output. This was not done in the CFSR. However, changes were made to take into account rising levels of CO_2 concentrations in the data assimilation system. First indications show that the correct time mean T2m has emerged by assimilating everything, except T2m itself (see "Climate trend" section) in the CFSR.

Providing gridded data from an independent analysis at some interface does have implications for the degree of coupling that one can truly claim to be present in CFSR. For instance, the land hydrology model is provided with six variables from the atmospheric model (wind, humidity, pressure, etc.), but the model-generated precipitation (which is traditionally too biased) is replaced by observed precipitation for added realism. Therefore, the term "semicoupled" is used in "The Land" section.

R1 versus CFSR. Because R1 had many users, we provide its most significant differences from CFSR in Table 1.

THE OBSERVATIONS. Reanalysis projects depend upon both historical and operational archives of observations and newly reprocessed sets of observations being produced at meteorological research centers around the world. Historical observations are often found stored in obsolete formats and mediums at various research institutions and national archives in various stages of incompleteness and/or disarray. A great deal of historical data mining and archeology has been accomplished, particularly at NCAR, NCDC, and NESDIS, in preparation for these and other climate research projects (Jenne and Woollen 1994; Woollen and Zhu 1997). In addition, a number of important modern datasets have been reprocessed and improved from operational archives at international centers, such as ECMWF, ESA, EUMETSAT, JMA, and BOM.

Observation sources. The CFSR project will be completed over two epochs of meteorological observing periods. Initially, CFSR has analyzed the TOVS or "modern" era, from 1979 through the present. Subsequently, the CFSR will be extended back in time to 1947 or earlier. Most of the observations for the period from 1948 through 1978 are available as products of the previous reanalysis preparation efforts at NCEP. Observations assimilated from 1978 through 1997 were drawn either from sources that were copied from previous projects or from datasets improved or made available since previous NCEP reanalysis projects were completed. The observations for the most recent period of the CFSR, 1997–2009, have mostly been drawn directly from the NCEP operational run history archives, which have been saved in tape silos beginning when the NCEP modernized real-time BUFR database was implemented.

Observation preparation. The bulk of the work involved in data preparation for CFSR was invested in 1) data movement and archiving large-volume datasets, especially those from satellite-based observing systems, and 2) the merging of datasets containing overlapping contents. An annotated diagram of the distribution of the 8 TB of data found in the 30-yr online data dump archive is shown in Fig. 2.

Conventional observing systems in the CFSR. The subsections below briefly describe some specific details about the conventional observing systems used in the CFSR, their sources, characteristics, and preparation. Charts that illustrate the observing system performances of some of the input datasets, with respect to the quality control reactions and the monthly RMS and mean fits to the analysis and first guess backgrounds, over the 31-yr period, can be found in the supplement (online at http://dx.doi.org/10.1175/2010BAMS3001.2) and, for a more complete set, within the CFSR atlas (online at http://cfs.ncep.noaa.gov/cfsr).

Radiosondes and pibals. From 1948 through 1997 a number of archives were combined for the CFSR assimilation, including operational archives from the NCEP NMC, ECMWF, JMA, USAF, and U.S. Navy, along with other military, academic, and national archives collected at NCAR and NCDC. For the CFSR radiosonde preparation, duplicates

Table 1. Comparison between configurations used in R1 and CFSR.

R1	CFSR	For details
T62 horizontal resolution (~200 km)	T382 horizontal resolution (~38 km)	Subsection *The Atmospheric Forecast Model*
Sigma vertical coordinate with 28 levels with top pressure ~3 hPa	Sigma–pressure hybrid vertical coordinate with 64 levels with top pressure ~0.266 hPa	Subsection *The Atmospheric Forecast Model*
Simplified Arakawa–Schubert convection	Simplified Arakawa–Schubert convection with momentum mixing	Subsection *Modification to shallow convection and vertical diffusion*
Tiedtke (1983) shallow convection	Tiedtke (1983) shallow convection modified to have zero diffusion above the low-level inversions	Subsection *Modification to shallow convection and vertical diffusion*
Seasonal and zonal mean climatological ozone for radiation	Prognostic ozone with climatological production and destruction terms computed from 2D chemistry models	Subsection *Upgrades to Radiation Parameterization*
Diagnostic clouds parameterized based on relative humidity	Prognostic cloud condensate from which cloud cover is diagnosed	Subsection *The Atmospheric Forecast Model*
Orographic gravity wave drag based on GLAS GFDL approach	Orographic gravity wave drag based on Kim and Arakawa (1995) approach and sub-grid scale mountain blocking following Lott and Miller (1997)	Subsection *The Atmospheric Forecast Model*
GFDL IR radiation with random cloud overlap and fixed CO_2 of 330 ppmv	AER RRTM IR radiation with maximum/random cloud overlap and observed global mean CO_2	Subsection *Upgrades to Radiation Parameterization*
GFDL SW based on Lacis–Hansen (1974) scheme with random cloud overlap and fixed CO_2 of 330 ppmv; no aerosols or rare gases	AER RRTM SW radiation with maximum/random overlap and observed global mean CO_2, aerosols including volcanic origin plus rare gases	Subsection *Upgrades to Radiation Parameterization*
Local K vertical diffusion both in PBL and free atmosphere with a uniform background diffusion coefficient	Non-local vertical diffusion in the PBL with local K in the free atmosphere with exponentially decaying background diffusion coefficient	Subsection *Modification to shallow convection and vertical diffusion*
Second-order horizontal diffusion	Eighth-order horizontal diffusion	Subsection *Modification to shallow convection and vertical diffusion*
Virtual temperature as prognostic variable	Specific enthalpy as a prognostic variable; more accurate thermodynamic equation	Subsection *Enthalpy as prognostic variable*
OSU two-layer land surface model	Noah four-layer land surface model	Subsection *The Land Surface Analysis*
Prescribed SST and sea ice as lower boundary condition	Coupled to GFDL MOM version 4 and a two-layer sea ice model	Subsection *The Ocean Model (MOM4)* and *The Sea Ice Model*

were resolved by merging the contents of duplicate soundings instead of picking one sounding from one of the sources and discarding the others, as has been commonly done in reanalysis projects to date. Starting in May 1997, the NCEP operational run history archive supplied the bulk of the CFSR radiosonde data. An example of an observing system performance chart, for 500-mb radiosonde temperature, is shown in Fig. 3. Similar plots for the systems described below are found in the supplement (online at http://dx.doi.org/10.1175/2010BAMS3001.2).

AMMA special observations. A special observation program known as AMMA has been ongoing since 2001, which is focused on reactivating, renovating, and installing radiosonde sites in West Africa (Kadi 2009). The CFSR was able to include much of this special data in 2006, thanks to an arrangement with the ECMWF and the AMMA project.

AIRCRAFT AND ACARS DATA. The bulk of CFSR aircraft observations are taken from the U.S. operational NWS archives; they start in 1962 and are continuous through the present time. A number of archives from military and national sources have been obtained and provide data that are not represented in the NWS archive. Very useful datasets have been supplied by NCAR, ECMWF, and JMA. The ACARS aircraft observations enter the CFSR in 1992.

SURFACE OBSERVATIONS. The U.S. NWS operational archive of ON124 surface synoptic observations is used beginning in 1976 to supply land surface data for CFSR. Prior to 1976, a number of military and national archives were combined to provide the land surface pressure data for the CFSR. All of the observed marine data from 1948 through 1997 have been supplied by the COADS datasets. Starting in May 1997 all surface observations are taken from the NCEP operational archives. METAR automated reports also start in 1997. Very high-density MESONET data are included in the CFSR database starting in 2002, although these observations are not assimilated.

PAOBS. PAOBS are bogus observations of sea level pressure created at the Australian BOM from the 1972 through the present. They were initially created for NWP to mitigate the extreme lack of observations over the Southern Hemisphere oceans. Studies of the impact of PAOB data (Seaman and Hart 2003) show positive impacts on SH analyses, at least until 1998 when ATOVS became available.

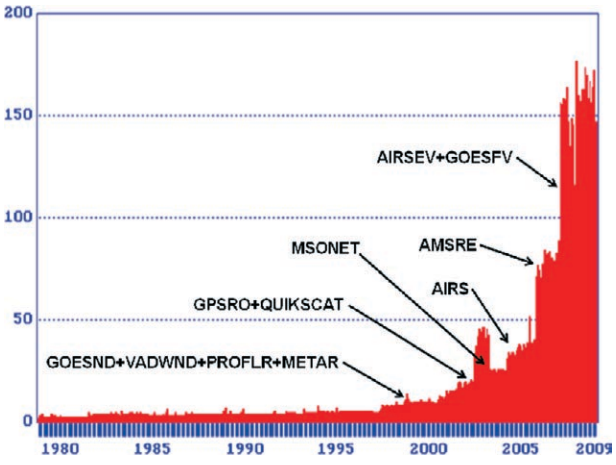

FIG. 2. Diagram illustrating CFSR data dump volumes, 1978–2009 (GB month^{-1}).

SATOB OBSERVATIONS. Atmospheric motion vectors derived from geostationary satellite imagery are assimilated in the CFSR beginning in 1979. The imagery from GOES, METEOSAT, and GMS satellites provide the observations used in CFSR, which are mostly obtained from U.S. NWS archives of GTS data. GTS archives from JMA were used to augment the NWS set through 1993 in R1. Reprocessed high-resolution GMS SATOB data were specially

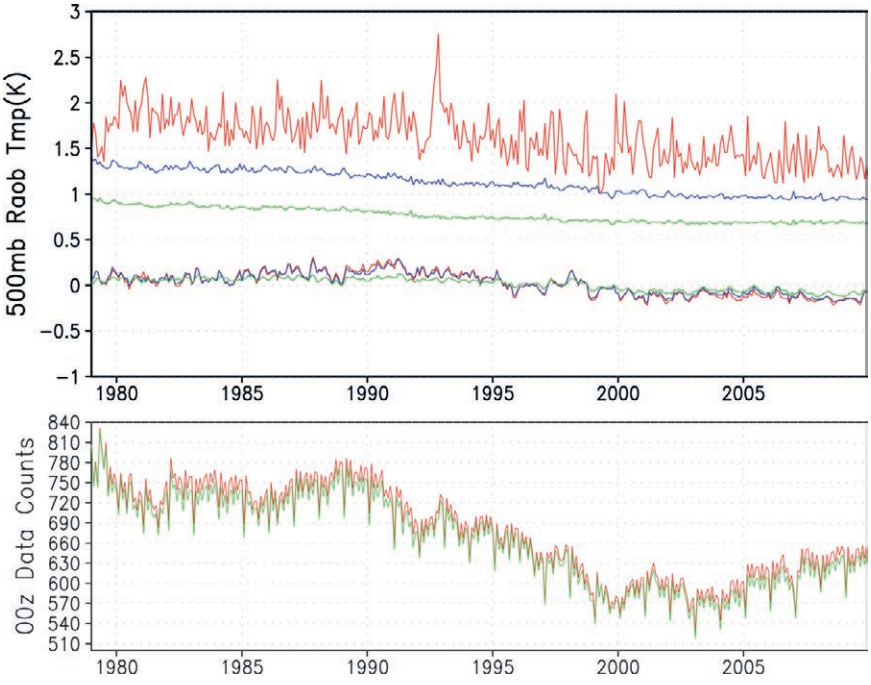

FIG. 3. Performance of 500-mb radiosonde temperature observations. (top) Monthly RMS and mean fits of quality-controlled observations to the first guess (blue) and the analysis (green). The fits of all observations, including those rejected by the QC, are shown in red. Units: K. (bottom) The 0000 UTC data counts of all observations (red) and those that passed QC and were assimilated (green).

produced by JMA for JRA-25, covering the years from 1987 through 2003. These were made available for the CFSR project and replaced the GTS GMS data in the assimilation database for that period. MODIS polar wind data are obtained from the NCEP archives for the CFSR starting in late 2004.

SSM/I OCEAN SURFACE WIND SPEED. Microwave imager data from the SSM/I are processed to derive ocean surface wind speed observations. SSM/I microwave radiance datasets, from DMSP satellites, were obtained from NCDC starting in 1993. Beginning in 1997 the DMSP data are available in the NCEP archives. The SSM/I brightness temperature data were converted to wind speeds by a neural net algorithm developed at NCEP and used in NCEP operations (Krasnopolsky et al. 1995; Gemmill and Krasnopolsky 1999; Yu et al. 1997).

SCATTEROMETER WINDS. Ocean surface wind datasets have been available from the European Space Agency *ERS-1*/AMI scatterometer since 1991 and from the *ERS-2*/AMI instrument since 1996. ESA has recently reprocessed ocean surface wind vectors from the *ERS-1* and *ERS-2* satellite archives, covering the years from 1991 through 2007. The reprocessed data were obtained for the entire period and assimilated in the CFSR. The NASA QuikSCAT SeaWinds scatterometer was launched in 1999. SeaWinds ocean surface vector wind data from the NCEP operational archives were assimilated in CFSR from 2001 until it went nonoperational in 2009. The NRL WindSat scatterometer data are assimilated in CFSR starting September 2008.

Satellite radiance–based observing systems in the CFSR. The NCEP operational GDAS has directly assimilated satellite radiances for a number of years, but CFSR is the first NCEP global reanalysis to do so. The historical TOVS and ATOVS archives were obtained from the NESDIS Web-based CLASS archive online. All of the other radiance data were obtained from the NCEP operational archives. Figure 4 is a CFSR usage chart of radiance-measuring instruments illustrating the time period during which each instrument was assimilated.

TOVS RADIANCES. The CFSR assimilated radiance data from satellite sounders with TOVS instruments onboard nine NOAA polar-orbiting satellites from *TIROS-N* to *NOAA-14* starting in 1978. The 1B datasets were calibrated using operational calibration coefficients stored in the files to convert the raw data counts into brightness temperatures

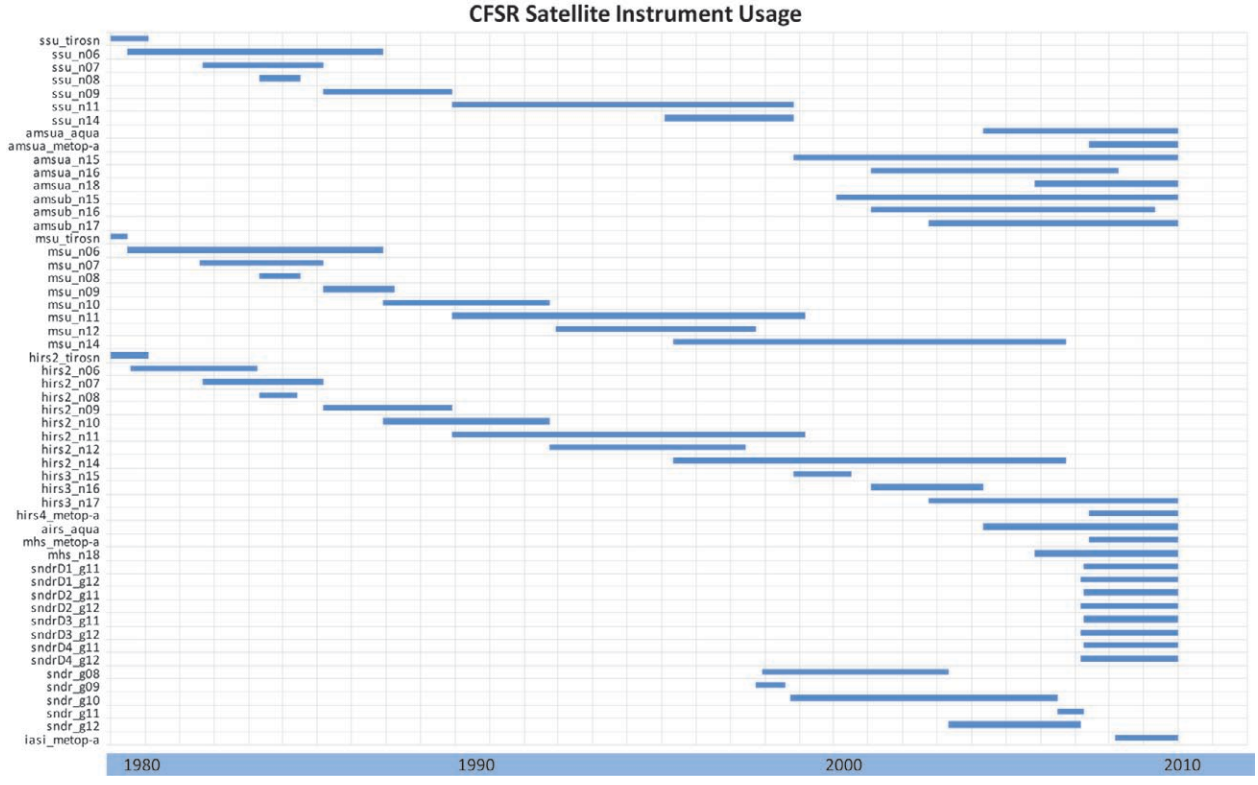

FIG. 4. Radiance instruments included in CFSR and the time period each was assimilated.

for the assimilation. The HIRS/2 and MSU from *NOAA-14* were assimilated until late 2006. *NOAA-11* and *NOAA-14* SSU were available until June 2004 and December 2005, respectively, but were only monitored in the assimilation following November 1998. The CFSR is the first NCEP reanalysis system to assimilate SSU radiances (see "Use of the SSU in the CFSR" section).

RECALIBRATED MSU RADIANCES. NESDIS has recently developed a postlaunch MSU calibration algorithm using simultaneous nadir overpasses (Zou et al. 2006, 2009). By the time the CFSR was begun, the recalibration of MSU channels 2, 3, and 4 from *NOAA-10* to *NOAA-14* was completed, and the recalibrated data were included in CFSR.

ATOVS RADIANCES. Advanced TOVS instrumentation became available in 1998 with the launch of the *NOAA-15* satellite and continued on with *NOAA-16* and *NOAA-17*. Evidence from a number of studies (e.g., English et al. 2000) suggests that ATOVS data have been an important factor in large forecast improvements seen over the last decade.

GOES RADIANCES. GOES sounder radiances have been available since the launch of *GOES-8* in 1994. CFSR started assimilating superobed sounder radiances from *GOES-8* and *GOES-9* in 1997 when the NCEP operational archive began. Full-resolution sounder data were introduced into CFSR assimilation in 2007.

AQUA AIRS, AMSU-A, AND AMSR-E DATA. The *Aqua* satellite was launched by NASA in May 2002. Several advanced infrared sounders from NASA were on board—AIRS and AMSU-A, along with a microwave scanning radiometer, AMSR-E, built by the Japanese NSDA. For operational use, the AIRS data were thinned by a factor of 40 in the horizontal arrangement and by a factor of 10 in the channel set. CFSR has followed this protocol and assimilated the AIRS data in this configuration beginning in late 2004. In mid-2007 the operational NCEP GDAS considered all nine AIRS spots in each AMSU field of view for assimilation and the CFSR followed suit.

METOP IASI, AMSU-A, AND MHS DATA. The *MetOp-A* satellite was launched in October 2006 and became operational in May 2007. The instruments carried onboard *MetOp-A* include IASI, the highest-resolution infrared sounding interferometer currently in orbit, the AMSU-A sounder, and the MHS instrument, which replaces AMSU-B. *MetOp-A* data were introduced into the CFSR assimilation in January 2009.

CHAMP/COSMIC GPS radio occultation data. The CFSR assimilates newly recalculated GPS RO observations from the CHAMP mission (Wickert et al. 2001) from May 2001 to December 2007. Recalculated observations from the U.S.–Taiwan COSMIC mission, launched in April 2006, are assimilated starting in July 2006. COSMIC is a six LEO microsatellite constellation that provides around 2,000 daily worldwide atmospheric soundings (Cucurull and Derber 2008; Cucurull 2010).

Observation quality control. The CFSR uses the NCEP operational observation QC procedures, which are summarized in Table S1 in the supplement (online at http://dx.doi.org/10.1175/2010BAMS3001.2; for more specific information on NCEP QC procedures see www.emc.ncep.noaa.gov/mmb/data_processing/prepbufr.doc/document.htm).

Performance charts including QC summaries for many of the observing systems used in the CFSR can be found on the CFSR Web site (online at http://cfs.ncep.noaa.gov/cfsr).

THE ATMOSPHERE. *The atmospheric data assimilation.* R1 and R2 were run with modified versions of the mid-1995 operational GDAS system, based on a sigma coordinate spectral prediction model and the 3DVAR SSI (Parrish and Derber 1992; Derber et al. 1991). These systems assimilated a fixed set of conventional observations along with TOVS/ATOVS temperature retrievals (Smith et al. 1979). Both projects continue as CDAS.

EVOLUTION OF THE NCEP GDAS. In anticipation of the eventual replacement of TOVS by the ATOVS instruments in 1998, retrieval assimilation was replaced with the direct assimilation of cloud-cleared radiances in the SSI (Derber and Wu 1998; McNally et al. 2000). Sources of radiance data other than ATOVS, such as the GOES sounder, AIRS, and AMSR-E, were added to the SSI assimilation as they became available. The operational GDAS underwent another major upgrade in May 2007 involving both the method of data assimilation and the vertical coordinate of the global prediction model. The sigma coordinate system (Phillips 1957) of the global spectral model, in use since the early 1980s, was replaced with a hybrid sigma–pressure system (Juang 2005). The SSI was replaced by the GSI (Kleist et al. 2009), which permits a more general treatment of the background errors,

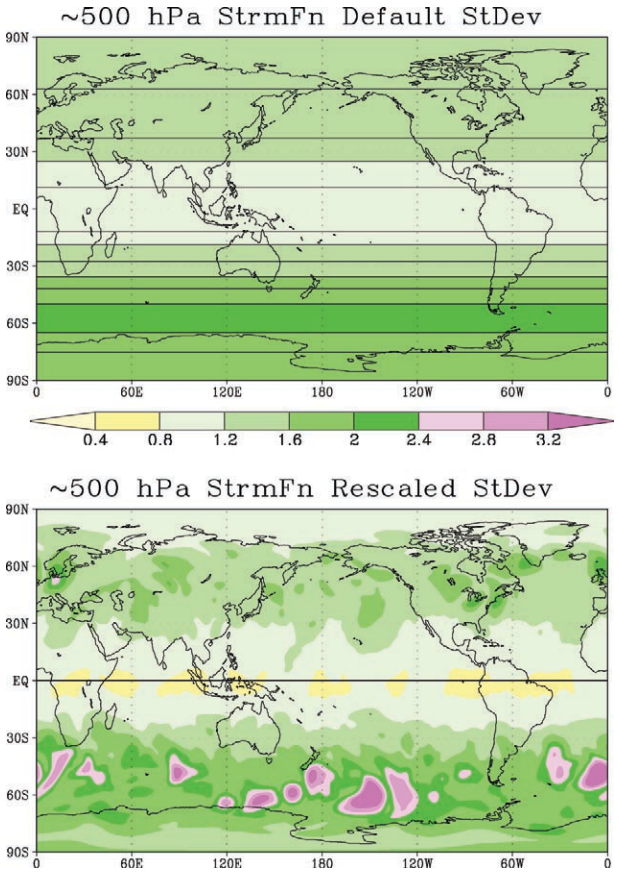

Fig. 5. (top) Static, zonally invariant, 500-hPa streamfunction (1e6/s) background error valid at 0000 GMT 06 Nov 2007; (bottom) flow-dependent adjusted background standard deviation.

such as flow dependence and anisotropic structure functions (Purser et al. 2003a,b). The 2007 operational GSI would form the basis for the CFSR.

Development of the GSI for the CFSR. Several innovative features were built into the operational version of the GSI implemented in May 2007 but were not activated because of resource limitations on implementation testing. Three features in particular were expected to be activated in the period that it was expected to take to complete the CFSR, so they were included from the start.

The first of these was to apply flow dependence to the background error variances (D. Kleist 2009, personal communication) in an effort to improve upon the climatological estimates that were previously in use (Wu et al. 2002; Kleist et al. 2009). The static variances undergo a simple rescaling based on the 6-h tendency in the model forecast, where the variances are increased (decreased) where the model tendencies are relatively large (small). The rescaling is performed level by level for each variable independently and is done in such a way as to approximately preserve the global mean variance as specified by the static estimate (i.e., it is not designed to increase or decrease the global mean error variance on a cycle-to-cycle basis). This procedure transforms the simple latitude- and height-dependent fixed variances into a fully three-dimensional, time-varying estimate, as shown in Fig. 5.

The second dormant GSI feature that is enabled in the CFSR is FOTO (Rančić et al. 2008). Many observation types are available throughout the 6-h assimilation window, but 3DVAR does not account for the time aspect. FOTO is a step in this direction. The algorithm takes advantage of the fact that tendency information for the guess fields (3-, 6-, and 9-h model forecasts) is readily available. The observational operators in the minimization procedure are generalized to use time tendencies of state variables.

The third GSI feature that is enabled in the CFSR is nonlinear VarQC (Andersson and Järvinen 1999), which replaces the OIQCBUFR program (Woollen 1991) that was used in R1 and R2 (Woollen et al. 1994). In the VarQC procedure, conventional GSI observation innovations (defined as the difference between the observations and the 6-h guess forecast) must first pass gross error checks. Then an innovation weight is computed based on its consistency with the solution of the variational minimization based on all of the available observations, including radiances, with additional input coming from the probabilities of error for the various observations. Any observation with a weight of 0.25 or greater is used in the minimization, in contrast to a typical pass/fail QC procedure where observations with a comparable weight of less than approximately 0.7 would be rejected from the process completely.

Another innovative feature of the CFSR GSI is the use of the historical concentrations of CO_2 (online at http://gaw.kishou.go.jp) when the historical TOVS instruments were retrofitted into the CRTM. Table 2 lists the values of CO_2 concentrations (ppmv) used to calibrate satellites back to 1979.

Use of the SSU in CFSR. The SSU instruments, onboard the majority of TOVS satellites, provide unique 29-yr observations for studying stratospheric temperatures. The SSU is a step-scanned infrared spectrometer with three modulated cell pressures for the original 15-μm CO_2 absorption band to be shifted up and split into three weighting functions, approximately located at 15, 5, and 1.5 hPa, for SSU channels 1, 2, and 3, respectively. However, historical use of the SSU radiances posed a challenge because of this

Table 2. Carbon dioxide concentrations (ppmv) used in the CRTM.

Satellite platform	Launch date*	Mission end date*	Mission mean (ppmv)**
TIROS-N	13 Oct 1978	1 Nov 1980	337.10
NOAA-6	27 Jun 1979	19 Sep 1983	340.02
NOAA-7	23 Jun 1981	7 Feb 1985	342.86
NOAA-8	28 Mar 1983	26 May 1984	343.67
NOAA-9	24 Dec 1984	13 Feb 1998	355.01
NOAA-10	17 Sep 1986	17 Sep 1991	351.99
NOAA-11	24 Sep 1988	Jun 2004	363.03
NOAA-12	14 May 1991	Standby	365.15
GOES-8	13 Apr 1994	5 May 2004	367.54
GOES-9	23 May 1995	28 Jul 1998	362.90
GOES-10	25 Apr 1997	West (in transition)	370.27
NOAA-14 to *NOAA-18*	May 1995	Dec 2007	380.00
IASI (*MetOp-A*)	19 Oct 2006	Current	389.00
NOAA-19	1 Jan 2010	Current	391.00

* Davis (2007)

** Historical CO_2 observed data available at WMO Global Atmospheric Watch (http://gaw.kishou.go.jp)

complicated sensor response and a leaking problem in the instrument's CO_2 cell pressure modulator, which caused the radiances from each satellite to exhibit a unique drift in time (Kobayashi et al. 2009). The CRTM, with its advanced surface emissivity model and radiative solver (Liu and Weng 2006), was used to quantitatively correct the leaking effect. By comparison to the detailed line-by-line calculation, the root-mean-square error resulting from the fitting and interpolation of the CO_2 cell pressure in the fast transmittance model is less than 0.1 K (Liu and Weng 2009). The SSU radiative transfer calculations were then compared to the SSU radiances from *NOAA-14*. The input temperature profiles are taken from the EOS *Aura* MLS product for November 2004, a completely independent data source. The MLS temperature product precision throughout the stratosphere is generally less than 1 K. More than 7,000 match-up data points are found, and all of the data points are analyzed. Figure 6 illustrates the results. The SSU and the MLS measurements are very consistent. The bias and rms error in the brightness temperature calculations at SSU channels 1 (black), 2 (red), and 3 (green) are less than (or equal to) 1.5 K.

Satellite bias correction spinup for CFSR. The direct assimilation of radiances represents one of the major improvements of the CFSR over R2. However, substantial biases exist when observed radiances are compared to those simulated by the CRTM depiction of the guess. These biases are complicated and relate to instrument calibration, data processing, and deficiencies in the radiative transfer model. A variational satellite bias correction scheme was introduced by Derber and Wu (1998) to address this issue when direct assimilation of radiances began at NCEP. This scheme has been continually developed and is used in the GSI system adapted for the CFSR. Before the radiances of a new instrument can be assimilated, its unique set of starting bias corrections must be determined by a separate spinup assimilation. In the case of CFSR, each set of historical instruments in Fig. 4 required an individual spinup. Because the TOVS instruments had never been assimilated by

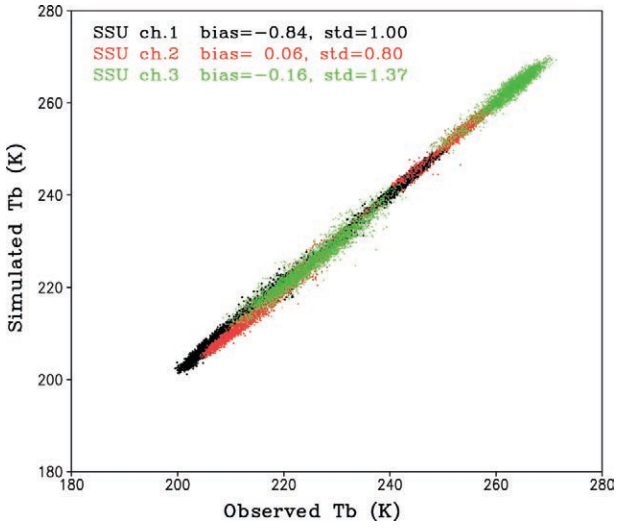

Fig. 6. Comparisons of the SSU brightness temperature at channel 1 (black), channel 2 (red), and channel 3 (green) between calculations and measurements for Nov 2004.

a GSI-based GDAS, a preliminary set of tests were run (not shown), which determined that a 3-month spinup was required prior to the introduction of those historical instruments in the CFSR. A detailed explanation of this important step is given in section 2 in the supplement (online at http://dx.doi.org/10.1175/2010BAMS3001.2).

Examples of the bias correction values actually applied to the CFSR over the TOVS period of the CFSR, 1979–98, may be seen in globally averaged, 4-times-daily averaged time series for MSU channels 1–4 and SSU channels 1–3 in Fig. 7. (The spinup of the SSU channels was done at the same time.) The one measure of the successful spinup procedures is the lack of discontinuities in the transitions between successive instruments. The breaks in the MSU time series are a result of the recalibration that was applied beginning in 1986, as noted in the "Recalibrated MSU Radiances" section.

Transition to real-time CFSR. Once CFSR stream 4 (see the "Overview" above) was completed in February 2009, a decision point was reached. The operational GSI had gone through several upgrades during the CFSR execution, the latest in February 2009 being a major addition to the CRTM to simulate the hyperspectral channels of the IASI instrument, onboard the new ESA MetOp satellite. The IASI radiances had become operational in March 2009.

To continue to meet the goal of providing the best-available initial conditions to the CFS, in the absence of staff and resources to maintain the CFSR GSI into the future, it was decided to make the transition to the CDAS mode of CFSR. The operational GSI, with both present and future implementations, would replace the CFSR GSI, and the coupled prediction model would be "frozen" to that of the CFS version 2. Historical observational datasets would be replaced with the operational data dumps. One consequence of the switch to the operational GSI would be that the period of March 2009 forward would be run without FOTO because the anticipated operational implementation of FOTO did not happen.

QBO problem in the GSI. The QBO can only be fully depicted in assimilation systems by sufficient direct wind observations, because the underlying physical mechanism is based on the dissipation of upwardly propagating gravity waves (Lindzen and Holton 1968), which are filtered out by the hydrostatic assumption. Soon after CFSR production began, it was noted that streams 2 and 3 completely missed the QBO wind transition. This was unexpected based on the ability of R1, R2, and CFSR streams 1 and 4 (starting in 1979 and 2004 respectively) to capture the QBO wind patterns. While searching for a comprehensive solution, it was noted that the ERA-40 tropical stratospheric wind profiles were readily available for the streams in question, included the stratospheric layers needed, and, qualitatively, adequately depicted both the QBO and semiannual oscillation. In order that the streams could proceed with a reasonable QBO signature, it was decided that the ERA-40 stratospheric wind profiles should be used as bogus observations for the period from 1 July 1981 to 31 December 1998.

Streams 1 and 4 had a common denominator. Both had more vertically resolved and temporally complete tropical radiosonde wind observations at the beginning of the stream than those in stream 2. Stream 1 benefited from the enhanced FGGE observation system, and stream 4 benefitted from the automation of modern radiosonde data collection, which results in more reports reaching the GTS and more stratospheric levels in the individual reports. The solution to this problem became apparent from consultations with several GMAO MERRA team members, after determining that the MERRA reanalysis, which uses the same GSI assimilation component, depicted the QBO very well. Prior to starting the MERRA project, GMAO had experienced a similar problem when analyzing the QBO in an earlier gridpoint analysis system. The problem was resolved by enlarging the horizontal length scale of the zonal wind correlation function in the tropical stratosphere (Gaspari et al. 2006). When the GMAO assimilation system was

switched to the GSI, the tropical stratospheric stream function variances of the background error reflected the changes made to fix the problem in the earlier system. When comparable background error variances were tested in the GSI for a case where the CFSR had failed to capture the QBO, the wind transition was successfully analyzed (not shown).

Tropical cyclone processing. As global prediction models have become more highly resolved and have included more realistic boundary layer, water cycle, and radiative parameterizations, the genesis and evolution of tropical storms has dramatically improved. Equally important is the ability of improved observing systems and analysis procedures to depict the details of tropical storm circulations and their accompanying steering flow.

The first global reanalysis to assimilate historical tropical storm information was JRA-25 (Onogi et al. 2007). It assimilated synthetic wind profiles (Fiorino 2002) surrounding the historical storm locations of Neumann (1999).

A unique feature of the CFSR is its approach to the analysis of historical tropical storm locations. The CFSR applied the NCEP tropical storm relocation package (Liu et al. 1999), which is a key component of the operational GFS analysis and prediction of tropical storms. By relocating a tropical storm vortex to its observed location prior to the assimilation of storm circulation observations, distortion of the circulation by the mismatch of guess and observed locations is avoided. M. Fiorino 2007, personal communication) provided the CFSR with the historical set of storm reports (provided to NCEP by the National Hurricane Center and the U.S. Navy Joint Typhoon Warning Center) converted into the operational format. Figure 8 shows the yearly total of tropical storm reports stacked by the eight geographical basins.

A measure of the ability of the assimilation system to depict observed tropical storms is to quantify whether or not a reported storm is detected in the guess forecast. Figure 9 shows the time series of the

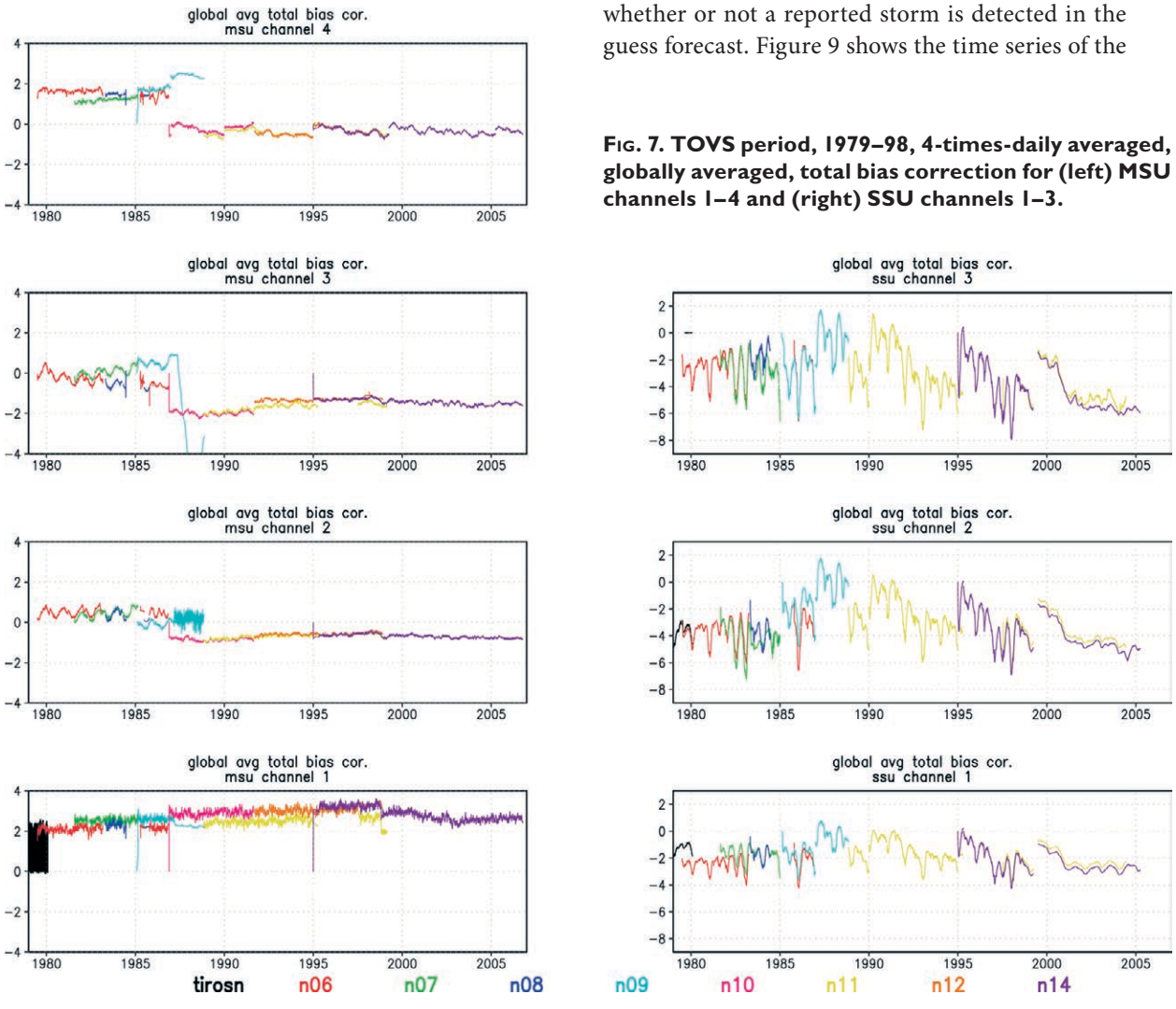

Fig. 7. TOVS period, 1979–98, 4-times-daily averaged, globally averaged, total bias correction for (left) MSU channels 1–4 and (right) SSU channels 1–3.

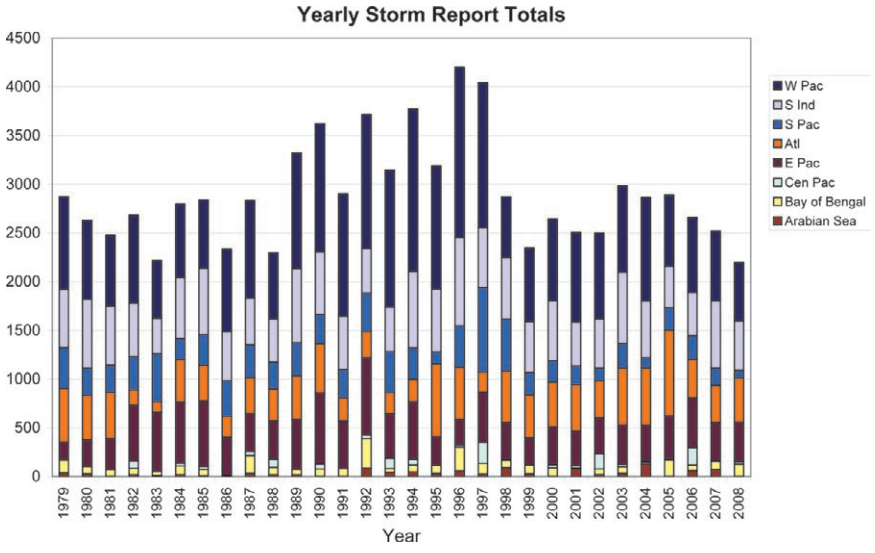

Fig. 8. The yearly total of tropical storm reports stacked by the eight geographical basins (from top to bottom): western Pacific (W Pac), southern Indian (S Ind), southern Pacific (S Pac), North Atlantic–Caribbean (Atl), eastern Pacific (E Pac), central Pacific (Cen Pac), Bay of Bengal, and Arabian Sea.

percentage of detected tropical storms plotted globally and for selected Northern Hemisphere basins. A noticeable improvement starts in 2000, coincident with the full utilization of the ATOVS satellite instruments, such that between 90% and 95% of reported tropical storms are detected.

The atmospheric forecast model. R1 (Kalnay et al. 1996; Kistler et al. 2001) was based on the operational medium-range forecast model of January 1995. R1 had a T62 horizontal resolution with 28 sigma layers in the vertical with the top layer near 3 hPa. Since then the NCEP GFS AM has undergone many major revisions and improvements. The current operational CFS version 1 (Saha et al. 2006) is based upon the GFS AM of 2003. Changes from R1 that are included in CFS version 1 are made to the boundary layer vertical diffusion (Hong and Pan 1996) and simplified Arakawa–Schubert cumulus convection (Pan and Wu 1995; Hong and Pan 1998), with cumulus momentum mixing and orographic gravity wave drag (Kim and Arakawa 1995; Alpert et al. 1988, 1996). The land surface model used in CFS version 1 is based on the two-layer OSU land model (Pan and Mahrt 1987). In addition, the cloud condensate is a prognostic variable (Moorthi et al. 2001) with a simple cloud microphysics parameterization (Zhao and Carr 1997; Sundqvist et al. 1989). Both large-scale condensation and the detrainment of cloud water from cumulus convection provide sources for cloud condensate. The fractional cloud cover used in the radiation calculation is diagnostically determined by the predicted cloud condensate based on the approach of Xu and Randall (1996). The CFS version 1 also has ozone as a prognostic variable with a simple parameterization for ozone production and destruction based on 10-day mean climatological data supplied by NASA GSFC. The SW radiation is parameterized following the NASA approach (Chou et al. 1998; Hou et al. 1996, 2002) and the LW radiation following the GFDL approach (Fels and Schwarzkopf 1975; Schwarzkopf and Fels 1991). Both radiation parameterizations use random cloud overlap with shortwave radiation being called every hour and longwave radiation every 3 h. CFS version 1 (and the underlying GFS AM) produced more realistic moisture prediction, which may have contributed to a better depiction

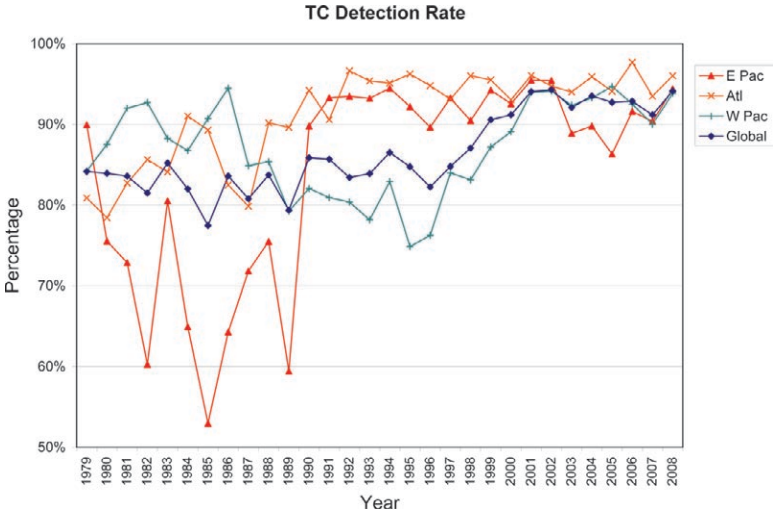

Fig. 9. Time series of the percentage of detected tropical storms plotted globally and for selected Northern Hemisphere basins: the Atlantic–Caribbean (ATL), western Pacific (W Pac), and eastern Pacific (E Pac).

of no-rain areas and much-improved tropical and hurricane track prediction.

Additional improvements to the GFS AM include fast and accurate LW radiation parameterization based on the RRTM developed at AER (Mlawer et al. 1997). It is also coupled to a four-layer Noah land surface model (Ek et al. 2003) and a two-layer sea ice model (Wu et al. 2005). In addition to gravity wave drag, the model now includes a parameterization of mountain blocking (Alpert 2004) following the subgrid-scale orographic drag parameterization by Lott and Miller (1997). The GFS AM now takes advantage of the ESMF-based modern computer algorithms (Collins et al. 2005). An update of the ozone production and destruction terms is done by using the monthly mean data provided by the NRL (McCormack et al. 2006). The horizontal resolution is T382 with 64 hybrid vertical layers, with the top layer ~0.2 hPa. The vertical structure of the model levels for the 28 sigma-layer model (left panels) used in R1 and the 64-layer sigma–pressure hybrid model (right panels) used in the GFS/CFSR is shown in Fig. 10 as a cross section at 90°E and as a cross section across North America at 40°N (Fig. S25 in the supplement, online at http://dx.doi.org/10.1175/2010BAMS3001.2). The top panels in both figures are plotted as a function of pressure to emphasize the resolution in the troposphere. In the bottom panels, the vertical axis is log(pressure) to emphasize the stratosphere. It can be seen that the new vertical structure adds more resolution in the stratosphere and does not have the topography signature above 100 hPa.

The version of AM used in the CFSR has some additional enhancements. It now uses C_pT as a prognostic variable in place of virtual temperature (Akmaev and Juang 2008; Juang 2009), with a generalized hybrid vertical coordinate with a sigma–pressure option. The vertical discretization is based on the generalized hybrid vertical coordinate (Juang 2005). It also uses RRTM shortwave radiation with maximum random cloud overlap (Iacono et al. 2000; Clough et al. 2005) and a simple modification of shallow convection/vertical eddy diffusion parameterizations that helps in improving marine stratus prediction off the west coasts of continents. In the following, some details on some of the newer features of the model are provided.

Upgrades to radiation parameterization. The new RRTMG-LW and RRTMG-SW parameterizations are adapted from AER (e.g. Mlawer et al. 1997; Iacono et al. 2000; Clough et al. 2005). In the CFSR implementation we invoke both SW and LW radiations at 1-h intervals. A maximum-random cloud-overlapping scheme is used for cloudy sky radiative transfer, and a climatological aerosol scheme provides the global distribution of aerosol optical depth. Stratospheric volcanic aerosols are also included. For CO_2 amount, historical monthly mean observed data available from the WMO Global Atmosphere Watch web site (online at http://gaw.kishou.go.jp) are used. These data are reprocessed into a 15° latitude × 15° longitude horizontal grid that is used in the CFSR. In addition to the major radiatively absorbing gases (water vapor, ozone, and CO_2), many other radiatively active greenhouse gases are also included. They are prescribed as global annual mean values adapted from the NIST.

The RRTMG-LW employs a computationally efficient correlated-k method for radiative transfer calculations. It contains 16 spectral bands with various numbers of quadrature points (*g* points) in each of the bands that sum up to a total of 140 *g* points. Active gas absorbers include H_2O, O_3, CO_2, CH_4, N_2O, O_2, and four types of halocarbons (CFCs). The RRTMG-SW contains 14 spectral bands with various numbers of *g* points in each of the bands to a total of 112. It uses a fast two-stream radiative transfer scheme, and includes many absorbing gases (H_2O, O_3, CO_2, CH_4, N_2O, O_2).

Enthalpy as prognostic variable. The generalized vertical hybrid coordinate version of GFS used in CFSR has been formulated to take into account the spatial and temporal variations of the gas constant and specific heat at the constant pressure of dry air, as well as all of the model gas form tracers in all of the prognostic equations. Instead of solving spatial and temporal variations of specific heat at constant pressure and temperature separately in the thermodynamics equation, specific enthalpy h (where $h = C_pT$, where C_p is the averaged specific heat at constant pressure of dry gas and all gas tracers, and T is the temperature) is introduced as the thermodynamic prognostic variable to solve C_p and T together (Juang 2010, manuscript submitted to *Mon. Wea. Rev.*). With this approach, the virtual temperature used in the CFSR AM is replaced by specific enthalpy for all prognostic equations. The thermodynamic equation is now more accurate in considering dry air and all of the gas tracers. It is also valid for the entire atmosphere to even higher altitudes than 600 km, where the dry gas density may be less but other gas tracers are not. However, current physical parameterizations in the model have not been updated to take advantage of the variations in C_p and the gas constant R.

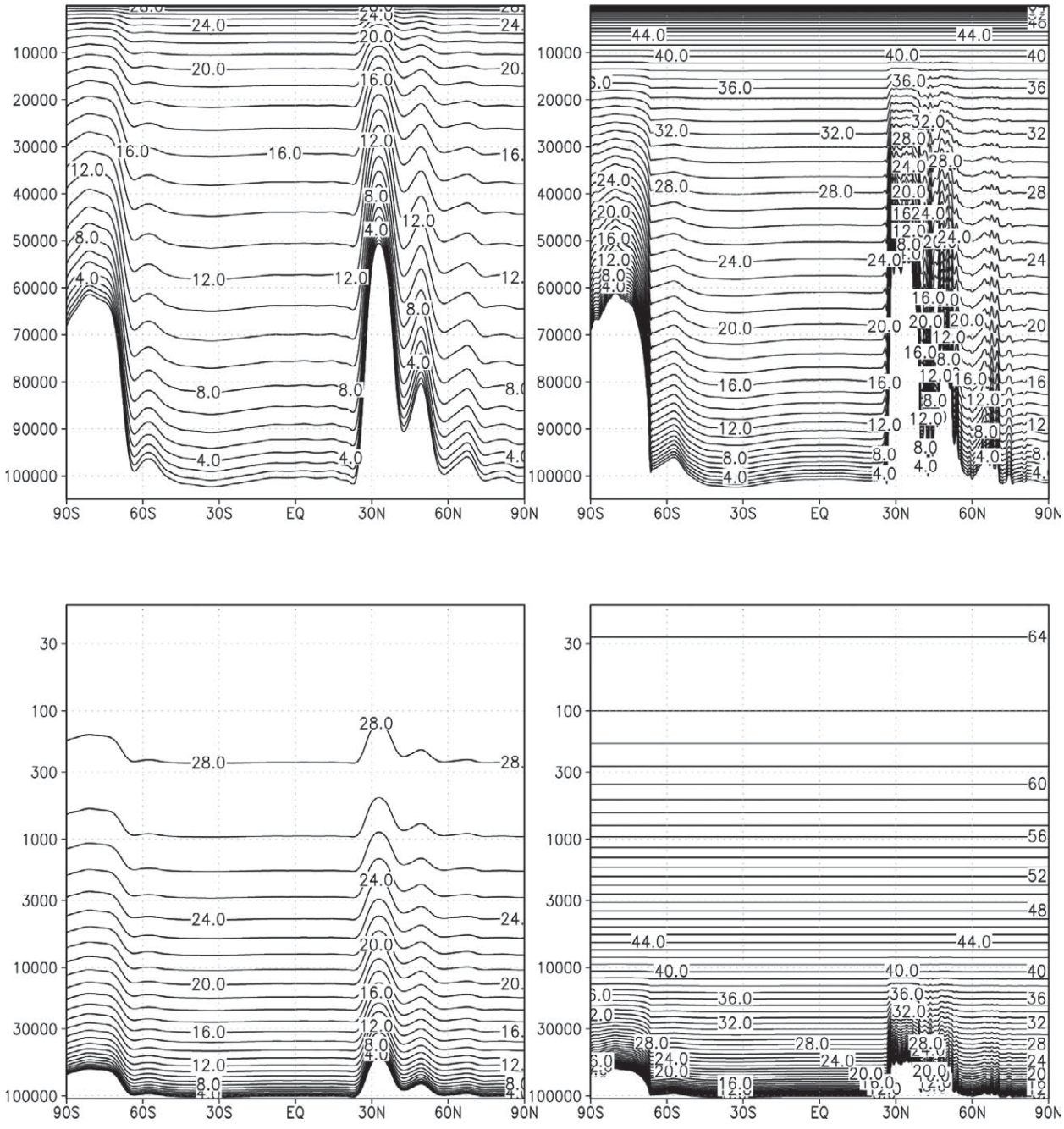

Fig. 10. The vertical structure of model levels as a meridional cross section at 90°E. (left) RI (28 sigma layers) and (right) CFSR (64 sigma–pressure hybrid layers). (top) A linear function of pressure to emphasize resolution in the troposphere; (bottom) log(pressure) is plotted to emphasize stratosphere.

Modification to shallow convection and vertical diffusion. CFS version 1 has very poor prediction of marine stratus near the west coast of the major continents. This poor performance may be attributed to the parameterization of shallow convection and background vertical eddy diffusion (Moorthi et al. 2010). The shallow convection parameterization follows Tiedtke et al. (1983). It is applied wherever the deep convection parameterization is not active. In this scheme, the highest positively buoyant level below the 0.7 P_s (where P_s is surface pressure) level for a test parcel from the top of the surface layer is defined as the shallow convection cloud top. The cloud base is the LCL for the same test parcel. Enhanced vertical eddy diffusion is applied to temperature and specific humidity within this cloud layer. The diffusion

coefficients are prescribed with a maximum value of 5 $m^2 s^{-1}$ near the center, approaching zero near the edges. The GFS also includes a background eddy vertical diffusion to enhance mixing close to the surface, where eddy diffusion calculated by the PBL parameterization is considered inadequate, particularly in the stable regime. The diffusion coefficient decreases exponentially with pressure, with the surface value set to 1.0 $m^2 s$. The combination of shallow convection and background diffusion reduces the strength of low-level inversions off the west coast of the continents, thus reducing the prediction of marine stratus.

In the version of the AM used in CFSR, two simple modifications (Moorthi et al. 2010) are used to improve the prediction of marine stratus. The modifications are as follows: 1) limit the shallow convection top to be below the low-level inversion when the condition for cloud-top entrainment instability (Deardorf 1980; Randall 1980) is not satisfied; and 2) set the background vertical diffusion to zero above low-level inversions. The combination of these two modifications, although not perfect, does improve the prediction of marine stratus.

THE OCEAN. *The sea surface temperature analysis.* Two daily SST analysis products have been developed using OI. Both products have a spatial grid resolution of 1/4°. One product uses AVHRR infrared satellite SST data. The other uses AVHRR and AMSR on the NASA Earth Observing System satellite SST data. Both products also use in situ data from ships and buoys and include a large-scale adjustment of satellite biases with respect to the in situ data. The in situ and corrected satellite data are analyzed using an OI procedure. The correlation scales range from 50 to 200 km, with smaller scales in higher latitudes (especially in western boundary current regions) and larger scales in the tropics. Because of AMSR's near all-weather coverage, there is an increase in OI signal variance when AMSR is added to AVHRR.

The AVHRR-only product uses Pathfinder AVHRR data (currently available from September 1981 through December 2005) and operational AVHRR data for 2006 onward. The AMSR + AVHRR product begins with the start of the AMSR data in June 2002. In this product, the primary AVHRR contribution is in regions near land where AMSR is not available. However, in cloud-free regions, use of both infrared and microwave instruments can reduce systematic biases because their error characteristics are independent.

In the CFS reanalysis the more accurate combined SST product was used when available. Thus, the AVHRR-only product was used from November 1981 through May 2002 and was replaced by the AMSR + AVHRR product from June 2002 onward. Both products are available as versions 1 and 2. Version 1 of the daily OI is described in Reynolds et al. (2007). This version uses 1 day of satellite and in situ data with the satellite bias correction based on 7 days of satellite and in situ data. The changes from version 1 to version 2 (see www.ncdc.noaa.gov/oa/climate/research/sst/oidaily.php) are relatively small and primarily consist of additional temporal smoothing. The temporal smoothing includes using three consecutive days of satellite and in situ data where the middle day is weighted more strongly than the other 2 days and provides the date of the analysis. The temporal smoothing also includes additional smoothing of the satellite bias corrections using 15 days of satellite and in situ data, instead of 7. In addition, ship SSTs are corrected relative to the buoy SSTs by subtracting 0.14°C from all ship observations before they are used to bias correct the satellite data. Thus, all of the observations are bias corrected with respect to buoys.

These fields, as prepared by Reynolds et al. (2007) at NCDC, do not have temperature values over land. Missing grid points were filled in via interpolation at NCEP to ease conversion of the field to the CFS model grid. Some segments of the CFS reanalysis were run before the full time series of the version 2 OI daily SST was finalized. For some of these periods, an early release of that product was used.

Ideally, the daily OI products would have been used for the entire CFS reanalysis period. However, these products were not available for the earliest part of the reanalysis—January 1979 through October 1981. For this period, the SST fields prepared for the ERA-40 project were used. These fields were derived from the Met Office Hadley Center's monthly mean HadISST dataset, as described in Fiorino (2004).

The ocean model (MOM4). The oceanic component is the MOM version 4p0d (Griffies et al. 2004), which is a finite difference version of the ocean primitive equations configured under the Boussinesq and hydrostatic approximations. The model uses the tripolar grid developed by Murray (1996). Northward of 65°N it uses a rotated bipolar grid that places two poles over land, thus eliminating the singularity in the northern ocean, while southward of 65°N it uses a regular latitude × longitude grid. The horizontal layout is a staggered Arakawa B grid and geometric height is in the vertical. The ocean surface boundary is computed as an explicit free surface. The zonal resolution is 1/2°.

The meridional resolution is 1/4° between 10°S and 10°N, gradually increasing to 1/2° poleward of 30°S and 30°N. There are 40 layers in the vertical, with 27 layers in the upper 400 m, and the bottom depth is approximately 4.5 km. The vertical resolution is 10 m from the surface to the 240-m depth, gradually increasing to about 511 m in the bottom layer. Vertical mixing follows the nonlocal K-profile parameterization of Large et al. (1994). The horizontal mixing of tracers uses the isoneutral method developed by Gent and McWilliams (1990; see also Griffies et al. 1998). The horizontal mixing of momentum uses the nonlinear scheme of Smagorinsky (see Griffies and Halberg 2000). We have used a configuration for the MOM4p0d similar to the one we used for MOM version 3 in the current operational CFS (Saha et al. 2006), but we have approximately doubled the resolution and the MOM4p0d is fully global with an Arctic Ocean and interactive ice model, where the operational MOM3 is truncated at 64°N and 74°S.

The ocean analysis: GODAS. The GODAS at NCEP uses a 3DVAR scheme that has evolved from a version originally developed by Derber and Rosati (1989). It was first adopted to assimilate temperature data into a Pacific Ocean model and was subsequently modified to assimilate other datasets (Behringer et al. 1998; Ji et al. 2000; Behringer and Xue 2004; Behringer 2007). In the context of the CFSR, the GODAS assimilates temperature and salinity observations by minimizing the cost function as follows:

$$J = \tfrac{1}{2}\,(x - x_b)^T B^{-1} (x - x_b) + \tfrac{1}{2}\,[y - H(x)]^T R^{-1}\,[y - H(x)],$$

where the first term is the sum of the squared differences between the forecast x_b and the analysis x, weighted by the uncertainty of the forecast, and the second term is the sum of the squared differences between the observations y and the analysis $H(x)$, weighted by the uncertainty of the observations. The operator H is a linear interpolator between the model grid and the location of the observation. The matrix B is the background error covariance and the matrix R is the observational error covariance that includes the representation error, resulting from the part of the observed field that the model is unable to resolve. The background error covariance matrix B is univariate and thus blocks the diagonal with respect to temperature and salinity. The horizontal covariance is modeled using a diffusion equation approximating a Gaussian function that is stretched in the zonal direction, with the stretching being the greatest near the equator (Derber and Rosati 1989; Weaver and Courtier 2001). The vertical covariance is likewise modeled with a diffusion equation with length scales specified as a function of depth such that at any level the scale is twice the level thickness. The variance is set to be proportional to the square root of the local vertical temperature gradient computed from the forecast. The observational errors are assumed to be uncorrelated, so that R is a diagonal matrix of the estimated error variances of the observations. The errors assigned to a temperature profile vary with depth according to the square root of the vertical temperature gradient and are scaled to have values between 1° and 2.5°C. The standard error assigned to a salinity profile is a constant 0.1 psu at all depths.

Temperature and salinity profiles are assimilated at 6-h intervals using all of the observations from the previous 10-day interval. The more distant a profile is in time, the less weight it receives in the assimilation. This approach allows relatively sparse ocean observations to have a greater impact on the model state (Derber and Rosati 1989; Behringer et al. 1998).

An exception to the assimilation procedure is made in the top level of the model (5 m), where the ocean analysis is accomplished by simple relaxation to gridded fields of temperature and salinity. The purpose is to provide a stronger constraint on the ocean at the interface with the atmosphere. This is done by relaxing the sea surface temperature, every 6 h, to the daily mean ¼° SST, described in the "Sea Surface Temperature Analysis" section.

The ocean observations. TEMPERATURE PROFILES. The temperature observations used for assimilation are profiles from XBTs from fixed mooring arrays: TAO/TRITON (McPhaden et al. 1998) in the Pacific Ocean, PIRATA (Bourlès et al. 2008), and the RAMA in the tropical Indian Ocean and from the Argo profiling floats (Argo Science Team 2001). The XBT observations collected prior to 1990 have been acquired from the NODC's World Ocean Database 1998 (Conkright et al. 1999), whereas XBTs collected subsequent to 1990 have been acquired from the GTSPP (online at www.nodc.noaa.gov/GTSPP).

The distribution of these profiles in time and space has changed greatly over the last 30 yr. The most significant changes have been the completion of the TAO mooring array in the equatorial Pacific Ocean in the early 1990s and the rapid growth of the global Argo array after about 2002. Figure 11 illustrates the increase in the number of observations over the years and the increase in their depth of penetration. Figures 12 and 13 illustrate the changes between 1985

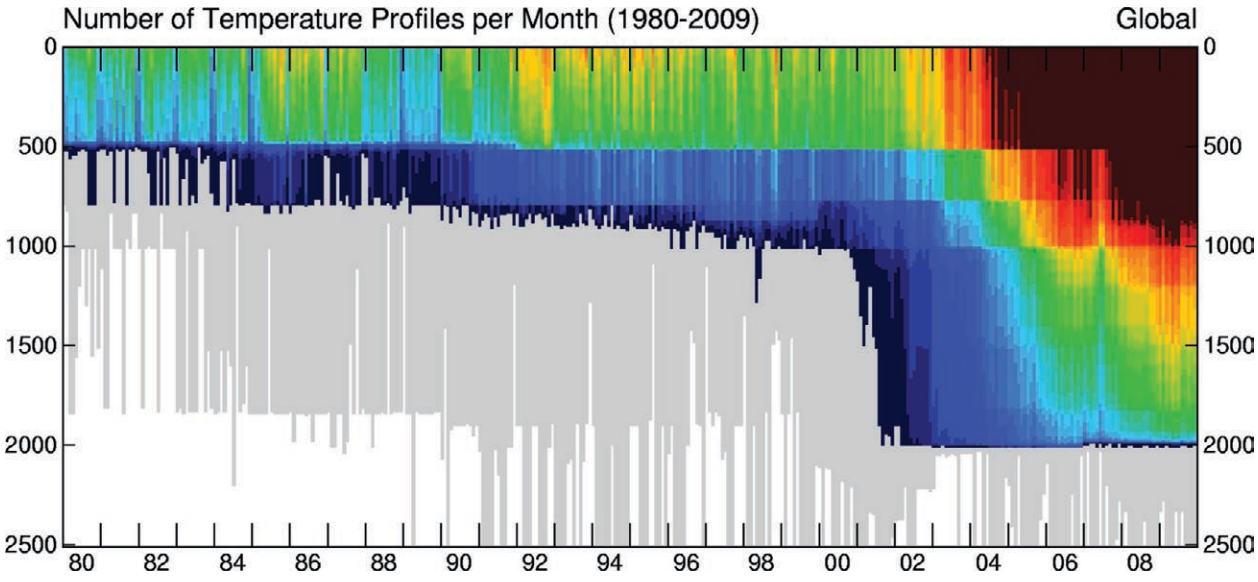

Fig. 11. The global number of temperature observations assimilated per month by the ocean component of the CFSR as a function of depth for the years 1980 through 2009. The stratification of colors reflects the penetration depths of different instrument types. The rapid growth of the Argo array after the year 2000 dominates the latter part of the figure. The contour interval is 250 observations.

when XBTs collected along shipping routes dominated the distribution and 2008 when the Argo array has made a nearly uniform global distribution possible.

Salinity profiles. Prior to the advent of the Argo array subsurface salinity observations were rare. In order to maintain water mass properties, as represented by the correlation between the potential temperature and salinity, we have chosen to assimilate synthetic salinity profiles as surrogates for direct observations. Each synthetic salinity profile is constructed from an observed temperature profile and the local climatological *T*–*S* correlation based on the World Ocean Database 1998 (Conkright et al. 1999). A similar result could be obtained by incorporating the *T*–*S* correlation in a bivariate assimilation. However, while the *T*–*S* correlation is strong in the deep ocean, it is weak in the near-surface ocean over much of the globe. A bivariate assimilation would allow the near-surface assimilation of temperature, for which we have direct observations, to be affected by an uncertain *T*–*S* correlation. The approach used here, which is univariate in temperature and salinity, avoids this problem.

Surface observations. As indicated above, temperature and salinity in the top level of the model are relaxed to gridded fields of observed SST and SSS. The SST data are the daily fields described in "The sea surface temperature analysis" section. The SSS data are a climatological map based on the World Ocean Database 1998 (Conkright et al. 1999). The SST and SSS fields are remapped to the model grid before being used in the CFSR.

THE SEA ICE. The accuracy of sea ice coverage is essential for a good reanalysis of the atmosphere and ocean over the polar regions. Global climate modeling studies have shown that sea ice concentrations strongly affect the climate in the Antarctic regions (e.g., Simmonds and Wu 1993). However, sea ice concentrations were prescribed in the previous NCEP reanalyses (both R1 and R2). The CFSR allows the sea ice concentration to be predicted in the forecast guess that provides closer sea ice–atmosphere linkages in the data assimilation system. This should improve the description of the main climatic patterns and trends over the polar regions in the CFSR.

The sea ice concentration analysis. The target in CFSR was to produce a global record of surface ice concentration for all points that may freeze anywhere on the globe, on a daily grid of 0.5° latitude × 0.5° longitude resolution throughout the period of the CFSR. Further, the aim was to produce the best analysis for the given day. This latter goal means that there are discontinuities in the production of the dataset, where newer datasets start to be used or newer methods of analysis are used or where, as for the Great Lakes, a high-quality dataset ended.

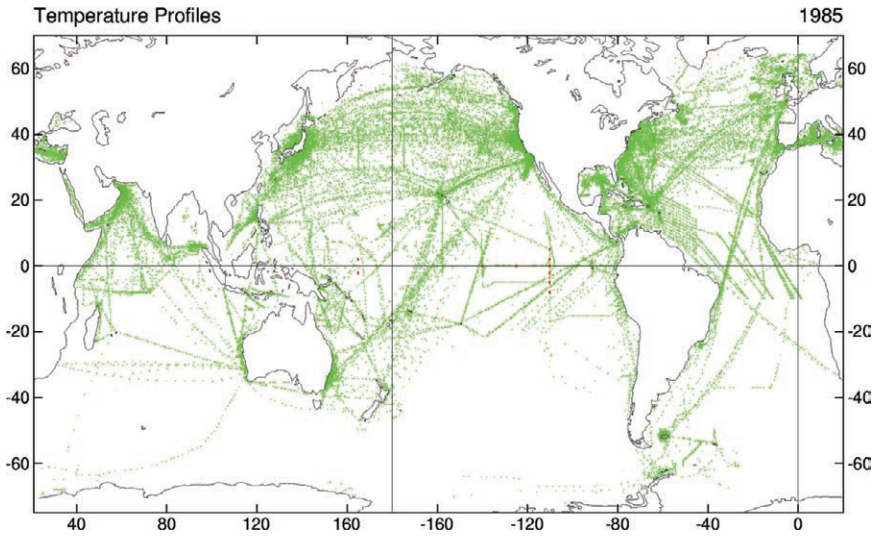

Fig. 12. The global distribution of all temperature profiles assimilated by the ocean component of the CFSR for the year 1985. The distribution is dominated by XBT profiles collected along shipping routes.

From the start of the CFSR through 31 December 1996, the global ice concentrations for most of the globe are simply regridded from Cavalieri et al. (1996, 2007; see GSFC Ice), except for 1) possibly ice-covered regions that lie outside that grid, 2) large Canadian lakes, 3) the Great Lakes, and 4) SST-based filtering of erroneous ice in the analysis.

For the Great Lakes, the data used are from Assel et al. (2002) from the start of the CFSR through the end of the dataset in spring 2002, and passive microwave thereafter. Those grids are available 1–3 times per week throughout the period when they are available. Concentrations were linearly interpolated between the observation dates, and those interpolated values are used here, averaged on to the target 0.5° grid from the native 2.55-km Mercator projection. For large lakes in Canada, the CIS analyses were used for all of the lakes that were analyzed (initially 34, in November 1995, increasing to 137 by October 2007) from November 1995 through 29 October 2007 (the start of CFSR processing). Again, concentrations were linearly interpolated in time between the observations. From 30 October 2007 to the present, the concentrations are the operational NCEP passive microwave sea ice concentration analyses (R. W. Grumbine 2010a, unpublished manuscript, hereafter GRUa).

Large water bodies that may freeze but that lie outside the domain analyzed in GSFC Ice (Cavalieri et al. 1996, 2007) in the 1978–96 period were analyzed by proxy, as was done for portions of the North American Regional Reanalysis (Mesinger et al. 2006). Proxies were generated anew for this work because the domain was much larger, and more data were now available. During the period from 1 January 1997 to 30 June 2006 (when both NCEP Ice and GSFC Ice were available), the NCEP ice analysis was used to identify points (one by one) that lay inside the GSFC ice domain and had high correlation to concentrations analyzed for points outside the GSFC ice domain but were still inside the NCEP domain. This includes large lakes, such as Lake Ladoga and Lake Onega, as well as the Caspian Sea. Because of changes in SST sources for filtering sea ice concentration analyses, the Aral Sea and Lakes Balkhash, Hulun Nur, and some others could not be consistently analyzed and were assigned zero ice concentration. This will bias surface heat fluxes high in periods where the real lake is ice covered. Some lakes cannot

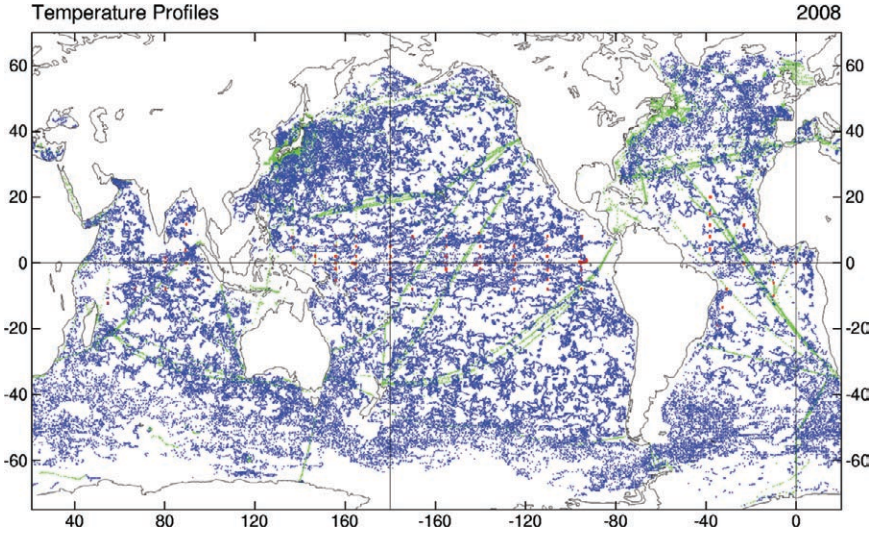

Fig. 13. As in Fig. 12, but for the year 2008. The Argo array (blue) provides a nearly uniform global distribution of temperature profiles.

be observed strictly by modern passive microwave because of land contamination issues, and they were assigned land flags in the CFSR if no other data were available; this includes Lake Athabasca, Lake of the Woods, Lake Nipigon (outside the period of CIS data), Iliamna Lake, and Lake Vanern.

For 1 January 1997–29 February 2000, the global ice concentration analysis was the NCEP operational ice analysis [Grumbine (1996); again, this is outside the Great Lakes and Canadian lakes]. From 1 March 2000 to 29 October 2007 the sea ice analysis is the newer NCEP sea ice analysis system (GRUa) that is applied to archived passive microwave data for DMSP *F-13*, *F-14*, and *F-15*. The old NCEP system was based on the NASA Team 1 algorithm (Cavalieri 1992), as was GSFC Ice (Cavalieri et al. 1996, 2007). The newer system is based on the Team 2 algorithm (Markus and Cavalieri 2000). In the newer NCEP system (GRUa), the sea ice concentration for each day and for each hemisphere (both northern and southern) is computed by regression of the Stokes-like parameter $(T85V^2 - T85H^2)^{0.5}$ (where T85V is the 85-GHz brightness temperature at vertical polarization, and T85H is brightness temperature for the horizontal polarization) against the Team 2–derived concentration for those points that are greater than 100 km from land and are poleward of 60° latitude. The regression provides an unbiased estimate and, because of the small footprint of the 85-GHz channel, a higher-resolution estimate, permitting analysis that is closer to the coast and inside smaller lakes than would otherwise be possible with the pure Team 2 algorithm. This operational system used the SSM/I on DMSP *F-13*, *F-14*, and *F-15*, while those were all available; *F-14* stopped providing data in October 2008 and *F-15* suffered progressively severe corruption of the 22-GHz channel (used for weather filtering, not for sea ice concentration computation) in late 2008 and was removed from NCEP sea ice production on 5 March 2009. AMSR-E was added to the operational sea ice system on 13 May 2009, using the AMSR-E Team 2 algorithm with January 2009 tie points, as described in Markus and Cavalieri (2009). That date was, unfortunately, concurrent with a data flow outage from AMSR-E and data corruption in *F-13*. This simultaneous failure degraded the quality of the sea ice analysis in May 2009. From June 2009 to the present, sufficient data have been available for analysis. February through April was reanalyzed retrospectively using *F-13* and AMSR-E, so that the CFSR ice is based on these, even though AMSR-E did not come in to the operational NCEP sea ice analysis until 13 May.

The passive microwave weather filters are imperfect, meaning that ice concentrations can be reported from the microwave for reasons other than ice being on the surface, so a sea surface temperature filter is also used (Grumbine 1996). The sea ice concentrations were, in general, produced before the SST analyses used for the CFSR. Therefore, an a posteriori filter was used for retrospective analyses through 29 October 2007 (Grumbine 2009). The usual SST filtering was also done using AVHRR-only analysis (Reynolds et al. 2007) for 4 January 1985–10 February 2000. The RTG low-resolution analysis (Thiebaux et al. 2003) was used from 11 February 2001 through 29 October 2007 and RTG high-resolution analysis (Gemmill et al. 2007) was used thereafter. A more detailed document on the CFSR sea ice analysis is in preparation (R. W. Grumbine 2010b, unpublished manuscript).

The sea ice model. The sea ice model is from the GFDL Sea Ice Simulator, with slight modifications. Its model grid is identical to the ocean model grid; there are three layers for the sea ice model, including two equal layers of sea ice and one layer of snow. In each ice grid there are five categories of possible sea ice thicknesses (0–0.1, 0.1–0.3, 0.3–0.7, 0.7–1.1 m, and the category greater than 1.1 m).

Sea ice dynamics are based on Hunke and Dukowicz (1997) using the elastic–viscous–plastic technique to calculate ice internal stress. Ice thermodynamics are based on Winton (2000). It is possible for ice to be transferred between the snow layer and the two ice layers, conservatively, when there is snowfall, evaporation, freezing, or melting. When sea ice forms over the ocean it releases latent heat and salt to the ocean. Details can be found in Griffies et al. (2004). Because of the lack of observations of sea ice thickness and motion covering the CFSR period starting in 1979, only the sea ice concentration is "assimilated" in CFSR. The 6-h model guess field and the analyzed sea ice concentration are used to produce a new initial condition at each analysis cycle. During the merging process, quality control is applied to prevent a failure when there is feedback between the sea ice analysis and the SST analysis; this is done on the model grid after an interpolation is made for SST and sea ice. When SST from the analysis is warmer than 275.3 K, no sea ice is allowed to exist. When the observed sea ice concentration is ≥15%, the sea ice concentration is reset to the observed value in the guess field. When the observed sea ice concentration is <15%, sea ice in the guess field is removed. In summer, the melt pond effect on ice albedo is considered in the Arctic. When there are serious problems in the observed sea

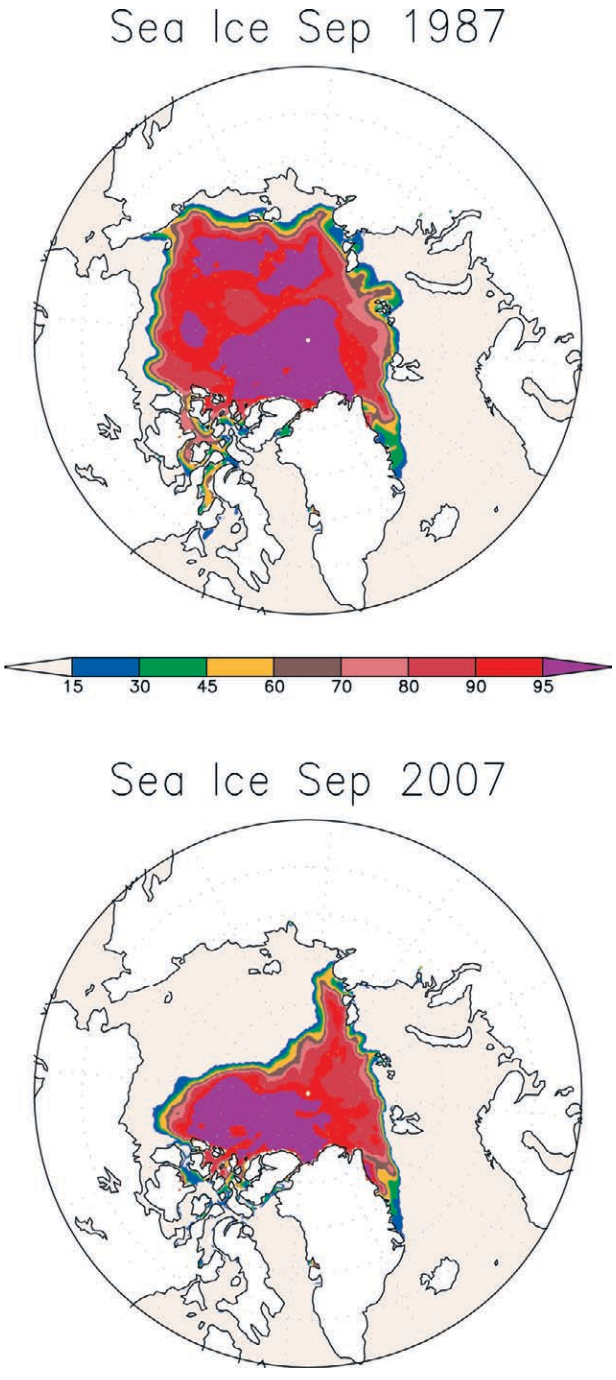

Fig. 14. Monthly mean sea ice concentration for the Arctic from CFSR (6-h forecasts) for Sep (top) 1987 and (bottom) 2007. (Units: %)

ice analysis, only the model prediction was used. This occurred during the period of 1–13 May 2009.

Because sea ice concentration is assimilated in the CFSR, the resulting initial conditions are close to the observations for sea ice concentration and sea ice coverage. Figure 14 shows the CFSR sea ice concentrations for September 1987 and 2007 for the Arctic (from 6-h forecasts). Record minimum Arctic sea ice was observed in September 2007. Sea ice extent, defined as the total area with sea ice present (including the open water) for which each grid cell has at least 15% sea ice, is shown in Fig. 15 for the Arctic and Antarctic in the CFSR (from 6-h forecasts). A large reduction in sea ice is obvious for the summers of both 2007 and 2008 over the Arctic. Interannual variability is relatively small for both hemispheres as a whole. The overall trend is slightly positive for the Antarctic and negative for the Arctic, which is consistent with previous studies. Because of the realistic sea ice distribution and other upgrades, the coupled CFSR has improved many aspects of the analysis of sea ice concentration over the polar regions, compared to the previous R1 and R2 (Wu and Grumbine 2010, unpublished manuscript).

THE COUPLER. *Parallel-programming model.* The CFS model, which runs on multiple processes with message-passing tools, uses a parallel-programming model called MPMD. There are three programs (executables) in the CFS: the atmospheric model (GFS), the ocean model (MOM version 4), and the coupler, each of which has its own data flow. The three programs run independently, but they exchange data as follows: the GFS runs on the atmospheric time step Δa (3 min), MOM version 4 runs on a *fast* sea ice time step Δi (also 3 min) for the sea ice model, and a slow ocean time step Δo (30 min) for both the ocean model and *slow* sea ice time step, while the coupler runs on a time step Δc, where $\Delta c = \max(\Delta a, \Delta i)$. At every coupler time step Δc, the coupler will receive data from both the GFS and MOM version 4 sea ice model and will send the needed data back to them, respectively. At every ocean time step, in addition to the data exchanged between GFS and sea ice, the coupler will also receive accumulated variables (fluxes) from the GFS and send them to the ocean model, while receiving data from the ocean and sea ice model and sending them back to the GFS.

Grid architecture. In the CFS, the atmospheric model GFS uses a different grid structure from the grids of the ocean model MOM version 4 and the MOM version 4 sea ice model. The CFS makes use of the atmospheric grid in the dummy atmospheric model (ATM) of MOM4. The atmospheric grid in the dummy AM is set to the same grid as the GFS; the CFS coupler redistributes the data received from the GFS grid to the MOM version 4 AM grid. Inside MOM version 4, the data on the grid of the dummy AM is interpolated with a mask to the MOM version 4 sea

ice grid through the exchange grid of the surface boundary layer. When the MOM version 4 sea ice model needs to exchange data with the ocean model, the data are redistributed between the two models. This structure is shown in Fig. 16.

Passing variables. In the CFS, both instantaneous and accumulated variables are exchanged between the GFS and MOM version 4 sea ice and ocean models, through the coupler. The instantaneous variables from the GFS to the sea ice model include downward shortwave and longwave radiation, and bottom model layer temperature, wind, humidity, pressure, and snowfall. The accumulated variables from the GFS to the ocean are net downward short- and longwave radiation, sensible and latent heat flux, wind stress, and precipitation. The variables sent from the sea ice/ocean back to the atmosphere are sea surface temperature, sea ice fraction, and thickness and snow depth.

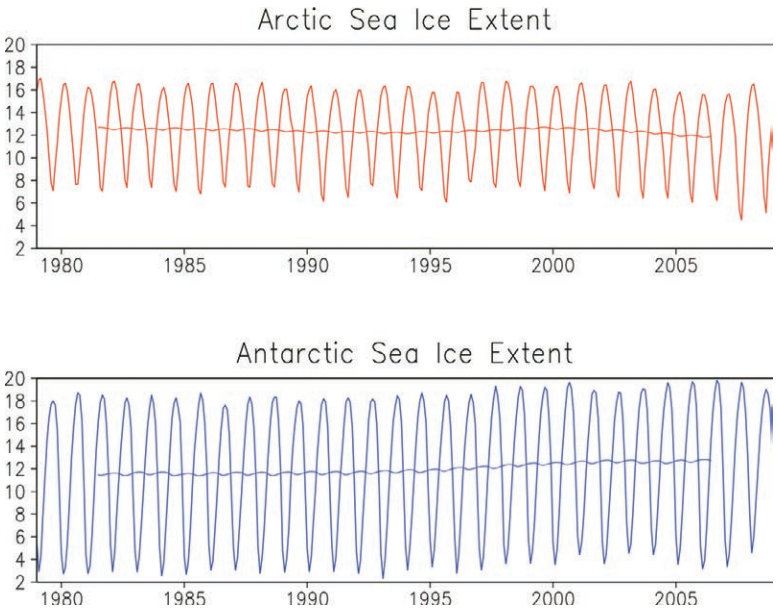

Fig. 15. Monthly mean sea ice extent (10^6 km^2) for the (top) Arctic and (bottom) Antarctic from CFSR (6-h forecasts). Five-yr running mean is added to detect long-term trends.

THE LAND. *The precipitation analysis.* Two sets of global precipitation analyses are used in the CFSR land surface analysis. The pentad dataset of CMAP (Xie and Arkin 1997) defines 5-day mean precipitation on a 2.5° latitude × 2.5° longitude grid over the globe by merging information derived from gauge observations, as well as satellite observations in infrared and passive microwave channels. The other dataset used is the CPC unified global daily gauge analysis, constructed on a 0.5° latitude × 0.5° longitude grid over the global land through the interpolation of quality-controlled rain gauge reports from ~30,000 stations collected from the GTS and many other national and international collections (P. Xie et al. 2010, unpublished manuscript). The OI algorithm of Xie et al. (2007) is employed to partially account for the orographic enhancements in precipitation. In addition to the analyzed values of precipitation, the number of reporting stations and the ending time of daily accumulation are also included in the dataset. Both analyses are generated for the entire CFSR analysis period from 1979 to the present and continue to be updated on a real-time basis.

The snow analysis. Snow liquid equivalent depth was updated using analysis data from the Air Force Weather Agency's SNODEP model (Kopp and Kiess 1996) and the NESDIS IMS (Helfrich et al. 2007). SNODEP uses in situ observations, an SSM/I-based detection algorithm, and its own climatology to produce a global analysis of physical snow depth once per day at 47-km resolution. Analysts may further adjust the analysis. SNODEP has been operational

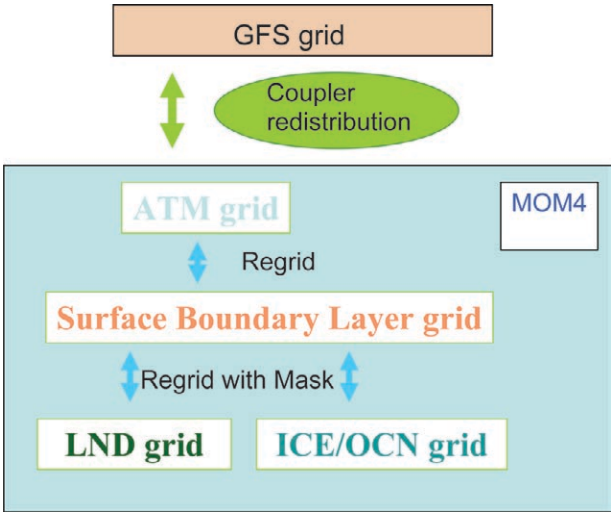

Fig. 16. CFS grid architecture in the coupler. AM: MOM version 4 atmospheric model (dummy for CFS), SBL: surface boundary layer where the exchange grid is located, LAND: MOM version 4 land model (dummy for CFS), ICE: MOM version 4 sea ice model, and OCN: MOM version 4 ocean model.

since 1975 and the data are available for the entire reanalysis period. The IMS data are a manually generated Northern Hemisphere snow cover analysis produced once per day. Analysts use surface data, geostationary and polar-orbiting imagery, and microwave-based detection algorithms to determine whether an area is either snow covered or snow free. IMS data are available at 23-km resolution starting in February 1997 and at 4-km resolution starting in February 2004.

The land surface analysis. The LSM used in CFSR is the Noah LSM (Ek et al. 2003), which was implemented in the NCEP GFS for operational medium-range weather forecast in 2005. Within CFSR, Noah is implemented both in the fully coupled land–atmosphere–ocean model to make the first-guess land–atmosphere simulation and in the semicoupled CFSR GLDAS to perform land surface analysis. The semicoupled GLDAS is forced with the CFSR atmospheric data assimilation output and observed precipitation. GLDAS interacts with the reanalysis once per day, instead of every time step, as in its fully coupled counterpart. The NASA LIS infrastructure (Peters-Lidard et al. 2007) is employed to execute CFSR–GLDAS. This semicoupled GLDAS LIS has been configured with the identical setup as in the fully coupled CFS–Noah, including the same T382 global Gaussian grid specification, land–sea mask, terrain height, soil and vegetation classes, and soil and vegetation parameters.

Compared to R1 and R2, this CFSR–GLDAS LIS uses observed global precipitation analyses as direct forcing to the land surface analysis, rather than the typical analysis approach of using precipitation from the assimilating background atmospheric model, or using observed precipitation to "nudge" soil moisture (R2). The pentad CMAP analysis and daily gauge analysis are used. Considering global gauge distribution and the strength of the satellite-based precipitation analysis, an optimal precipitation forcing is generated by blending the two precipitation analyses with the CFSR background 6-hourly GDAS precipitation. The blending function is latitude dependent, which favors the satellite-based CMAP analysis in the tropics, the gauge analysis in the midlatitudes, and the model precipitation in high latitudes. Moreover, an even heavier weight is assigned to the gauge analysis in regions of a dense gauge network, namely, North America, western Europe, and Australia.

Every LSM is characterized by a specific equilibrium land surface climatology. That equilibrium can be quite different from model to model. The spinup time required to drive an LSM to its equilibrium is much longer than for the troposphere. Experiments have estimated that at least 3–5 yr might be required to spin up the four-layer CFSR–Noah land surface states, if initialized with a previous global reanalysis in which a different LSM [e.g., the two-layer OSU LSM (Pan and Mahrt 1987) in R1/R2] was used. Because the same Noah LSM has been included in the operational GFS since 2005, the 2-yr (2006 and 2007)-averaged GFS land surface states for each given calendar day of the start dates of the CFSR streams were used as land initial conditions. An additional 12-month spinup period was executed prior to the CFSR production.

The CFSR–GLDAS LIS is executed once over each 24-h land surface analysis cycle at 0000 UTC model time, instead of the 6-h cycles of the atmospheric analysis in GDAS and the oceanic analysis in GODAS. The reason is that GLDAS LIS is anchored to the *daily* gauge precipitation analysis; hence, the gain of executing on a 6-h cycle is limited. The Noah simulation is made for the past 24 h using GDAS atmospheric forcing and the blended precipitation forcing. After completion of the 24-h execution, the simulated soil moisture and soil temperature of all four Noah soil layers are inserted into the CFSR 0000 UTC restart file (the so-called surface file) as the land surface initial conditions for the next CFSR 0000 UTC cycle.

The IMS and SNODEP data were used to produce daily analyses of physical snow depth on the model physics grid over land. The data were horizontally interpolated using a "budget" method (Accadia et al. 2003) in order to preserve the total water volume. In the Southern Hemisphere, and globally prior to February 1997, these analyses were created solely from the SNODEP data. In the Northern Hemisphere, starting in February 1997, a combination of SNODEP and IMS was used. IMS data were introduced because they more accurately depict snow cover compared to SNODEP, especially along mountain ridges (because of the higher resolution). Therefore, in regions where the IMS and SNODEP analyses did not agree, the IMS determined whether there was snow or not in the daily analysis. More specifically, if the IMS indicated snow cover, the analyzed depth was set to 2.5 cm or the SNODEP value, whichever was greater; and, if IMS indicated a region was snow free, the analyzed depth was set to zero.

The model snow was updated at 0000 UTC by comparing the first guess to the daily analysis. The analyzed physical depth was converted to liquid equivalent

depth using a 10:1 ratio. If the first-guess depth was greater than twice (or less than half) the analyzed depth, then the model depth was set to twice (half) the analyzed value. Otherwise, the model snow was not modified. In contrast to directly replacing the model snow with the analysis, this method results in a smoother evolution of the snowpack and reduces the artificial addition of water when the land surface model erroneously melts the snow too quickly. Daily analyses were not available for 65 days because of missing IMS or SNODEP data. On these days, the model snow was simply cycled. GLDAS LIS also updates its snow fields (snow liquid equivalent and physical depth) to the values of the snow analysis at 0000 UTC.

Figures 17 and 18 show the CFSR resulting global 2-m volumetric soil moisture climatology for May and November (1980–2008), respectively. It is consistent with our current knowledge about the large-scale soil moisture climatology of wet and dry regions. The contrast between May and November also illustrates the seasonal variation corresponding to precipitating and drying seasons of various regions, for instance, North America, the Amazon region, and India. Further discussion will be given in J. Meng et al. (2010, unpublished manuscript).

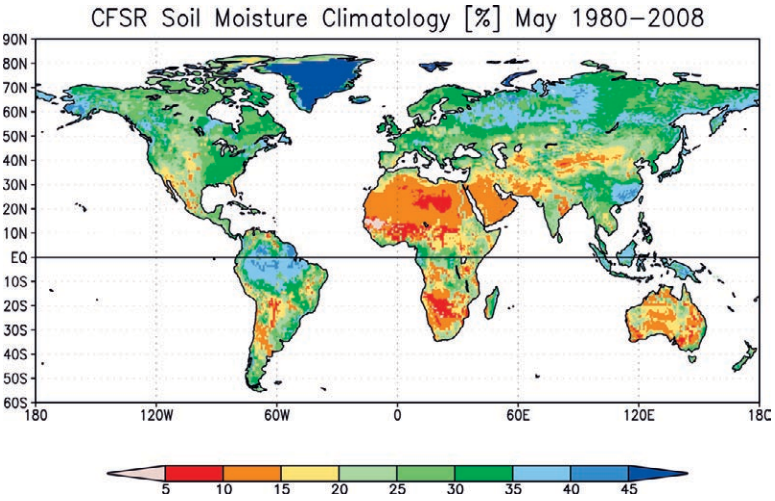

Fig. 17. The 2-m volumetric soil moisture climatology of CFSR for May averaged over 1980–2008.

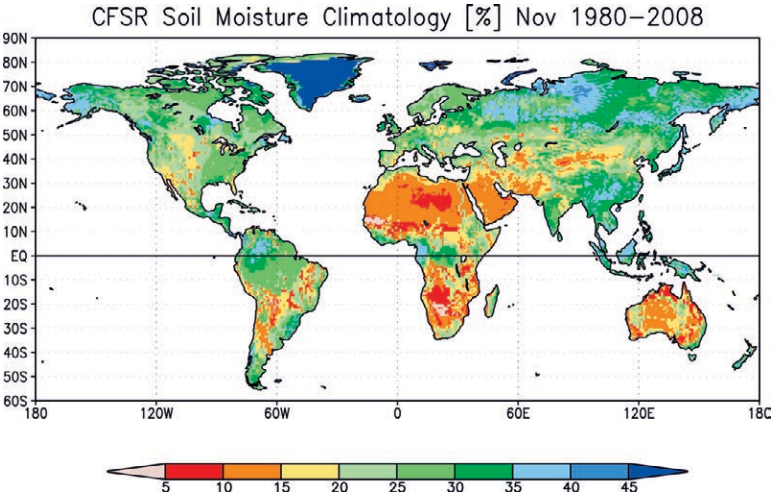

Fig. 18. As in Fig. 17, but for Nov.

PRELIMINARY RESULTS OF THE ATMOSPHERIC ANALYSIS. A set of papers analyzing the CFSR in depth will be submitted later; here we will give some preliminary results.

Climate trends. The motivation in the 1990s was to use reanalysis data for climate studies. However, in spite of the constant data assimilation system used in R1, artificial changes may have been introduced by ingesting data from constantly changing observational platforms, thereby leading to reservations about the use of R1 in the context of climate change detection (Chelliah and Ropelewski 2000). Aside from input data changes, the constant CO_2 of 330 ppmv, throughout R1 (from 1948 to the present) has raised concerns (Cai et al. 2009) about the utility of R1 and R2 for anthropogenic climate change studies. Because CFSR has increasing CO_2, we can study its time trace of global mean temperatures. Figure 19 compares the time series of the annual global mean land (2 m) temperatures in R1 and CFSR to the time series from observed GHCN CAMS data (Fan and Van den Dool 2008). It is obvious that CFSR has less bias than R1 and an upward trend that not only is stronger than in R1 but also appears to be quite realistic compared to GHCN CAMS. The linear trends are 0.66, 1.02, and 0.94 K $(31\ yr)^{-1}$ for R1, CFSR, and GHCN CAMS, respectively. What makes the comparison particularly compelling is that neither R1 nor CFSR assimilated any 2-m temperature data over land. In both systems, the upward trend is the result of the assimilation of other types of data (mainly at upper levels). In addition, by including increasing CO_2 specifications in

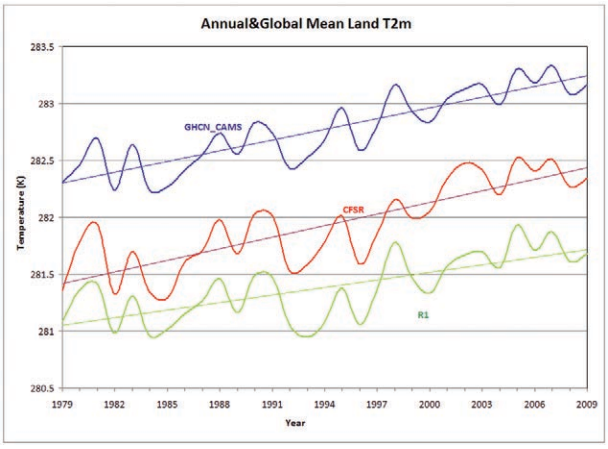

Fig. 19. The annual global mean 2-m temperature over land in R1 (green), CFSR (red), and GHCN CAMS (blue) over the period of 1979–2009. Units: K. Least squares linear fits of the three time series against time (thin lines). The linear trends are 0.66, 1.02, and 0.94 K (31 yr)$^{-1}$ for R1, CFSR, and GHCN CAMS, respectively. (Keep in mind that straight lines may not be perfectly portraying climate change trends).

the CRTM calibration of satellites and in the forecast model, the trend in the CFSR almost matches the independent observed trend. It thus appears that CFSR may be more useful for climate change studies than either R1 or R2. It is also apparent that data assimilation alone captures only 2/3 of the observed upward temperature trend over land. The interannual detrended anomalies in all three datasets are very similar (not shown). Both R1 and CFSR display upward trends over the oceans (not shown) of about 0.3 K over 1979–2009. This is much less than over land, a puzzle that is yet to be fully understood.

Medium-range forecast skill. An integral part of the CFSR job suite was a once-daily-at-0000-GMT 120-h medium-range global prediction run at the CFS version 2 resolution of T126L64. The primary purpose of the forecast is to monitor the forecast scores as a measure of quality of the CFSR analyses. A standard measure of medium-range forecast skill is the hemispheric 500-hPa (AC), taken for both hemispheres (NH and SH). The time series of 120-h scores could be assessed against those compiled by the operational GFS, R1/CDAS, and the set of R2/CFS hindcasts and the CFSR-Lite.

Global reanalysis forecast skill comparisons. The daily 120-h forecast scores for each hemisphere have been averaged for each year in the period of 1979–2008 for the CFSR, CFSR-Lite (through 2006), operational GFS, CFS R2 hindcasts (1981–2006), and CDAS/R1. The NH scores are on the right-hand side of Fig. 20; the SH scores are on the left-hand side. The configuration of the systems was as follows:

- CFSR: T382L64 (GSI)
- CFSR-Lite: T62L64 (GSI)
- CFS R2 hindcasts: CFS, T62L64; R2, T62L28 (SSI)
- CDAS/R1: T62L28 (SSI)
- GFS: 1984, R40L12; 1985, R40L18; 1987, T80L18; 1991, T126L28; 2000, T170L42; 2002, T254L64; 2005, T382L64.

Please note that the resolution changes of the GFS may include changes to the forecast model and/or data assimilation. A history of changes to the NCEP GFS/GDAS may be found online (at www.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html).

NH results. The CDAS/R1 scores are remarkably steady over the past 30 yr, with very slight improvement since 2000, which may be related to the automation of rawindsondes, the debut of ACARS ascent and descent profiles, and the improvement of ATOVS retrievals over those of TOVS. Note the comparable scores for CDAS/R1 (blue) and the operational GFS (red) in the mid-1990s, the period when R1 was assembled and run based on the GFS of that era. The CFS R2 hindcast scores are a slight improvement over CDAS/R1, which may be attributed to the increase of vertical resolution from 28 to 64 layers in the CFS predictions and ex post facto hindcast bias correction. CFSR-Lite has a modest but clear improvement over the CFS R2 hindcasts, a result of the following multiple factors: 1) a vertical resolution increase of the assimilation system, from 28 to 64 layers; 2) direct assimilation of radiances replacing retrievals; 3) GSI versus SSI; and 4) forecast model improvements between 2003 and 2007. The CFSR represents a significant improvement over the lower-resolution systems for the entire period of 1979–2008 and demonstrates that the level of skill of the comparable operational GFS has been maintained since 2000. The dramatic jump in CFSR skill over the CFSR-Lite emphasizes the importance of horizontal resolution in global data assimilation. The slight improvement of the operational GFS over the CFSR, since the GFS resolution increased to T382L64 in 2005, may well reflect the importance of resolution in the medium-range prediction model. The upward slope of the CFSR scores reflects both the conventional data improvements noted above and the TOVS-to-ATOVS radiance transition.

SH RESULTS. The most noticeable difference between SH and NH scores (Fig. 20) is the slope of the increase in skill over the period of 1979–2008 compared to those for the NH. The SH scores may be viewed partly as a proxy for oceanic observing systems (e.g., satellite winds and radiances), and the NH as a proxy for land-based systems (e.g., rawindsondes and ACARS profiles). The CDAS/R1 and GFS scores were not consistently sustained above the 0.60 skillful level until the early 1990s. As in the NH, CFS R2 hindcasts only marginally improved over CDAS/R1. The quantum leap in skill came with CFSR-Lite scores, which emphasized the importance of the direct assimilation of radiances, with a comparable level of skill maintained throughout the TOVS period of 1979–98. The superiority of the ATOVS instruments over the TOVS is noted by the jump in CFSR-Lite scores since 2000. The CFSR scores are modestly better than the CFSR-Lite scores for the TOVS period, but they are dramatically better in the ATOVS period, a suggestion that the higher-resolution CFSR and GFS assimilation systems are exploiting information in the ATOVS systems that is not available from TOVS. Note that the NH and SH CFSR and GFS scores at the end of the period are nearly comparable, an indication that the medium-range skill of the prediction of synoptic-scale features is now comparable over oceans and continents.

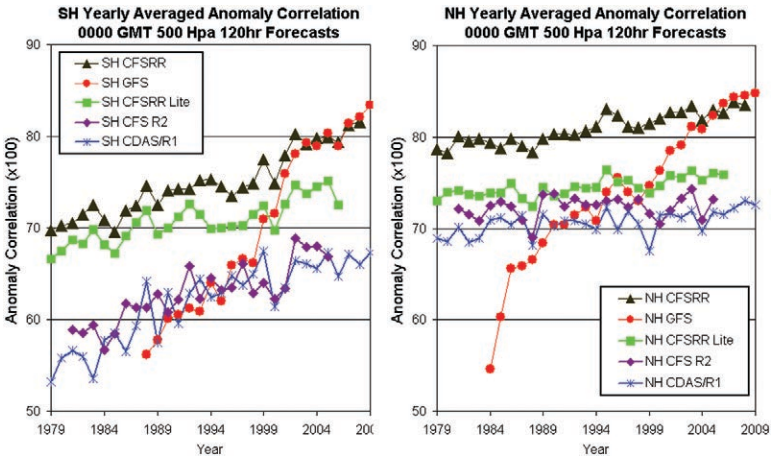

FIG. 20. Yearly averaged (left) Southern Hemisphere and (right) Northern Hemisphere 0000 GMT 120-h forecast anomaly correlations for CFSR (black triangles), GFS (red circles), CFSR-Lite (green squares), CFS R2 (purple diamonds), and CDAS R1 (blue stars).

Atmospheric mass. ATMOSPHERIC MASS BALANCE. Monitoring atmospheric mass as part of a reanalysis effort is now common; see Fig. 10 in Uppala et al. (2005) and Fig. 1 in Trenberth and Smith (2005), which compare traces of surface pressure, precipitable water, and dry pressure in ERA-40, ERA-15, and R1. Compared to previous reanalyses, the CFSR is very well behaved in terms of the atmospheric mass or, very nearly equivalent, the surface pressure. One check is to study the time variation of the "dry pressure," which should nearly be conserved. The dry component of surface pressure is defined locally by subtracting the precipitable water (in all of its forms) from the total surface pressure, converted to pressure units. Indeed, the global mean of dry pressure is almost constant, with a standard deviation of 0.10 hPa around its mean value of 983.01 hPa over the 1979–present period (not shown). The global mean of the total pressure varies somewhat more, with a standard deviation of 0.16 hPa around its mean of 985.39 hPa, because the water content of the atmosphere is free to vary. However, even this variation is less than 1 hPa from maximum to minimum in the global mean. In earlier reanalyses (Trenberth and Smith 2005), these variations were much larger and a consistency between total pressure and precipitable water was lacking.

The mass balance of the atmospheric water component, with its input by evaporation (E) and output by precipitation (P), remains somewhat worrisome, even in the CFSR. Global average monthly mean P, E, and $E–P$ are shown in Fig. 21. The global mean P is always larger than E (by a non-negligible few tenths of millimeters per day) and this imbalance grows around 1998, probably related to the ingest of new data systems, like AMSU. The decrease of global mean $E–P$ after 1998 appears to be due to a change over the oceans (Figs. 21b,c) and results from both an increase in precipitation (Fig. 21a) and a decrease in evaporation (Fig. 21b). Another feature is an increase in the amplitude of the seasonal cycle of $E–P$ over land after 1999 (Fig. 21c, blue curve), which appears to be due to a weaker seasonal amplitude in precipitation (Fig. 21a). Further analyses are required to investigate the causes of these features in the CFSR. Assimilation of observations could violate the governing equations, and the hydrological imbalance in the CFSR is one prominent example of this violation.

ATMOSPHERIC TIDES. We now discuss the 2D atmospheric mass distribution in one aspect, namely, the atmospheric thermal (solar radiation induced) tides in the CFSR. The previous reanalyses have had many types of users. Among them, there are geodesists and

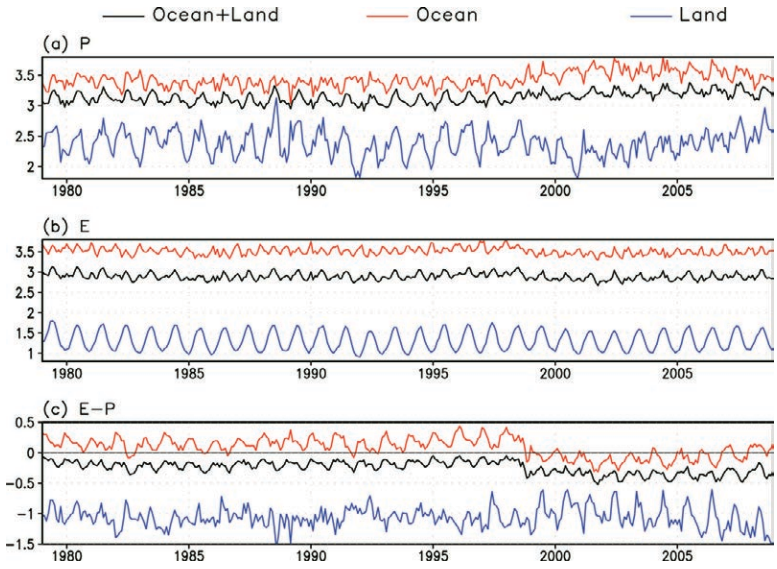

Fig. 21. Global average of monthly mean (a) precipitation, (b) evaporation, and (c) evaporation minus precipitation. Averages over ocean (red), land (blue), and ocean plus land (black) are plotted. (Units: mm day^{-1})

oceanographers (Ponte and Ray 2002; Ray and Ponte 2003; Velicogna et al 2001; de Viron et al. 1999), who require global atmospheric surface pressure, especially atmospheric tides. Compared to R1, the new CFSR offers several very significant advances for such users. First, and perhaps foremost, is the availability of hourly output. Figure 22 gives an example of the solar tides in CFSR for March 1998 (an arbitrarily picked month). The 24 global maps in one figure give a compact display of the tides. Each map is obtained from data available on the model T382 Gaussian grid. In the red (blue) areas, the pressure is higher (lower) than the daily mean. The data at 0000, 0600, 1200, and 1800 UTC are from the actual analysis made at these times, while the data in between these cycles are from the coupled guess forecast. One can clearly see the red and blue areas propagate westward around the Earth in 1 day. A cross section along the equator shows a dominant wavenumber two pattern, corresponding to the semidiurnal solar oscillation. This type of map is available for each month from 1979 to the present at the CFSR Web site (http://cfs.ncep.noaa.gov/cfsr).

Most reanalyses have output only every 6 h; imagine Fig. 22 with only the 0000, 0600, 1200, and 1800 UTC entries present. Previously, interpolation methods had to be devised (Van den Dool et al. 1997) to extract hourly tides from 6-hourly R1 data. While these methods worked satisfactorily, it is better to obtain hourly data directly from the assimilating model. Thanks to larger data storage devices, this can be done now. Nevertheless, users should keep in mind that only the 0000, 0600, 1200, and 1800 UTC fields are actual analyses, while the in-between hourly data are model forecasts. The tides appear accurate in terms of their amplitude (the phase error was never a big problem). R1 had a semidiurnal variation with an amplitude about 40% too strong (see Table 1 in Dai and Wang 1999), compared to the global station analysis by Haurwitz and Cowley (1973). This amplitude error has disappeared completely in the CFSR.

Accuracy of the surface pressure analysis. The accuracy of the surface pressure analysis in the CFSR, as expressed by computing error bars on the analyzed surface pressure fields, is difficult to judge directly. It is too expensive to conduct logical cross-validation experiments, such as making a reanalysis with, say, 10% of the data being withheld in turn, to judge accuracy. Instead, we discuss the root-mean-square difference of the 6-h forecast guess against the surface observations. This measure is only slightly higher than the accuracy of the analysis. Figure 23 shows annually compiled fit-to-obs statistics for SH ocean (blue) and NH land (red) for 1979–2008. The accuracy of the surface pressure analysis over the SH ocean has steadily improved, from an error of 2 hPa in 1979 to only about 1.0 hPa at present. The improvement is gradual and relates to a gradual increase in observations (and not just surface pressure observations) over the SH oceans, which were sparse before there were satellites. At the end of the period, the error in the analysis over the SH oceans is on par with that over the NH land. Figure 23 shows that NH land has an error of about 1.1 hPa from 1979 to 1996, without the gradual improvement seen in the SH. Then, in May/June 1997, a sudden large increase in METAR data helped reduce the error in the NH to below 1 hPa. In general, one may conclude that the accuracy of surface pressure analyses in the hemispheres (whether it be land or ocean) is getting closer. This has been noticed when real-time operational 5-day anomaly correlation scores for the two hemispheres became comparable, after a long history of lagging scores in the SH (see red lines for the GFS in Fig. 20). A more thorough analysis of atmospheric mass in the CFSR is forthcoming in a separate paper (van den Dool and Saha 2010).

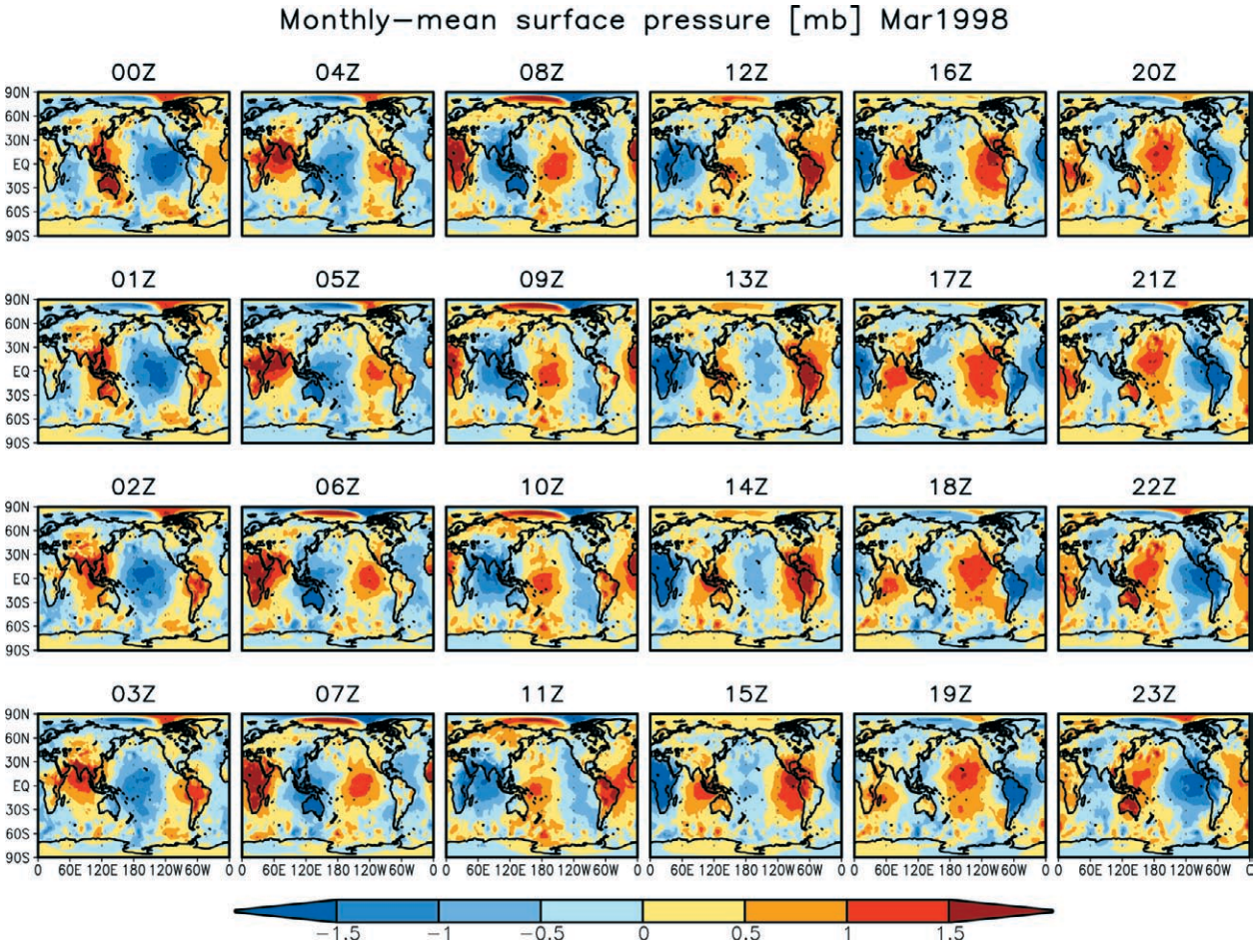

Fig. 22. Monthly mean hourly surface pressure with the daily mean subtracted for the month of Mar 1998 in a collage of 24 global maps. In red (blue) areas, pressure is higher (lower) than the daily mean (units: hPa). Time starts in the upper left (0000 UTC), and then proceeds down to 0300 UTC, and then continues at the top of the next column of the global maps, and so on.

The upper atmosphere. The CFSR is an improvement upon R1 and R2 because of increased horizontal (T382 versus T62) and vertical (L64 versus L28) resolution. The L28 model has only eight layers above 100 hPa, with none being above 1 hPa. By contrast, the L64 model has 22 layers above 100 hPa, with 2 layers being above 1 hPa. Because of the vertical limitations, R1 and R2 only produce usable analyses up to 10 hPa. However, the CFSR can produce useful analyses up to 1 hPa.

Another notable difference between R1 and/or R2 and the CFSR is that the CFSR GSI uses satellite radiances rather than derived temperature or moisture profiles. This allows the GSI greater freedom to generate adjustments to the temperature, moisture, and ozone fields to best match the observed radiances. The use of satellite radiances from ozone-sensitive IR channels mandates the need for quality ozone profile information.

Ozone, however, is one of the few parameters that the GSI assimilates as a physical quantity rather than as radiance. The CFSR assimilates the version-8 SBUV profile and total ozone retrievals (Flynn et al. 2009) from *Nimbus-7* and SBUV/2 profile and total ozone retrievals from *NOAA-9*, *-11*, *-14*, *-16*, *-17*, *-18*, and eventually *NOAA-19*. The ozone layers and total ozone values assimilated into the CFSR have not been adjusted to account for biases from one satellite to the next. Hence, the ozone products generated from the CFSR should not be used for trend detection.

Previously, it was shown how the number of radiosonde reports per year and the number of sondes reporting above 100 hPa increased in the late 1990s. The GSI had difficulty determining equatorial winds because of the low density and quality of radiosonde observations prior to 1998. Hence, the CFSR had difficulty analyzing a proper QBO and SAO wind signal before that time. ERA-40 winds above 20 hPa

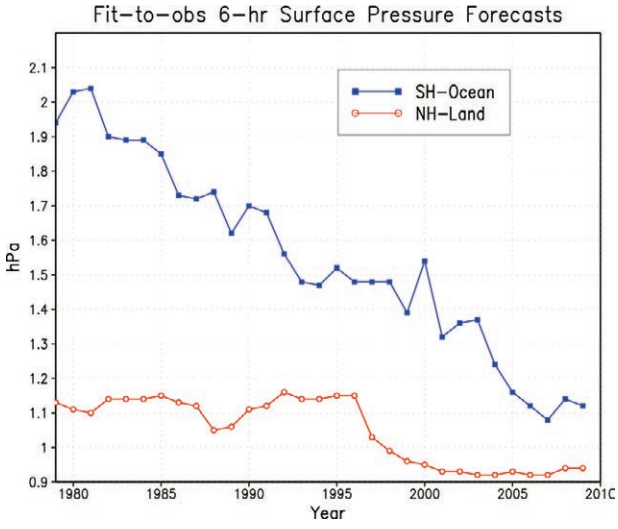

FIG. 23. The fit of 6-h forecasts of instantaneous surface pressure against irregularly distributed observations. Shown are annually compiled fit-to-obs data for 1979–2009 (hPa). SH ocean (blue) and NH land (red) are shown.

from 20°S to 20°N were assimilated until 1998, when AMSU data became available and radiosonde coverage was better. Figure 24 shows the global mean temperature anomalies from 1000 to 1 hPa from January 1979 through May 2009. Issues of using the SSU data from 1979 to 1998 have been evident in both the ERA-40 and JRA-25 temperature anomalies (Onogi et al. 2007). Extensive collaboration between the JCSDA, NASA, and ECMWF resulted in minimizing the effects of the SSU CO_2 pressure changes and the frequent loss of satellite channels. However, in the CFSR assimilation process, all three of the SSU channels were subject to bias correction. Unfortunately, these corrections, of the highest channel (3) in particular, resulted in feedback that produced a gradual warming of almost 10°C in the upper atmosphere over time. The end and beginning of each stream is quite evident as warm anomalies are transitioned to cold biases. The upper-atmosphere anomalies are greatest at the lowest pressures and reach down to 30 hPa. These anomalies decrease in depth after 1998, as the CFSR progresses to the present. Positive temperature anomalies from El Chichón (April 1982) and Mt. Pinatubo (June 1991) volcanic eruptions are apparent in Fig. 24 between 100 and 30 hPa, and they last about 2 yr. Evident in the troposphere is a gradual change from negative anomalies to positive anomalies between 700 and 200 hPa.

Figure 25 shows the equatorial (5°S–5°N) zonal mean *U* wind component time series from 1979 to 2009. Evident are the annual equatorial easterlies in the troposphere. The QBO variability of the *u* component is evident in the lower stratosphere, while the SAO is evident in the upper stratosphere to lower mesosphere. As earlier stated, ERA-40 stratospheric wind profiles were used as bogus observations from July 1981 through December 1998. As a result, a difference plot (not shown) between the CFSR and ERA-40 winds would show very little difference at pressures less than 20 hPa between 30°N and 30°S.

To see if the SBUV(/2) ozone profile data are correctly assimilated into the CFSR, differences between the monthly mean CFSR total ozone and that of the observed SBUV(/2) total ozone matching the satellites and time periods used in the CFSR are presented in Fig. 26. Only 67% of the differences are within an allowable 2 DU of the observed total ozone values. Most of the larger differences occur in the high latitudes and tropics. Further investigation revealed that most of the difference between the CFSR and observed SBUV(/2) ozone occurs above 10 hPa. At that point it was discovered that the ozone layer observational background errors in the CFSR were set too large in the upper stratosphere by as little as 2 times at 10 hPa to as much as 60 times at 0.2 hPa. The result is that SBUV(/2) ozone layer observations do not alter the CFSR's first guess above 10 hPa, and the model climatology is used. Although the integrated effect upon the total ozone is small (a few Dobson units), this does account for most of the differences we see in Fig. 26 at the high latitudes and in the tropics.

Future reanalyses would benefit from not bias adjusting the SSU radiances and having fewer streams. Improvements to the GSI to handle sparse data in the tropics would result in a better analysis of the QBO and SAO winds. Intersatellite bias adjustments must come from improved versions of the SBUV(/2) ozone data record. How the GSI uses the ozone data and alters its first guess must be explored further.

PRELIMINARY RESULTS FROM THE OCEAN. *Precipitation–SST relationship.* Previous studies have shown that there is a close relationship between precipitation and SST on intraseasonal time scales in the tropical Indian Ocean and in the western Pacific Ocean. In Fig. 27, we show the temporal lag correlation coefficient between precipitation and SST in the tropical western Pacific region in two generations of NCEP reanalysis data. Data for the boreal winter (November–April) over the period of 1979–2008 are bandpass filtered for 20–100 days after removing the climatological mean.

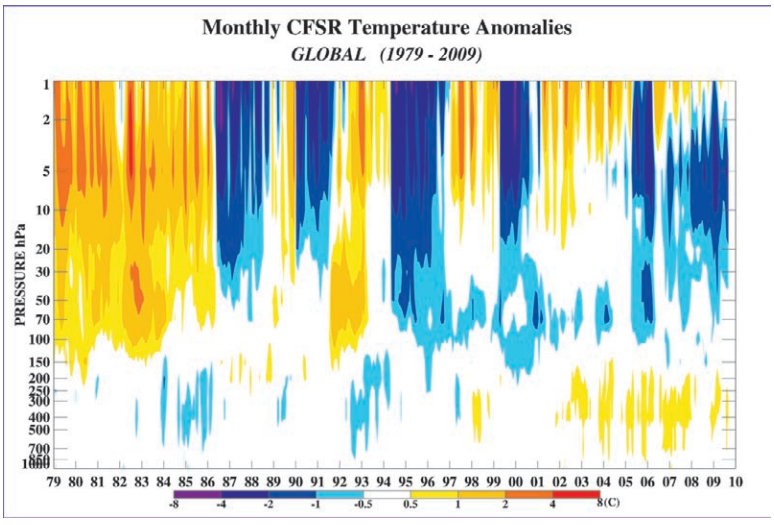

Fig. 24. Global mean temperature anomalies from 1000 to 1 hPa from Jan 1979 through May 2009. (Units: K)

Clearly, the precipitation–SST relationship in CFSR is consistent with observational data: at lag 0, precipitation has a weak negative correlation with SST. The positive correlation of SST and precipitation gradually increases with a warming of the SST (negative lags on the horizontal axis in the figure), reaching a maximum at lag day –7. On the other hand, the cooling of SST by precipitation reaches a maximum at lag day 5 (positive lags on the horizontal axis in the figure). In R1 and R2, the increase in precipitation resulting from warming of the SST is too quick because of a lack of coupling in the data assimilation system or a problem with using an observed SST that was a weekly averaged product. In the coupled CFSR, this deficiency is largely corrected. These results are consistent with the observational study of Woolnough et al. (2000) and the coupled versus uncoupled model comparison studies of Pegion and Kirtman (2008) and Fu and Wang (2004).

The oceanic component. This subsection focuses on comparisons between the oceanic component of CFSR and observations, globally and in point-wise fashion, for the equatorial section.

Equatorial cross sections of temperature are shown for the CFSR in the top panel of Fig. 28 (for the years 1982–2008), and in the bottom panel its differences from the *WOA05* (Locaranini et al. 2006) are shown. The 20°C isotherm, which at the equator is a proxy for the location of the seasonal thermocline, slopes upward toward the east, from approximately 150-m depth in the western equatorial Pacific Ocean to approximately 20-m depth in the equatorial cold tongue region. The difference plot indicates a more diffuse thermocline for the CFSR than in the observations, such that the CFSR section is cooler than the *WOA05* section above 20°C and warmer below. Large differences (>1°C) are found in the thermocline in the equatorial Atlantic, in the western Pacific warm pool region, and in the Pacific cold tongue region.

The corresponding equatorial cross section of the zonal velocity in the CFSR is shown in Fig. 29. The isopleths slope upward toward the east in the equatorial Atlantic, Pacific, and Indian Oceans. The core of the undercurrent is approximately centered on the 20°C isotherm, and the maximum mean velocity is about 0.85 m s^{-1} at 130°W in the equatorial Pacific. The strong westward flow in the near-surface layers are accompanied by a weaker, but broader, eastward flow below. Figure 30 shows differences between the vertical profiles of temperatures and zonal velocities of CFSR and TAO mooring data in the equatorial Pacific Ocean. Clearly, the CFSR is colder than the observations below the seasonal thermocline in the western and central equatorial Pacific, and warmer than the observations above the thermocline in the eastern equatorial Pacific. The magnitude of the undercurrent in the CFSR is about 10%–15% less than that observed. The CFSR also has difficulty capturing

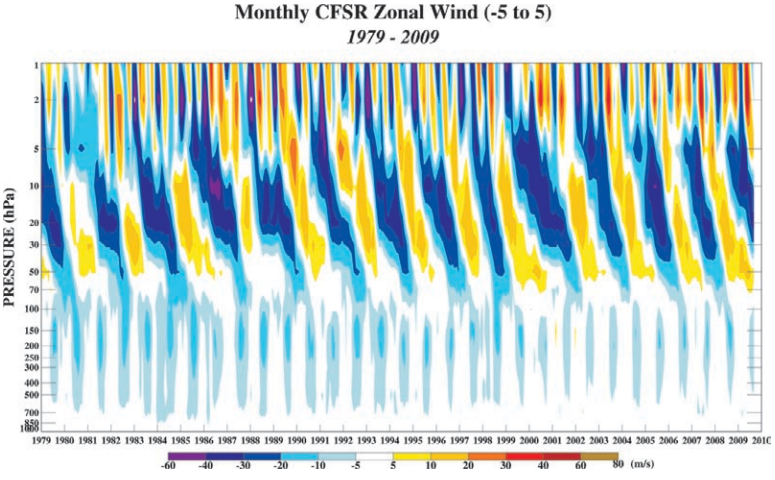

Fig. 25. Equatorial (5°S–5°N) zonal mean *U* wind component time series from 1979 to 2009. (Units: m s^{-1})

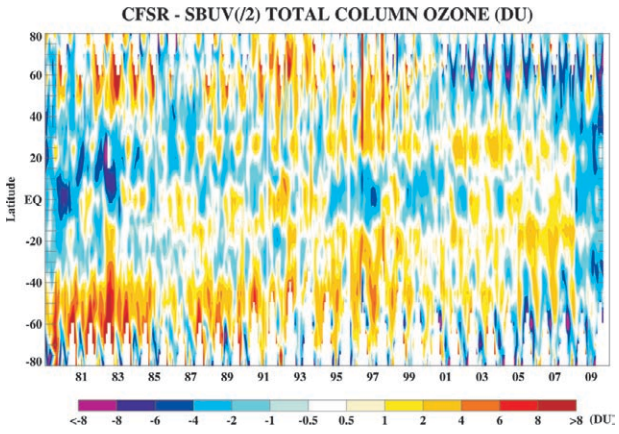

Fig. 26. Monthly zonal mean total ozone differences (DU) between the CFSR and observed monthly zonal mean total ozone from *Nimbus-7* SBUV (Jan 1979–Dec 1989), *NOAA-11* SBUV/2 (Jan 1990–Aug 1993), *NOAA-9* SBUV/2 (Sep 1993–Jun 1997), *NOAA-11* SBUV/2 (Jul 1997–Dec 2000), *NOAA-16* SBUV/2 (Jan 2001–Dec 2004), and *NOAA-17* SBUV/2 (Jan 2005–Dec 2009). SBUV(/2) total ozone observations cannot be generated in the polar night regions, hence a difference value of zero is assigned to these latitudes and months.

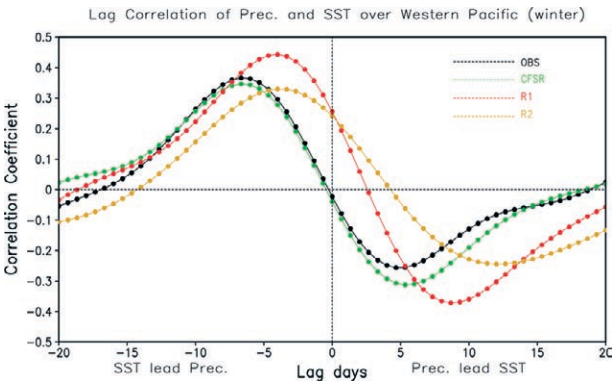

Fig. 27. Temporal lag correlation coefficient between precipitation and SST in the tropical western Pacific (averaged over 10°S–10°N, 130°–150°E) in R1 (red), R2 (brown), CFSR (green), and observation (black). GPCP daily precipitation and Reynolds ¼° daily SST are used as observational data. Negative (positive) lag in days on the x axis indicates the SST leads (lags) the precipitation. Data for the boreal winter (Nov–Apr) over the period 1979–2008 are bandpass filtered for 20–100 days after removing the climatological mean.

the near-surface westward flow in the central basin and the slow eastward flow in the western equatorial Pacific that is seen in the TAO data.

The differences between the vertically averaged temperatures (from the surface to 300-m depth) in CFSR and observations from the 2005 World Ocean Atlas (Locarnini et al. 2006) for 1979–2008 are shown in Fig. 31. The CFSR is colder almost everywhere by approximately 0.5°–1°C between 30°S and 30°N. While the TAO velocity data allow pointwise comparisons in the equatorial Pacific, the surface drifters drogued at 15 m (Lumpkin and Pazos 2006) are useful for global comparisons of surface velocities of CFSR to pseudo-observations, as shown in Fig. 32. This is the case even though the CFSR maps were made from averages of Eulerian velocities on the model grid, while the drifter maps were constructed from the pseudo-Lagrangian motion of drifters nonuniformly distributed in time and space. The drifter velocities are distinctly stronger (larger eastward velocities) than CFSR in the Antarctic circumpolar current, the western boundary currents, and the equatorial zone of the Indian Ocean. Also, the drifter velocities are more divergent than the CFSR in the cold tongue region of the eastern Pacific.

The spatial patterns of the first two EOFs of the CFSR SSH are shown at the top of Fig. 33 for

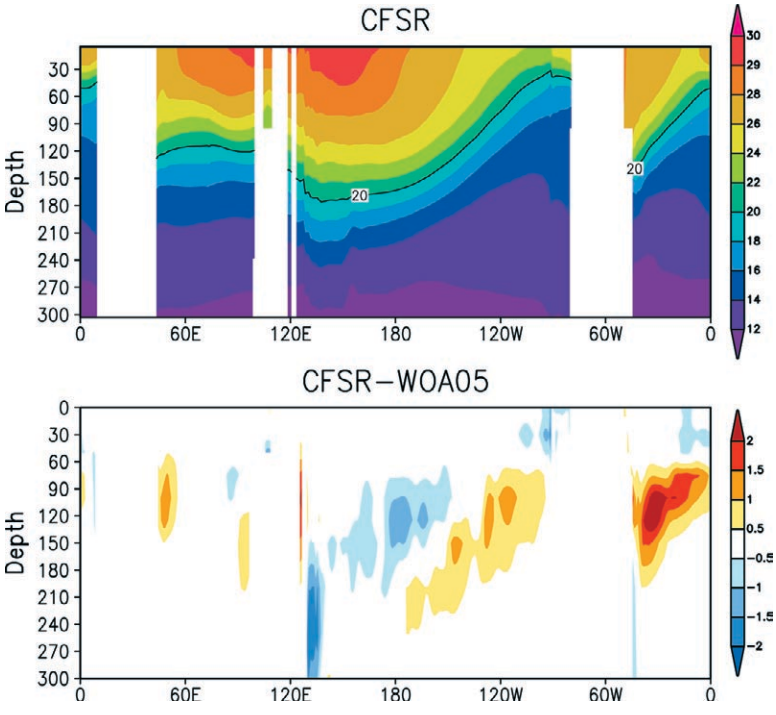

Fig. 28. The (top) subsurface temperature mean for an equatorial cross section for CFSR and (bottom) differences between CFSR and observations from the *WOA05* (Locarnini et al. 2005). (Units: K)

satellite altimeter data from TOPEX/*Jason-1* in the middle, while the bottom panels show the temporal amplification factors for the first two EOFs for CFSR and TOPEX/*Jason-1* date. The EOFs were computed for the period of 1993–2008, the period for which TOPEX/*Jason-1* data are available. The first EOF for CFSR and TOPEX/*Jason-1* data represents the warm phase of ENSO, as is indicated by the maximum value in 1997, while the second EOF represents the cold phase of ENSO, as indicated by its maximum value in 1998. A comparison of the first two EOFS for CFSR SSH and upper-ocean heat content (not shown) highlights the dominance of thermosteric variability of SSH over other physical processes. The major difference between the EOFs of CFSR and TOPEX/*Jason-1* is the double maximum found in the spatial plot of the first EOF of CFSR when compared to TOPEX/*Jason-1*. Also, the first EOF represents only 32% of the variability in CFSR compared to 42% for TOPEX/*Jason-1*. The second EOF is very similar for both CFSR and TOPEX/*Jason-1*.

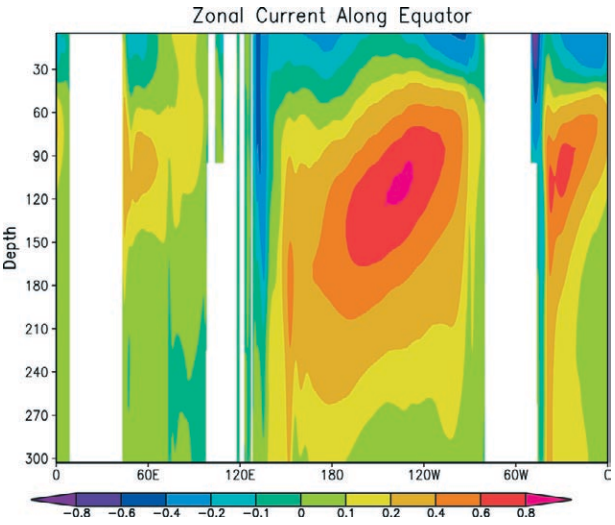

Fig. 29. The subsurface zonal velocity for an equatorial cross section for CFSR for the years 1979–2008. (Units: m s^{-1})

CONCLUDING REMARKS. In this paper, we have described a new global reanalysis (CFSR) produced at NCEP covering the period from 1979 to the present. Sections in this paper describe the observational datasets, forecast models, and data assimilation systems used in this endeavor. The data itself, and its dissemination by NCDC and NCAR, are discussed in detail in the sections "Data access" and "Data description" of the online supplement. We also briefly present some preliminary results. The accuracy of this new reanalysis is hard to assess directly but, using the 5-day forecast scores as a measure of the accuracy of initial states (a reasonable assumption), we conclude that CFSR is considerably more accurate than the previous global reanalysis made at NCEP in the 1990s. It is more comprehensive because it includes analyses of both the ocean and sea ice, and it has higher resolution in space and time. The accuracy increases over time, especially in the Southern Hemisphere, where the use of satellite radiance data becomes very important.

Many known errors in the observational data ingest and execution of previous reanalyses were corrected in the CFSR. Many of the input datasets have been improved by years of quality control and by exposure to successive reanalyses at various centers, most notably at ECMWF, NCEP, JMA, and NASA. Undoubtedly, some errors may still persist and more errors may be discovered in the future, but this is all part of the converging process of making increasingly improved analyses of the Earth's system. A thorough study is

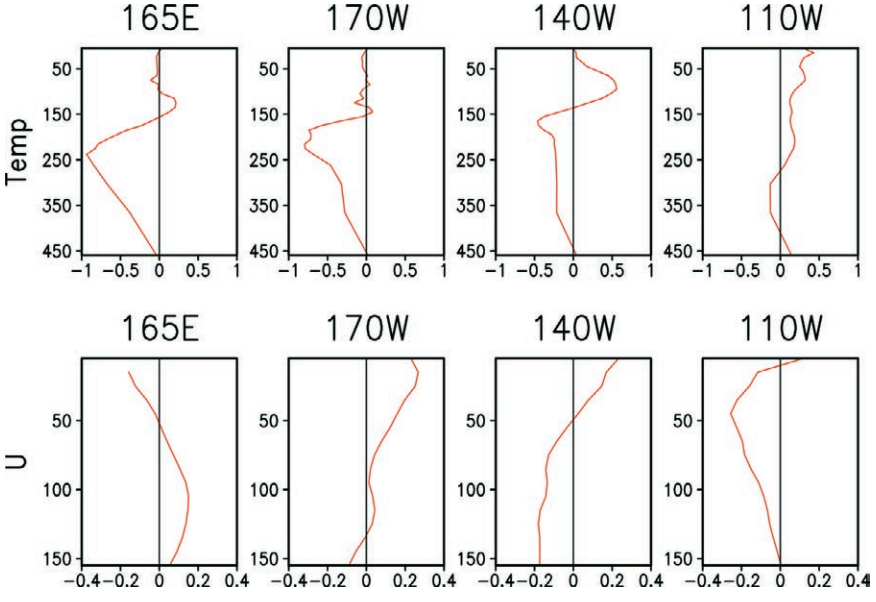

Fig. 30. Vertical profiles of the differences between CFSR and TAO observations for (top) subsurface temperature and (bottom) zonal velocity for four locations in the equatorial Pacific Ocean. (Units: m s^{-1})

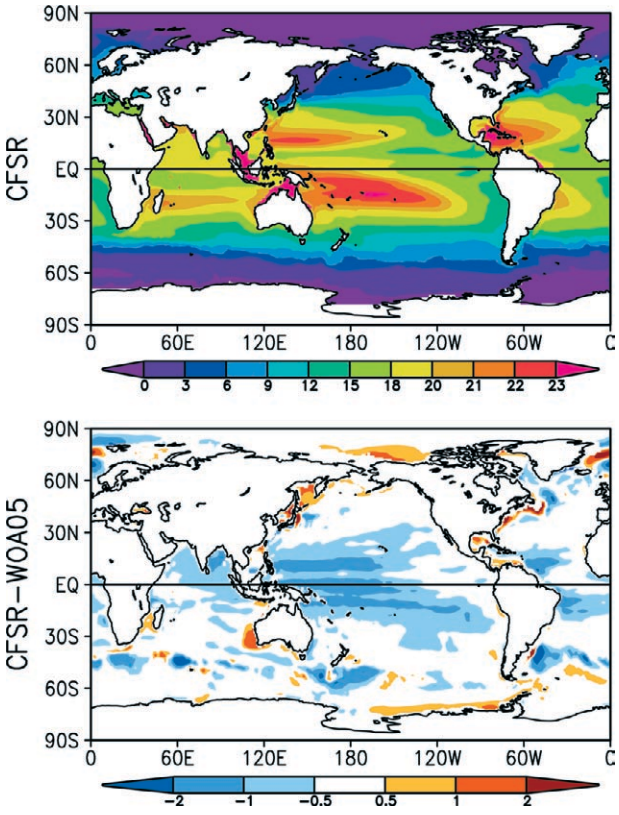

Fig. 31. The (top) vertically averaged temperature (from the surface to 300-m depth) for CFSR for 1979–2008, and (bottom) differences between CFSR and observations from *WOA05* (Locarnini et al. 2006). (Units: K)

also required of the benefits of coupling the atmosphere to the ocean and the sea ice; but, at first blush, this aspect appears to have worked very well in the CFSR. The SST–precipitation correlation has improved in the tropics. Some problems related to executing the project in many streams (as is common to speed up the process) still exist. All subcomponents with longer time scales (such as the deep soil, deep ocean, and the top of the atmosphere) have discontinuities where one stream ends and the next one begins; this in spite of a full 1-yr overlap between the streams.

Future developments include the following three projects:

1) CFSR was conducted mainly to create initial conditions for the coupled atmosphere–ocean–land–sea ice reforecasts of the CFS version 2 forecast model, over the period of 1982–present. This project is underway at NCEP. The design of these reforecasts is as follows: From every fifth day in the calendar, there will be four 9-month "seasonal" forecasts from 0000, 0600, 1200, and 1800 UTC. From every day, there will also be shorter predictions—one run to the first season (~123 days) at 0000 UTC and three runs to 45 days from 0600, 1200, and 1800 UTC. The emphasis on the shorter subseasonal predictions, for the MJO and week 3–6 forecasts, is to bridge the weather–climate gap and is the main reason for a high-resolution reanalysis to be conducted. There is consensus that, given a forecast model at a lower resolution (say T126), the skill of the forecasts benefits from the highest possible resolution of the initial state.
2) Given the pace of model and data assimilation development, we expect a new global reanalysis to be conducted at NCEP once every 7 years or so. However, there is serious thought being given to immediately conduct CFSRL: a "light" (with a reduced horizontal resolution of T126) version of the reanalysis that was just completed. It would be done in a single stream to overcome the discontinuities found in the CFSR for the deep ocean, deep soil, and top of the atmosphere. It is possible that the CFSRL will be finished in 1 year, in time for CPC to use it when they change their climate normals to the last 30-yr period from 1981 to 2010.
3) A final activity to be conducted when the reforecast project is complete is to apply the reanalysis system, as used here, to the historical period of 1948–78. The CFSR is the successor of R2, and, when extended back to 1948, will also be the successor of R1. It is possible this will be done in one-stream "light" mode.

ACKNOWLEDGMENTS. We acknowledge the U.S. Climate Program Office (reanalysis grant) for funding part of this project. We thank the staff of EMC, CPC, NCEP Central Operations (NCO), and NESDIS for their various contributions in completing this large endeavor in record time. We thank Kenneth E. Mitchell, now retired from NCEP/EMC, for his guidance in updating the land system used in the CFSR. We acknowledge the Geophysical Fluid Dynamical Laboratory (GFDL) for the ocean and sea ice model used in the CFSR. We thank George VandenBerghe, Carolyn Pasti, Doris Pan, Don Avart, Cameron Shelton, and others for their outstanding computer support. We acknowledge the efforts of Augustin Vintzileos (EMC) and Adrian Simmons (ECMWF) to give us the AMMA observations in time for use in the CFSR. The CFSR Science Advisory Board (chaired by Jeff Anderson) gave useful guidance at the start of the project. A very thorough and constructive internal review was provided by Glenn White (EMC). Reviews by Jeff Anderson, Michael Bosilovich, and Chet Ropelewski were very helpful. As with other reanalysis projects, this is the collective work of many

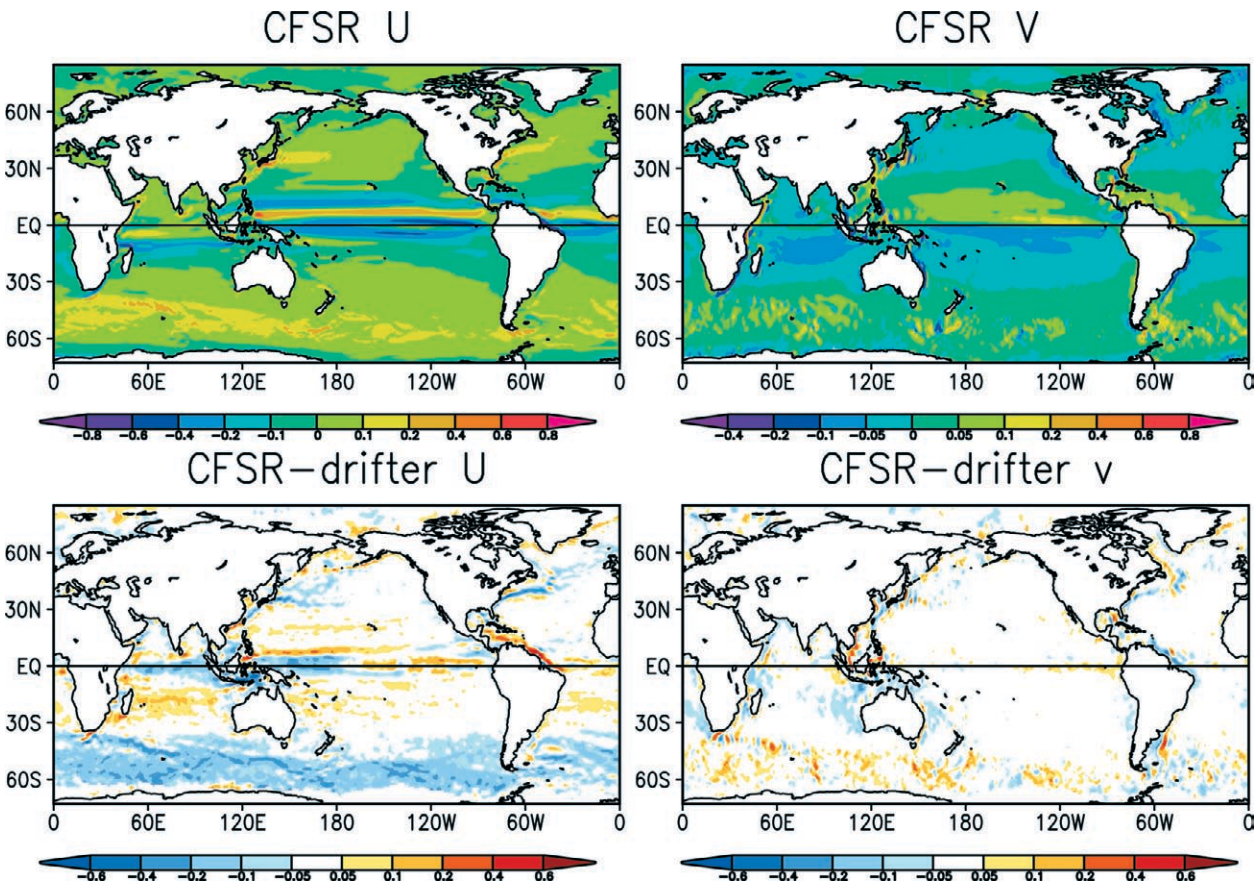

Fig. 32. (top left) Zonal and (top right) meridional surface velocities for CFSR and (bottom) differences between CFSR and drifters from the surface velocity program of TOGA. The drifter data, a pseudo-observation dataset obtained from inverting location data of the surface drifters, has the important asset of being global. (Units: m s^{-1})

people from many organizations. We express our gratitude to all of them.

APPENDIX: ACRONYMS

3DVAR	Three-dimensional variational data assimilation
4DVAR	Four-dimensional variational data assimilation
AC	Anomaly correlation
ACARS	Aircraft Communications Addressing and Reporting System
AER	Atmospheric and Environmental Research, Inc.
AIRS	Atmospheric Infrared Sounder
AM	Atmospheric model
AMI	Active Microwave Instrument
AMIP	Atmospheric Model Intercomparison Project
AMMA	African Monsoon Multidisciplinary Analysis
AMSR-E	Advanced Scanning Radiometer-EOS
AMSU	Advanced Microwave Sounding Unit
ATOVS	Advanced TIROS Operational Vertical Sounder
AVHRR	Advanced Very High Resolution Radiometer
BOM	Australian Bureau of Meteorology
BUFR	Binary Universal Form for the representation of meteorological data
CAMS	Climate Anomaly Monitoring System
CDAS	Climate Data Assimilation System
CFS	Climate Forecast System
CFSR	Climate Forecast System Reanalysis
CFSR-Lite	CFSR "light"
CFSRR	Climate Forecast System Reanalysis and Reforecasts
CHAMP	Challenging Mini-satellite Payload
CIS	Canadian Ice Service
CLASS	Comprehensive Large-Array Data Stewardship System

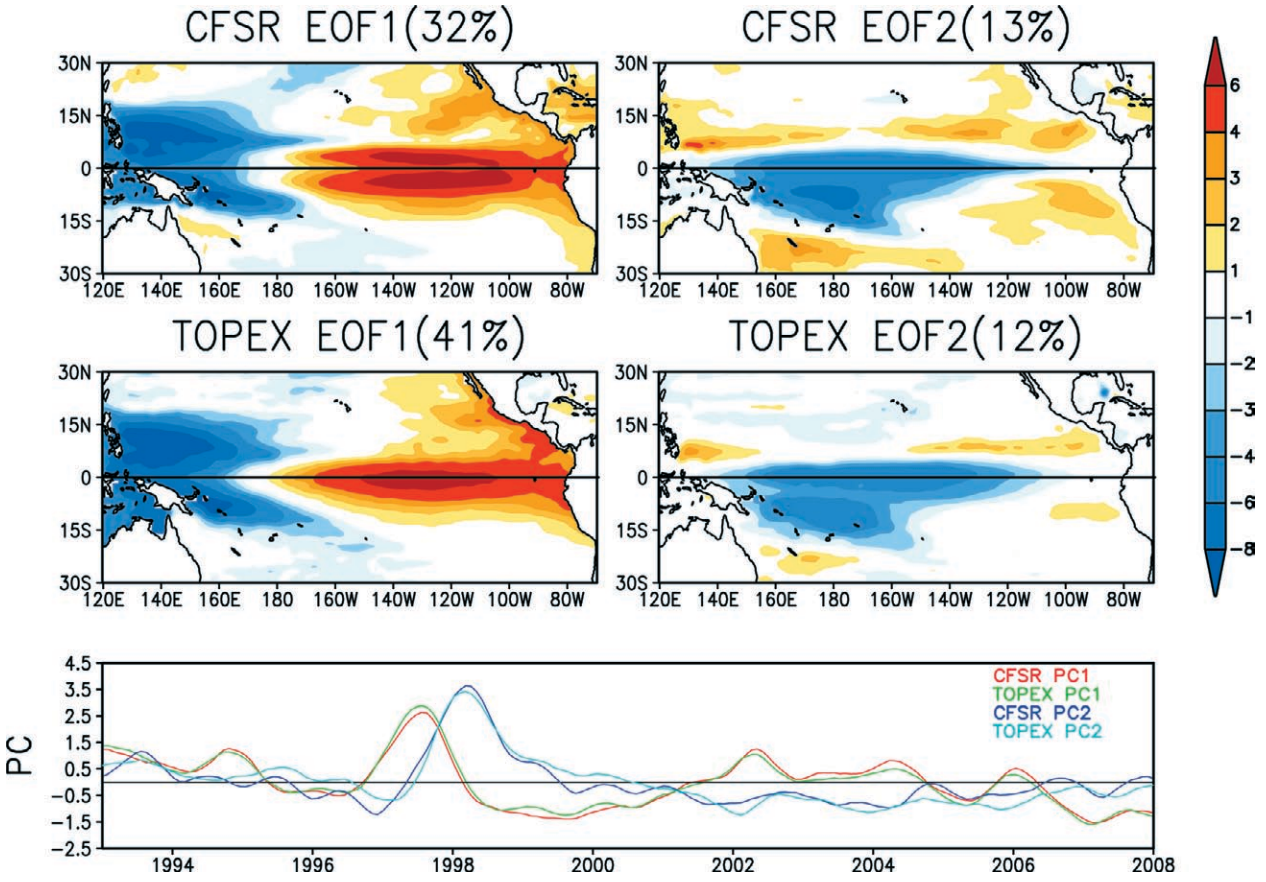

Fig. 33. The first two EOFs of the SSH variability for the (top) CFSR and for (middle) TOPEX satellite altimeter data for the period of 1993–2008. (bottom) The time series amplitude factors are plotted.

CMAP	CPC Merged Analysis of Precipitation
COLA	Center for Ocean–Land–Atmosphere Studies
COSMIC	Constellation Observing System for Meteorology Ionosphere and Climate
CPC	Climate Prediction Center
CpT	Enthalpy
CRTM	Community Radiative Transfer Model
DMSP	Defense Meteorological Satellite Program
ECMWF	European Centre for Medium-Range Weather Forecasts
EMC	Environmental Modeling Center
EnKF	Ensemble Kalman filter
EOS	Earth Observing System
ERA-15	15-yr ECMWF Re-Analysis
ERA-40	40-yr ECMWF Re-Analysis
ERA-Interim	ECMWF's Interim Re-Analysis
ERS	European Remote Sensing Satellite
ESA	European Space Agency
ESMF	Earth System Modeling Framework
ESRL	Earth System Research Laboratory
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
FGGE	First Global Atmospheric Research Program (GARP) Global Experiment
FOTO	First-order time interpolation to the observation
GDAS	Global Data Assimilation System
GFDL	Geophysical Fluid Dynamics Laboratory
GFS	Global Forecast System
GHCN	Global Historical Climate Network
GLDAS	Global Land Data Assimilation System
GMAO	Global Modeling and Assimilation Office
GMS	Geosynchronous Meteorological Satellite

GODAS	Global Ocean Data Assimilation System
GOES	Geostationary Operational Environmental Satellite
GRIB	Gridded binary
GSFC	Goddard Space Flight Center
GSFC Ice	GSFC Ice Cloud and Land Elevation Satellite
GSI	Gridded statistical interpolation
GTS	Global Telecommunication System
GTSPP	Global Temperature–Salinity Profile Project
HadISST	Hadley Centre Global Sea Ice and Sea Surface Temperature
HIRS	High-Resolution Infrared Sounder Unit
IASI	Infrared Atmospheric Sounding Interferometer
IMS	Interactive Multisensor Snow and Ice Mapping System
JCSDA	Joint Center for Satellite Data Assimilation
JMA	Japanese Meteorological Agency
JRA-25	Japanese 25-Year Global Reanalysis
LCL	Lifting condensation level
LDAS	Land Data Assimilation System
LEO	Low Earth Orbiting Satellite
LIS	Land Information System
LSM	Land surface model
LW	Longwave
MERRA	Modern Era Retrospective-Analysis for Research and Applications
METAR	Meteorological Aviation Report
MetOp	Meteorological Operation
MHS	Microwave humidity sounder
MJO	Madden–Julian oscillation
MLS	Microwave limb sounding
MODIS	Moderate Resolution Imaging Spectroradiometer
MOM	GFDL Modular Ocean Model
MPMD	Multiple Program Multiple Data
MSU	Microwave Sounder Unit
NARR	North American Regional Reanalysis
NASA	National Aeronautical and Space Administration
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NCEP	National Centers for Environmental Prediction
NESDIS	National Environmental Satellite, Data and Information Service
NIST	National Institute of Standards and Technology
NMC	National Meteorological Center
NOAA	National Oceanic and Atmospheric Administration
NODC	National Oceanographic Data Center
NOMADS	NOAA Operational Model Archive and Distribution System
NRL	Naval Research Laboratory
NSDA	National Space Development Agency of Japan
NWP	Numerical weather prediction
NWS	U.S. National Weather Service
OI	Optimum interpolation
OIQCBUFR	Optimum Interpolation based Quality Control of observations in BUFR format
OSU	Oregon State University
PAOBS	Paid Observation, data from a manual analysis
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PIRATA	Prediction and Research Moored Array in the Tropical Atlantic
QBO	Quasi-biennial oscillation
QC	Quality control
QuikSCAT	Quick Scatterometer
R1	NCEP/NCAR Reanalysis Global Reanalysis 1
R2	NCEP–DOE Reanalysis Global Reanalysis 2
RAMA	Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction
RO	Radio occultation
RRTM	Rapid radiative transfer model
RRTMG-LW	RRTM LW modified GCM version 2.3
RRTMG-SW	RRTM SW modified GCM version 2.3
RTG	Real-time Global
RTOVS	Revised TIROS Operational Vertical Sounder
SAO	Semiannual oscillation
SATOB	Satellite observations
SBUV	Solar Backscatter Ultraviolet Radiometer
SLP	Sea level pressure
SNODEP	Snow depth
SSH	Sea surface height
SSI	Spectral Statistical Interpolation
SSM/I	Special Sensor Microwave Imager

SSS	Sea surface salinity
SST	Sea surface temperature
SSU	Stratospheric Sounder Unit
SW	Shortwave
TAO	Tropical Atmosphere Ocean
TIROS	Television and Infrared Observation Satellite
TOPEX	Ocean Topography Experiment
TOVS	TIROS Operational Vertical Sounder
TRITON	Triangle Trans-Ocean Buoy Network
USAF	U.S. Air Force
VarQC	Variational quality control
WMO	World Meteorological Organization
WOA05	2005 World Ocean Atlas
XBT	Expendable bathythermograph

REFERENCES

Accadia, C., S. Mariani, M. Casaioli, and A. Lavagnini, 2003: Sensitivity of precipitation forecast skill scores to bilinear interpolation and a simple nearest-neighbor average method on high-resolution verification grids. *Wea. Forecasting,* **18,** 918–932.

Akmaev, R. A., and H.-M. H. Juang, 2008: Using enthalpy as a prognostic variable inatmospheric modeling with variable composition. *Quart. J. Roy. Meteor. Soc*., **134**, 2193–2197.

Alpert, J. C., 2004: Subgrid-scale mountain blocking at NCEP. *Proc. 20th Conf. on Weather and Forecasting,* Seattle, WA, Amer. Meteor. Soc., P2.4. [Available online at http://ams.confex.com/ams/pdfpapers/71011.pdf.]

——, M. Kanamitsu, P. M. Caplan, J. G. Sela, G. H. White, and E. Kalnay, 1988: Mountain induced gravity wave drag parameterization in the NMC medium-rangeforecast model. *Proc. Eighth Conf. on Numerical Weather Prediction*, Baltimore, MD, Amer. Meteor. Soc., 726–733.

——, S.-Y. Hong, and Y.-J. Kim, 1996: Sensitivity of cyclogenesis to lower tropospheric enhancement of gravity wave drag using the Environmental Modeling Center medium-range model. *Proc. 11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc., 322–323.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc*., **90**, 1283–1296.

Andersson, E., and H. Järvinen, 1999: Variational quality control. *Quart. J. Roy. Meteor. Soc*., **125**, 697–722.

Argo Science Team, 2001: The global array of profiling floats. *Observing the Ocean in the 21st Century*, C. J. Koblinsky and N. R. Smith, Eds., Australian Bureau of Meteorology, 248–258.

Assel, R. A., D. C. Norton, and K. C. Cronk, 2002: A Great Lakes ice cover digital data set for winters 1973-2000. NOAA Tech. Memo. GLERL-121, 46 pp.

Balaji,V., 2007: The FMS coupler architecture. *Proc. Ocean Model Coupling Workshop*, Princeton, NJ, 1–26.

Behringer, D. W., 2007: The Global Ocean Data Assimilation System (GODAS) at NCEP. Preprints, *11th Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface,* San Antonio, TX, Amer. Meteor. Soc., 3.3. [Available online at http://ams.confex.com/ams/pdfpapers/119541.pdf.]

——, and Y. Xue, 2004: Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. Preprints, *Eighth Symp. on Integrated Observing and Assimilation System for Atmosphere, Ocean, and Land Surface,* Seattle, WA, Amer. Meteor. Soc., 2.3. [Available online at http://ams.confex.com/ams/pdfpapers/70720.pdf.]

——, M. Ji, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. *Mon. Wea. Rev,* **126**, 1013–1021.

Bender, M. A., I. Ginis, R. Tuleya, B. Thomas, and T. Marchok, 2007: The operational GFDL coupled hurricane–ocean prediction system and summary of its performance. *Mon. Wea. Rev.*, **135**, 3965–3989.

Bengtsson, L., and Coauthors, 2007: The need for a dynamical climate reanalysis. *Bull. Amer. Meteor. Soc.,* **88,** 495–501.

Bosilovich, M., 2008: NASA's modern era retrospective-analysis for research and applications: Integrating Earth observations. *Earthzine,* 26 September 2008. [Available online at www.earthzine.org/2008/09/26/nasas-modern-era-retrospective-analysis/.]

Bourlès, B., and Coauthors, 2008: The PIRATA Program: History, accomplishments, and future directions. *Bull. Amer. Meteor. Soc.,* **89**, 1111–1125.

Cai, M., C. S. Shin, H. M. van den Dool, W. Wang, S. Saha, and A. Kumar, 2009: The role of long-term trends in seasonal predictions: Implication of global warming in the NCEP CFS. *Wea. Forecasting*, **24**, 965–973.

Cavalieri, D. J., 1994: Sea ice algorithm in NASA Sea Ice Validation Program for the Defense Meteorological Satellite Program Special Sensor Microwave Imager: Final Report. NASA Tech. Memo. 104559, 126 pp.

——, C. Parkinson, P. Gloersen, and H. J. Zwally, 2007: Sea ice concentrations from *Nimbus-7* SMMR and DMSP SSM/I passive microwave data, 1978-1996.

National Snow and Ice Data Center, Boulder, CO, digital media. [Available online at http://nsidc.org/data/nsidc-0051.html.]

Chelliah, M., and C. F. Ropelewski, 2000: Reanalyses-based tropospheric temperature estimates: Uncertainties in the context of global climate change detection. *J. Climate*, **13**, 3187–3205.

Chou, M. D. M. J. Suarez, C. H. Ho, M. M. H. Yan, and K. T. Lee, 1998: Parameterizations of cloud overlapping and shortwave single scattering properties for use in general circulation and cloud ensemble models. *J. Climate*, **11**, 202–214.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown, 2005: Atmospheric radiative transfer modeling: A summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transfer*, **91**, 233–244.

Collins, N., and Coauthors, 2005: Design and implementation of components of earth system modeling framework. *Int. J. High Perform. Comput. Appl.*, **19**, 355–356.

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

Compo, G. P., J. S. Whitaker, and P. D. Sardeshmukh, 2006: Feasibility of a 100-year reanalysis using only surface pressure data. *Bull. Amer. Meteor. Soc.,* **87,** 175–190.

Conkright, M. E., and Coauthors, 1999: World Ocean Database 1998, documentation and quality control version 2.0. National Oceanographic Data Center Internal Rep. 14, 127 pp.

Cucurull, L., 2010: Improvement in the use of an operational constellation of GPS radio-occultation receivers in weather forecasting. *Wea. Forecasting,* **25,** 749–767.

——, and J. C. Derber, 2008: Operational implementation of COSMIC observations into the NCEP's Global Data Assimilation System. *Wea. Forecasting,* **23**, 702–711.

Dai, A., and J. Wang, 1999: Diurnal and semidiurnal tides in global surface pressure fields. *J. Atmos. Sci.*, **56**, 3874–3891.

Davis, G., 2007: History of the NOAA satellite program. *J. Appl. Remote Sens.*, **1,** 012504, doi:10.1117/1.2642347.

Deardorff, J. W., 1980: Cloud top entrainment instability. *J. Atmos. Sci.*, **37**, 131–147.

Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347.

——, and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299.

——, D. F. Parrish, and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting*, **6**, 538–547.

de Viron, O., G. Schwarzbaum, F. Lott, and V. Dehant, 2005: Diurnal and subdiurnal effects of the atmosphere on the Earth rotation and geocenter motion. *J. Geophys. Res.*, **110**, B11404, doi:10.1029/2005JB003761.

Dworak, R., and J. Key, 2009: Twenty years of polar winds from AVHRR: Validation and comparison with ERA-40. *J. Atmos. Oceanic Technol.*, **48,** 24–40.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarplay, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.,* **108**, 8851, doi:10.1029/2002JD003296.

English, S. J., R. J. Renshaw, P. C. Dibben, A. J. Smith, P. J.Rayer, C. Poulsen, F. W. Saunders, and J. E. Eyre, 2000: A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Quart. J. Roy. Meteor. Soc.,* **126**, 2911–2931.

Fan, Y., and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948–present. *J. Geophys. Res.,* **113,** D01103, doi:10.1029/2007JD008470.

Fels, S., and M. D. Schwarzkopf, 1975: The simplified exchange approximation: A new method for radiative transfer calculations. *J. Atmos. Sci.,* **32,** 1475–1488.

Fiorino, M., 2002: Analysis and forecasts of tropical cyclones in the ECMWF 40-year reanalysis (ERA-40). ERA-40 Project Report Series, Vol. 3, ECMWF, 443 pp.

——, 2004: A multi-decadal daily sea surface temperature and sea ice concentration data set for the ECMWF ERA-40 reanalysis. ERA-40 Project Report Series, Vol. 12, ECMWF, 16 pp. [Available online at www.ecmwf.int/publications/library/ecpublications/_pdf/ERA40_PRS12.pdf.]

Flynn, L., and Coauthors, 2009: Measurements and products from the Solar Backscatter Ultraviolet (SBUV/2) and Ozone Mapping and Profiler Suite (OMPS) instruments. *Int. J. Remote Sens.*, **30**, 4259–4272.

Fu, X., and B. Wang, 2004: Differences of boreal summer intraseasonal oscillations simulated in an atmosphere–ocean coupled model and an atmosphere-only model. *J. Climate*, **17**, 1263–1271.

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

Gaspari, G., S. E. Cohn, J. Guo, and S. Pawson, 2006: Construction and application of covariance func-

tions with variable length fields. *Quart. J. Roy. Meteor. Soc.*, **132**, 1815–1838.

Gemmill, W. H., and V. M. Krasnopolsky, 1999: The use of SSM/I data in operational marine analysis. *Wea. Forecasting*, **14**, 789–800.

——, B. Katz, and X. Li, 2007: Daily real-time global sea surface temperature: A high-resolution analysis. RTG SST RH, MMAB Tech. Note 260, 39 pp.

Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155.

Gibson, J. K., P. Kallberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ERA description. ECMWF Re-Analysis Project Report Series, Vol. 1, 89 pp.

Griffies, S. M., and R. W. Hallberg, 2000: Biharmonic friction with a Smagorinsky viscosity for use in large-scale eddy-permitting ocean models. *Mon. Wea. Rev.*, **128**, 2935–2946.

——, A. Gnanadesikan, R. C. Pacanowski, V. Larichev, J. K. Dukowicz, and R. D. Smith, 1998: Isoneutral diffusion in a *z*-coordinate ocean model. *J. Phys. Oceanogr.*, **28**, 805–830.

——, M. J. Harrison, R. C. Pacanowski, and A. Rosati, 2004: Technical guide to MOM4. GFDL Ocean Group Technical Report No. 5, 337 pp. [Available online at www.gfdl.noaa.gov/~fms.]

Grumbine, R. W., 1996: Automated passive microwave sea ice concentration analysis at NCEP. NCEP OMB Tech. Note 120, 13 pp.

——, 2009: A posteriori filtering of sea ice concentration analyses. NCEP MMAB Tech. Note 282, 7 pp.

Hamill, T. M., J. S. Whitaker, and S. L. Mullen, 2006: Reforecasts: An important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc.*, **87**, 33–46.

Han, Y., P. van Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006: JCSDA Community Radiative Transfer Model (CRTM), version 1. NOAA Tech. Rep. NESDIS 122, 33 pp.

Haurwitz, B. and D. Cowley, 1973: The diurnal and semidiurnal barometric oscillations, global distribution and annual variation. *Pure Appl. Geophys.,* **102,** 193–222.

Helfrich, S. R., D. McNamara, B. H. Ramsay, T. Baldwin, and T. Kasheta, 2007: Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrol. Processes,* **21,** 1576–1586, doi:10.1002/hyp.6720.

Hong, S.-Y. and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.,* **124**, 2322–2339.

——, and ——, 1998: Convective trigger function for a mass-flux cumulus parameterization scheme. *Mon. Wea. Rev.*, **126**, 2599–2620.

Hou, Y. -T., K. A. Campana and S.–K. Yang, 1996: Shortwave radiation calculations in the NCEP's global model. *Int. Radiation Symp., IRS-96*, Fairbanks, AK, Deepak Publishing, 317–319.

——, S. Moorthi, and K. Campana, 2002: Parameterization of solar radiation transfer in the NCEP models. NCEP Office Note 441, 46 pp. [Available online at www.emc.noaa.gov/officenotes/FullTOC.html#2000.]

Iacono, M. J., E. J. Mlawer, S. A. Clough, and J.-J. Morcrette, 2000: Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR Community Climate Model, CCM3. *J. Geophys. Res.*, **105**, 14 873–14 890.

Jenne, R., and J. Woollen, 1994: The reanalysis database. *Extended Abstracts*, *10th Conf. on Numerical Weather Prediction*, Portland, OR, Amer. Meteor. Soc.

Ji, M., A. Leetmaa, and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies. *Mon. Wea. Rev.,* **123***,* 460–481.

Juang, H.-M., 2005: Discrete generalized hybrid vertical coordinates by a mass, energy and angular momentum conserving vertical finite-differencing scheme. NCEP Office Note 445, 33 pp. [Available online at www.emc.ncep.noaa.gov/officenotes/FullTOC.html#2000.]

Kadi, M., 2009: *The AMMA Observing Network Contribution to GCOS, GCOS Steering Committee: 17th Session, Paris, France, 27-30 October 2009.* 4 pp.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S. K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83,** 1631–1643.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50–Year Reanalysis: Monthly means CD–ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247–267.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009: Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.,* **137,** 1046–1060.

Kobayashi, S., M. Matricardi , D. Dee, and S. Uppala 2009: Toward a consistent reanalysis of the upper stratosphere based on radiance measurements from SSU and AMSU-A. *Quart. J. Roy. Meteor. Soc.*, **135**, 2086–2099.

Kopp, T. J., and R. B. Kiess, 1996: The Air Force Global Weather Central snow analysis model. Preprints, *15th Conf. on Weather Analysis and Forecasting,* Norfolk, VA, Amer. Meteor. Soc., 220–222.

Krasnopolsky, V. M., L. C. Breaker, and W. H. Gemmill, 1995: A neural network as a nonlinear transfer function model for retrieving surface wind speeds from the Special Sensor Microwave/Imager. *J. Geophys. Res.*, **100**, C6, 11 033–11 045.

Large, W. G., J. C. McWilliams and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, **32**, 363–403.

Lindzen, R. S., and J. R. Holton, 1968: A theory of quasi-biennial oscillation. *J. Atmos. Sci.,* **26,** 1095–1107.

Liu, Q., and F. Weng, 2006: Advanced doubling–adding method for radiative transfer in planetary atmospheres. *J. Atmos. Sci.*, **63**, 3459–3465.

——, and ——, 2009: Recent stratospheric temperature observed from satellite measurements. *SOLA,* **5,** 53–56, doi:10.2151/sola.2009-014.

——, T. Marchok, H. Pan, M. Bender, and S. Lord, 1999: Improvements in hurricane initialization and forecasting at NCEP with global and regional (GFDL) models. NOAA Tech. Proc. Bull. 472, National Weather Service, Office of Meteorology, Silver Spring, MD, 1–7. [Available online at www.nws.noaa.gov/om/tpb/472.htm.]

Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, and H. E. Garcia, 2006: *Temperature*. Vol. 1, *World Ocean Atlas 2005*, NOAA Atlas NESDIS 61, 182 pp.

Lott, F., and M. J. Miller, 1997: A new subgrid-scale orographic drag parameterization: Its performance and testing. *Quart. J. Roy. Meteor. Soc.*, **123**, 101–127.

Lumpkin, R., and M. Pazos, 2006: Measuring surface currents with Surface Velocity Program drifters: The instrument, its data, and some recent results. *Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics (LAPCOD),* A. Griffa et al., Eds., 39–67.

Markus, T., and D. J. Cavalieri, 2000: An enhancement of the NASA Team sea ice algorithm. *IEEE Trans. Geosci. Remote Sens.*, **38**, 1387–1398.

——, and ——, 2009: The AMSR-E NT2 sea ice concentration algorithm: Its basis and implementation. *J. Remote Sens. Soc. Japan*, **29**, 216–225.

McCormack, J. P., S. D., Eckermann, D. E. Siskind, and T. McGee, 2006: CHEM2D-OPP: A new linearized gas phase photochemistry parameterization for high altitude NWP and climate models. *Atmos. Chem. Phys.*, **6**, 4943–4972.

McNally, A. P., J. C. Derber, W.-S. Wu, and B. B. Katz, 2000: The use of TOVS level-1B radiances in the NCEP SSI analysis system. *Quart. J. Roy. Meteor. Soc.*, **126**, 689–724.

McPhaden, M. J., and Coauthors, 1998: The Tropical Ocean-Global Atmosphere (TOGA) observing system: A decade of progress. *J. Geophys. Res.,* **10**3, 14 169–14 240.

——, and Coauthors, 2009: RAMA: The Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction. *Bull. Amer. Meteor. Soc.,* **90**, 459–480.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360.

Mlawer E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102** (D14), 16 663–16 682.

Mo, T., M. D. Goldberg, D. S. Crosby, and Z. Cheng, 2001: Recalibration of the NOAA microwave sounding unit. *J. Geophys. Res.*, **106**, 10 145–10 150.

Moorthi, S., H. L. Pan, and P. Caplan, 2001: Changes to the 2001 NCEP operationalMRF/AVN global analysis/forecast system. NWS Technical Procedures Bulletin, Vol. 484, 14 pp. [Available online at www.nws.noaa.gov/om/tpb/484.htm.]

——, R. Sun, H. Xiao, and C. R. Mechoso, 2010: Low-cloud simulation in the Southeast Pacific in the NCEP GFS: Role of vertical mixing and shallow convection. NCEP Office Note 463, 28 pp. [Available online at www.emc.ncep.noaa.gov/officenotes/FullTOC.html#2000.]

Murray, R. J., 1996: Explicit generation of orthogonal grids for ocean models. *J. Comput. Phys.*, **126**, 251–273.

National Research Council, 2007: *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts.* National Academies of Science, 76 pp.

Neumann, C. J., 1999: The HURISK model: An adaptation for the Southern Hemisphere (A user's manual). SAIC Rep., Contract N00014-96-C-6015, 31 pp.

Onogi, K., and Coauthors, 2007: The JRA-25 Reanalysis. *J. Meteor. Soc. Japan*, **85**, 369–432.

Pan, H. L., and L. Mahrt, 1987: Interaction between soil hydrology and boundary layer development. *Bound.-Layer Meteor.*, **38**, 185–202.

——, and W.-S. Wu, 1995: Implementing a mass flux convective parameterization package for the NMC medium range forecast model. NMC Office Note 409, 40 pp. [Available online at www.emc.ncep.noaa.gov/officenotes/FullTOC.html#1990.]

Paolino, D., Q. Yang, B. Doty, J. Kinter, J. Shukla, and D. M. Straus, 1995: A pilot reanalysis project at COLA. *Bull. Amer. Meteor. Soc.,* **76**, 697–710.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Pegion, K., and B. Kirtman 2008: The impact of air–sea interactions on the simulation of tropical intraseasonal variability. *J. Climate,* **24**, 6616–6635.

Peters-Lidard, C. D., and Coauthors, 2007: High-performance Earth system modeling with NASA/GSFC's Land Information System. *Innovations Syst. Software Eng.*, **3,** 157–165. doi:10.1007/s11334-007-0028-x.

Phillips, N. A., 1957: A coordinate system with some special advantages for numerical forecasting *J. Atmos. Sci.*, **14**, 184–185.

Ponte, R. M., and R. D. Ray, 2002: Atmospheric pressure corrections in geodesy and oceanography: A strategy for handling air tides. *Geophys. Res. Lett.*, **29**, 2153, doi:10.1029/2002GL016340.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524–1535.

——, ——, ——, and ——, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548.

Rančić, M., J. C. Derber, D. Parrish, R. Treadon, and D. T. Kleist, 2008: The development of the first-order time extrapolation to the observation (FOTO) method and its application in the NCEP global data assimilation system. *Proc. 12th Symp. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS),* New Orleans, LA, Amer. Meteor. Soc., J6.1. [Available online at http://ams.confex.com/ams/pdfpapers/131816.pdf.]

Randall, D. A., 1980: Conditional instability of the first kind upside-down. *J. Atmos. Sci*, **37**, 125–130.

Ray, R. D., and R. M. Ponte, 2003: Barometric tides from ECMWF operational analyses. *Ann. Geophys.*, **21**, 1897–1910.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution blended analyses for sea surface temperature. *J. Climate,* **20**, 5473–5496.

Rocken, C., R., and Coauthors, 1997: Analysis and validation of GPS/MET data in the neutral atmosphere. *J. Geophys. Res.*, **102** (D25), 29 849–29 866.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517.

Schubert, S. D., W. Min, L. Takacs, and J. Joiner, 1997: Reanalysis of historical observations and its role in the development of the Goddard EOS Climate Data Assimilation System. *Adv. Space Res.* **19**, 491–501.

——, and Coauthors, 2008: Assimilating earth system observations at NASA: MERRA and beyond. *Third WCRP Int. Conf. on Reanalysis*, Tokyo, Japan, WCRP, 1–6. [Available online at http://wcrp.ipsl.jussieu.fr/Workshops/Reanalysis2008/Documents/V1-104_ea.pdf.]

Schwarzkopf, M. D., and S. Fels, 1991: The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rates and fluxes. *J. Geophys Res.,* **96** (D5), 9075–9096.

Seaman, R., and Hart, T., 2003: The history of PAOBs in the Australian Bureau of Meteorology. *Aust. Meteor. Mag.,* **52**, 241–250.

Smith, W. L., H. M. Woolf, C. M. Hayden, D. Q. Wark, and L. M. McMillin, 1979: The TIROS-N Operational Vertical Sounder. *Bull. Amer. Meteor. Soc.*, **60**, 1177–1187.

——, F. W. Nagle, C. M. Hayden, and H. M. Woolf, 1981: Vertical mass and moisture structure from TIROS-N. *Bull. Amer. Meteor. Soc.*, **62**, 388–393.

Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and cloud studies with mesoscale numerical weather prediction model. *Mon. Wea. Rev.,* **117**, 1641–1757.

Thiébaux, J., E. Rogers, W. Wang, and B. Katz, 2003: A new high-resolution blended real-time global sea surface temperature analysis. *Bull. Amer. Meteor. Soc.*, **84**, 645–656.

Tiedtke, M., 1983: The sensitivity of the time-mean large-scale flow to cumulus convection in the ECMWF model. *ECMWF Workshop on Convection in Large-Scale Models*, Reading, UK, ECMWF, 297–316.

Trenberth, K. E., and L. Smith, 2005: The mass of the atmosphere: A constraint on global analyses. *J. Climate*, **18**, 864–875.

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.

Van den Dool, H. M., S. Saha, J. Schemm, and J. Huang, 1997: A temporal interpolation method to obtain hourly atmospheric surface pressure tides in reanalysis 1979–1995. *J. Geophys. Res.*, **102** (D18), 22 013–22 024.

Velicogna, I., J. Wahr, and H. Van den Dool, 2001: Can surface pressure be used to remove atmospheric contributions from GRACE data with sufficient accuracy to recover hydrological signals? *J. Geophys. Res.*, **106** (B8), 16 415–16 434.

Weaver, A. T., and P. Courtier, 2001: Correlation modeling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc*., **127**, 1815–1846.

Wickert, J., and Coauthors, 2001: Atmosphere sounding by GPS radio occultation: First results from CHAMP. *J. Geophys. Res*., **105**, 7257–7273.

Woollen, J. S., 1991: New NMC operational OI quality control. Preprints, *Ninth Conf. on Numerical Weather Prediction*, Portland, OR, Amer. Meteor. Soc., 24–27.

——, E. Kalnay, L. Gandin, W. Collins, S. Saha, R. Kistler, M. Kanamitsu, and M. Chelliah, 1994: Quality control in the reanalysis system. Preprints, *Ninth Conf. on Numerical Weather Prediction*, Portland, OR, Amer. Meteor. Soc., 13–14.

Woolnough, S. J., J. M. Slingo, and B. J. Hoskins, 2000: The relationship between convection and sea surface temperature on intraseasonal time scales. *J. Climate*, **13**, 2086–2104.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916.

Wu, X., K. S. Moorthi, K. Okomoto, and H. L. Pan, 2005: Sea ice impacts on GFS forecasts at high latitudes. Preprints, *Eighth Conf. on Polar Meteorology and Oceanography,* San Diego, CA, Amer. Meteor. Soc., 7.4. [Available online at http://ams.confex.com/ams/pdfpapers/84292.pdf.]

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558.

——, M. Chen, A. Yatagai, T. Hayasaka, Y. Fukushima, and S. Yang, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607–626.

Xu, K. M., and D. A. Randall, 1996: A semiempirical cloudiness parameterization for use in climate models. *J. Atmos. Sci*., **53,** 3084–3102.

Yu, T.-W., M. D. Iredell, and D. Keyser 1997: Global data assimilation and forecast experiments using SSM/I wind speed data derived from a neural network algorithm. *Wea. Forecasting*, **12**, 859–865.

Zhao, Q. Y., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP models. *Mon. Wea. Rev.,* **125**, 1931–1953.

Zou, C.-Z., M. Goldberg, Z. Cheng, N. Grody, J. Sullivan, C. Cao, and D. Tarpley, 2006: Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *J. Geophys. Res.,* **111,** D19114: doi:10.1029/2005JD006798.

——, M. Gao, M. Goldberg, 2009: Error structure and atmospheric temperature trend in observations from the microwave sounding unit. *J. Climate,* **22**, 1661–1681.