Progress Report on the Global Data Processing System 2003

THE NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION
NATIONAL WEATHER SERVICE, U.SA.

1. Highlights For Calendar Year 2003

The most significant event of 2003 from an operational standpoint was the transition of NCEP operations to the new NCEP Central Computer System (CCS).

The new computer system uses a highly parallel architecture with 1536 processors. During the first three years of the contract, the supercomputer will, on average, provide 4.9 times the computational power of the current system. It will undergo incremental upgrades reaching 48 times the computational power of the current system by October 2009.

1.1 The Central Computer System

The initial phase of the system, which was installed in the IBM Gaithersburg, Maryland computer facility in July 2002, is comprised of 44 Regatta H servers, each with 32 1.3 GHz CPUs. An additional 128 processors were added in September 2003. The CCS is configured as two symmetric systems to increase overall system reliability. All operational codes were converted to the CCS by November 2002 and parallel operations began in December 2002. The CCS became operational on 19 May 2003.

1.2 Winter Storm Reconnaissance Program

The goal of the Winter Storm Reconnaissance (WSR) program is to enhance the meteorological observational network prior to significant winter events. The innovative aspect of this program is that observations are adaptive in that they are collected only in areas presumed to most influence forecasts for these winter events. General forecast material and ensemble-based objective guidance are used to identify forecast weather events associated with large uncertainty. For flight planning purposes, the route selection and dropsonde areas must be selected approximately 30 hours before flight time. An ensemble application is used to evaluate all pre-designed flight tracks to assist in the selection of the site or sites expected to reduce forecast error variance the most. In 2001, of the 63 targeted winter storm events, 70% of the cases showed a clear improvement in the quality of 24-96 hour lead time forecasts, measured in terms of error reduction in surface pressure, tropospheric wind, and precipitation forecasts. Based on these and other earlier positive results, the program was raised to operational level. In the winters of 2002 and 2003, adaptive WSR observations improved the targeted surface pressure forecasts in 60-65% of the cases. In 2004, the WSR program covers the mid-January - mid-March time period.and the next WSR program will be conducted from mid-January through mid-March 2003.

1.3 Major changes to the NCEP Production Suite

A timeline of the most significant additions and enhancements to the NCEP production suite is provided below:

March 11, 2003: Assimilation of NOAA 17 1B radiances, restored assimilation of NOAA-16 AMSU-A radiances, and assimilation of QuikSCAT winds superobbed at 0.5 degrees were added to the Global Data Assimilation System (GDAS).

April 29, 2003: The Geophysical Fluid Dynamics Laboratory (GDFL) Hurricane model was upgraded. The model is now run from the operational T254, 64 level Global Forecast System (GFS) analysis. Upgrades include extending ocean coupling over entire ocean domain, and using a more realistic Gulf Stream assimilation, increasing the number of vertical levels from 18 to 42, implementing the Simplified Arakawa-Schubert Cumulus Parameterization (SAS) scheme and GFS Boundary Layer scheme, including a more accurate pressure gradient computation; improving mass initialization, which reduces noise over mountains, improving the interpolation of winds over mountains, and improving the (GFS) vortex removal algorithm in the model initialization. During the 2003 hurricance season, the upgraded GFDL model had its lowest track errors ever achieved in the Atlantic and Eastern Pacific. The average GFDL track error in the Atlantic was 15% lower than in 2002 and 30% lower than in 2001.

May 12, 2003: The forecast range of NCEP non-hurricane ocean wave models was extended from 126 to 168-hours, and the number of cycles was increased from 2 per day to 4 per day.

May 19, 2003: The 8-km resolution Fire Weather / IMET support run using the Non-hydrostatic Mesoscale Model (NMM) in a relocatable window nested within the Eta became operational. The NMM 48-hour forecast runs four times per day over a small domain selected by the Boise Fire Weather Center, NWS Western Region (WR) and NCEP's Storm Prediction Center (SPC). Boise and WR dispatch IMETs to wild fires and WR picks up NMM output grids, prepares FX-NET graphics imagery and transmits them via satellite communications to the IMET laptops used in the field. SPC forecasters also use the NMM guidance for their 1-2 day Fire Weather Outlook.

May 27, 2003: The 3DVAR data assimilation replaced the Optimum Interpolation system in the Rapid Update Cycle (RUC) weather forecast and data assimilation system.

July 8, 2003: The Eta Model "Spring Bundle" was implemented. This bundle consists of upgraded gridscale microphysics, improved cloud-radiation interaction, direct analysis of WSR-88D radial velocity to the 3DVAR analysis, added NOAA-17 data and reduced thinning of satellite radiances in 3DVAR analysis, added GOES cloud-top pressures and Stage IV hourly precip data to the precipitation assimilation package, increased model output frequency and added products related to precipitation type and convection forecasts. These changes reduced daytime "warm" bias in temperatures and improved the accuracy of 24-hour quantitative precipitation forecasts.

August 8, 2003: The Rapid Radiative Transfer Model (RRTM) longwave radiation package from Atmospheric and Environmental Research, Inc., was installed in the NCEP Global Forecast System model. The new package accounts for more trace gases (CH4, N2O, CFC's) and has more accurate tropospheric water vapor absorption.

August 12, 2003: The North Pacific Hurricane (NPH) wave model became operational. The Western Region, Pacific Region, Ocean Prediction Center, and Tropical Prediction Center requested this implementation to improve their forecast services. The NPH uses the same quarter-degree lon/lat grid as the Eastern North Pacific wave model and is based on the WAVEWATCH III wave model core, with the addition of input winds from the GFDL Hurricane model when tropical systems exist.

August 19, 2003: NCEP completed its three phase implementation, begun in June, of Restricted Data handling to comply with WMO Resolution 40 and other agreements to restrict distribution of observations used in NCEP operations. Registration is required for those within NCEP to access the complete observational dataset including restricted data. No outside access is permitted.

September 16, 2003: A Short Range Ensemble Forecast (SREF) upgrade was implemented. Five new Eta members using the Kain-Fritsch convective parameterization were added to the ensemble. Also added were new cloud and convective products from Eta members. Several new convective fields and vertical levels were included in RSM members at the request of Aviation Weather Center, Storm Prediction Center, and Hydrometeorological Prediction Center. A number of new BUFR soundings were included in output from all 10 Eta members. This implementation increases ensemble diversity and adds new products needed for forecasting aviation weather and severe storms.

September 24, 2003: The Global Ocean Data Assimilation System (GODAS) was implemented. GODAS is quasi-global because the Arctic Ocean is excluded. GODAS uses version 3 of GFDL's Modular Ocean Model (MOM). It also uses a 3-dimensional variational (3DVAR) scheme to assimilate observations of temperature and salinity. GODAS will provide the ocean initial conditions for the seasonal-to-interannual forecasts that will be made with the new NCEP coupled ocean-atmosphere model, scheduled to become operational in 2004.

October 1, 2003: The NCEP Regional Reanalysis (RR) was completed. It creates a 23-year (1979-2002) set of 3 hourly analyses for North America using the Eta Mesoscale Forecast Model (32 km horizontal resolution, 45 vertical layers) and the Eta Data Assimilation System (EDAS). In the future, the RR will continue to run in real-time to add new data to the original data set. Archived RR products will be available to the scientific community at NCDC, NCAR, and the San Diego Supercomputing Center in the first half of 2004. The RR is superior to the completed NCEP/NCAR Global Reanalysis (GR), in both resolution and accuracy because of the various advances that have been made in regional modeling and data assimilation since the GR system was implemented in 1995. These advances include assimilation of precipitation, direct assimilation of satellite radiances, the use of additional data as well as improved data processing efforts, and several Eta Model improvements, particularly those developed within NCEP's GCIP-funded land-surface effort.

November 4, 2003: A fully portable, unified BUFRLIB merging unique components from 3 existing BUFRLIB versions (decoder, verification and endian-independent versions) was created and made operational for NCEP and the Weather Research and Forecast (WRF) system. This new version features standardized documentation and more descriptive error messages as well as many changes to existing routines and the addition of new routines. This enhancement is a prerequisite for ultimate use in the operational data assimilation of mesonet, the Global Positioning System-Integrated Precipitable Water (GPS-IPW), the Radio Acoustic Sounding System (RASS), the Cooperative Agency Profiling (CAP) and boundary layer wind profiler, NEXRAD Level II.5 superobs and Atmospheric Infrared Sounder (AIRS) radiances.

November 20, 2003: A package of several GDAS changes was implemented. Radiance assimilation was improved by a new fast radiative transfer scheme, a new sea surface emissivity model for IR radiances and the addition of GOES-12 sounder radiances. The time interpolation of the guess includes multiple surface fields and multiple sigma guess files. NEXRAD radial winds and lidar line of sight winds were also added to the analysis.

1.4 Operational Performance of the Production Forecast Suites

On-time production of numerical forecast products continued to improve during 2003 The average on-time production on the IBM SP increased to 99.00% in 2003 as compared to 98.88% in 2002. June 2003 recorded a record monthly average of 99.97%. Reliability of receipt of products at the NWS Telecommunication Gateway were greatly improved in 2003, increasing to an average 98.5% from 97.8%. The loss of all telecommunications circuits for 16 hours caused a significant drop in average reliability in July, followed by a 6 hour site wide power outage at the IBM CCS Gaithersburg, MD, facility in August. For the remainder of the year on time generation of forecast guidance remained consistently above 99.5% .

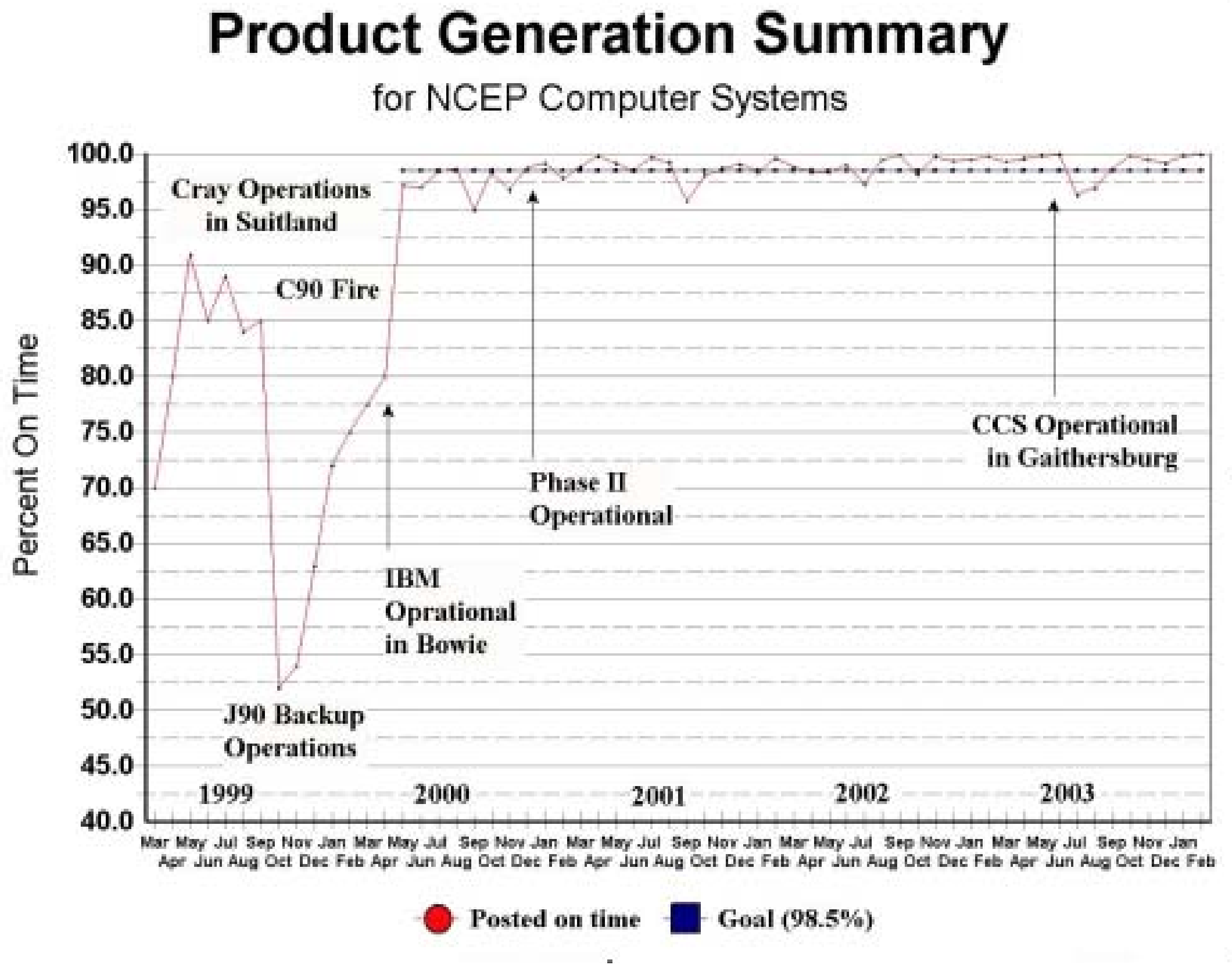

Figure 1. NCEP model guidance product generation - March 1999 to December 2003.

2. Equipment

2.1 IBM Regatta Cluster

The IBM Regatta Cluster is comprised of 44 P690 servers and two P655 servers. Each P690 server comes with 32 1.3 GHz CPUs and 32 Gbytes of memory, except 64 Gbytes in two servers used for I/O. Each P655 server contains 64 processors and 128 Gbytes of memory. During 2003, two additional Regatta Servers were added to the system to support the Community Multi-scale Air Quality (CMAQ) effort. The current configuration of the IBM CCS is contained in Table 1.

The system is configured with a high speed (gigabit) ethernet ring. This enables NCEP to run operations on a portion of the entire system for optimal performance and to share a common NFS file system. A back-end Hierarchical Storage Manager containing 7.0 terabytes (TB) of disk cache and two Storage Tek Silos with a tape archive capacity of about 2 Pbytes is also usable by the entire system.

Table 1. IBM Systems

Processors	1536 1.3 GHz
Memory	1792 GB
Operating Systems	AIX
Disk Storage	42 TB disk subsystem
CPU	Power-4 1.3 GHZ 4 CPUs/Logical Partition

2.2 Additional Components

Independent from the IBM CCS, two additional computers are part of NCEP's Central Computer System. They are an SGI Origin 2000/32 system and an Origin 3000/16 system which share over one TB of disk space. They also share a third Storage Tek silo with eight 9840 drives. Two of these systems are used to support several non-operational projects. A second Origin 2000 is dedicated to running the ocean data assimilation program and the ocean-atmosphere coupled model which is used to prepare seasonal outlooks.

2.3 Communications

Figure 2 shows the communications network used by NCEP to exchange information within the NWS and to provide analysis and forecast products to users via the Internet. NCEP sends and receives data from the IBM facility in Gaithersburg, Maryland via high speed OC3 (155 Mbits/sec) circuits. Communications between the World Weather Building (WWB) and the Telecommunications Gateway computer system at the Silver Spring Metro Center (SSMC) in Silver Spring, Maryland were upgraded from Fast Network Services (FNS) running at 10 Mbits/sec to the faster OC3 circuits. The WWB site is connected to three remote NCEP centers, the Storm Prediction Center (SPC) in Norman, Oklahoma, the

Aviation Weather Center (AWC) in Kansas City, Missouri, and the Tropical Prediction Center (TPC) in Miami, Florida by a T3 line (45 Mbits/sec) through an ATM cloud to three T3 circuits, each with a capacity of 12 Mbits/sec. Access to the Internet from WWB is provided by a direct circuit (155 Mbits/sec) through the Metropolitan Area Network (MAN) ATM cloud to SSMC. From there, information is sent to the Internet over commercial ISP OC3 lines to an Internet Service Provider (ISP).

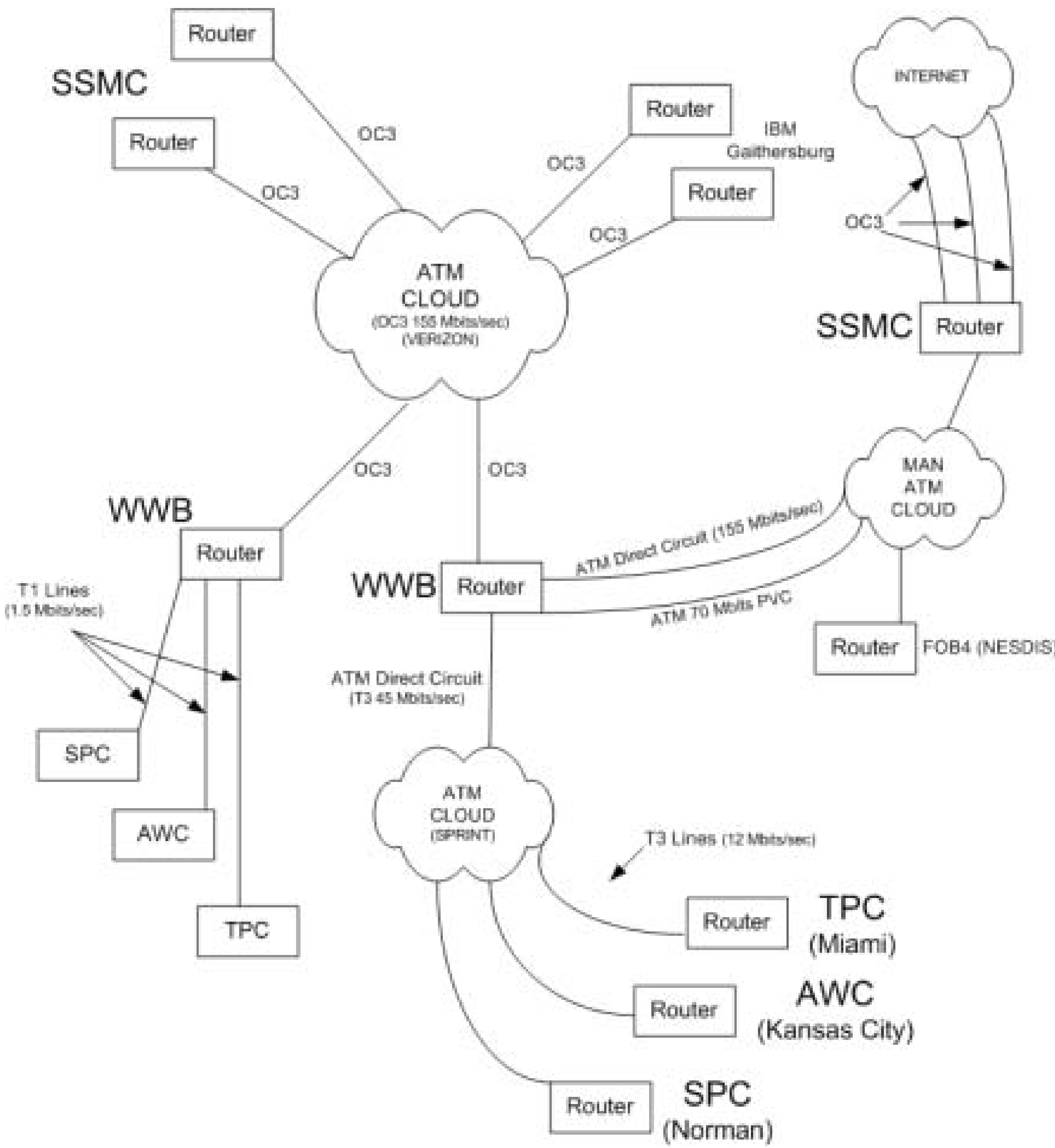

Figure 2. NCEP Network Configuration - 2003

3. Observational Data Ingest and Access System

3.1 Status at the End of 2003

3.1.1 Observational Data-Ingest

NCEP receives the majority of its data from the Global Telecommunications Systems (GTS) and the National Environmental Satellite, Data, and Information Service (NESDIS). The GTS and aviation circuit bulletins are transferred from the NWS Telecommunications Gateway (NWSTG) over two 56 Kbits/sec lines to two co-located Linux workstations, through an interfacing x.25 pad, where the software accumulates the incoming data-stream in files. Each file is open for 20 seconds, after which the file is queued to the Distributive Brokered Network (DBNet) software to send these bulletins to NCEP's Central Operations (NCO). Files containing GTS observations are networked to one of two IBM decoder workstations located at the BCC. There the data-stream file is parsed for bulletins which are then passed to the Local Data Manager (LDM). The LDM controls continuous processing of a bank of on-line decoders by using a bulletin header pattern-matching algorithm. Files containing GTS gridded-data are parsed on the Linux workstations, "tagged by type" for identification, converted to WMO BUFR (Binary Universal Form for the Representation of meteorological data) format and then transferred directly to the IBM CCS by DBNet. There, all observations for assimilation are stored in accumulating data files according to the type of data. Observational data in 24 hour "tank" files remain on-line for up to 10 days before migration to offline cartridges. Observational data in 1 month "tank" files (consisting of some oceanographic data) remain on-line for up to 6 months before migration. While online, there is open access to them for accumulating late arriving observations and for research and study.

3.1.2 Data Access

The process of accessing the observational data base and retrieving a select set of observational data is accomplished in several stages by a number of FORTRAN codes. This retrieval process is run operationally many times a day to assemble "dump" data for model assimilation. The script that manages the retrieval of observations provides users with a wide range of options. These include observational date/time windows, specification of geographic regions, data specification and combination, duplicate checking of bulletins, "part" merging, and parallel processing. The primary retrieval software performs the initial stage of all data dumping by retrieving subsets of the database that contain all the database messages valid for the time window requested by a user. The retrieval software looks only at the date in Section 1 of the BUFR message to determine which messages to copy. BUFR messages in the database are designed such that all reports (i.e., BUFR "subsets") in the message will contain an observation time with the same year, month, day and hour as the date in Section 1 of the message. The observation minute for the BUFR subsets in the message may range from zero to 59. The Section 1 minute is always zero.

This results in an observing set containing possibly more data than was requested, but allows the software to function very efficiently since individual BUFR subsets do not have to be unpacked and time-checked at this point. A final 'winnowing' of the data to an observational set with the exact time window requested is done by the duplicate checking and merging codes applied to data as the second stage of the process. Any data on the reject list receive a quality mark indicating that these data should not be used in operations. However, the flagged data are still processed and monitored. Finally, manual quality marks, in the form of purge or keep flags, and corrections are applied to the data extracted. The quality marks are provided by personnel in two NCEP groups: the NCO Senior Duty Meteorologists (SDMs) and the Ocean Prediction Center (OPC).

4. Quality Control System

4.1.1 Interactive Phase

During the first phase, interactive quality control is accomplished by the Ocean Prediction Center (OPC) and the NCO/Senior Duty Meteorologists (SDMs). The OPC personnel use a graphical interactive system called CREWSS (Collect, Review, and Edit Weather data from the Sea Surface) which provides an evaluation of the quality of the marine surface data provided by ships, buoys (drifting and moored), Coastal Marine Automated Network (CMAN) stations, and tide gauge stations based on comparisons with the model first guess, buddy checks vs. neighboring platforms, the platform's track, and a one week history for each platform. The OPC personnel can flag the data as to the quality or correct obvious errors in the data, such as incorrect hemisphere, misplaced decimal, etc. The quality flags and corrections are then uploaded to the IBM CCS and are stored there in an ASCII file for use during the data retrieval process. The SDM performs a similar process of quality assessment for upper-air profile (rawinsonde, PIBAL, dropwindsonde) temperature, wind and moisture data; single-level aircraft [AIREP, PIREP, AMDAR (Aircraft Report, Pilot Report, Aircraft Meteorological Data Relay) , Meteorological Data Collection and Reporting System (MDCRS), reconnaissance temperature, wind and moisture data; satellite-derived wind data; and surface land and marine pressure, temperature, moisture and wind data. The SDMs use an interactive program which initiates the offline execution of two of the quality control programs (described in the next paragraph) and then review the programs' decisions before making assessment decisions. The SDMs use satellite pictures, meteorological graphics, continuity of data, and input comparisons (buddy checks) to decide whether or not to override automatic data QC flags.

During the past several years, the procedure described above was used daily to screen data that were suspected of containing errors. While CREWSS continues to be used daily, since the Fall of 2003, interactive quality control by the SDM's is only applied selectively in cases where there are indications of egregious errors.

4.1.2 Automated Phase

Separate automated quality control algorithms for rawinsonde, non-automated (AIREP and PIREP) and automated AMDAR aircraft, automated MDCRS ACARS aircraft, wind profiler, and NEXRAD Vertical Azimuth Display (VAD) reports are also executed. The purpose of these algorithms is to eliminate or correct erroneous observations which arise for various reasons including location, transcription or communications errors. Attempts are made, when appropriate, to correct commonly occurring types of errors. Rawinsonde temperature and height data pass through a complex quality control process using extensive hydrostatic, baseline, and horizontal and vertical consistency checks based upon differences from the 6-hour forecast. Corrections and quality marks are applied to the rawinsonde data. A complex quality control algorithm performs the quality control for all levels as a whole, rather than considering the mandatory levels first, and then the significant levels. In addition, an improvement was made to the way in which the hydrostratic residuals are calculated and used. AIREP, PIREP, and AMDAR aircraft reports are also quality controlled through a track-checking procedure by an aircraft quality control program. In addition, AIREP and PIREP reports are quality controlled in two ways: isolated reports are compared to the first guess and groups of reports in close geographical proximity are inter-compared. Both of the above quality control programs are run offline by the SDM as well as in automated QC codes. MDCRS ACARS data receive automated QC that flags a number of common errors in this data type, such as calm winds, spurious reports at 0.0 degrees East or North, and spurious reports above 100 hPa. Wind profiler reports are quality controlled with a complex quality control program using multiple checks based on differences from 6-hour forecasts, including increment, vertical statistical, temporal statistical, and combined vertical-temporal consistency checks. Finally, VAD wind reports are quality controlled with a similar type of program which also includes an algorithm to account for contamination by bird migration.

The final part of the quality control process in the Global Forecast System (GFS) production suite is for all data types to be checked using an optimum interpolation based quality control algorithm, which uses the results of both phases of the quality control. As with the complex quality control procedures, this program operates in a parallel rather than a serial mode. That is, a number of independent checks (horizontal, vertical, geostrophic) are performed using all admitted observations. Each observation is subjected to the optimum interpolation formalism using all observations except itself in each check. A final quality decision (keep, toss, or reduced confidence weight) is made based on the results from all individual check and any manual quality marks attached to the data. Results from all the checks are kept in an annotated observational database.

In the Eta/EDAS production suites, the 3DVAR analysis performs a comprehensive multi-platform quality control of the observations and generates statistical and graphical output of the results.

5. Monitoring System

5.1 Status at the End of 2003

5.1.1 Monthly Data Counts

During the first week of each month the average daily data counts for the previous month are tabulated and displayed on a web site, "www.ncep.noaa.gov/NCO/DMQAB/counts". The statistics are divided into two classes, satellite data and non-satellite data. The counts in each class are further divided into categories based on general data type and subcategories according to specific observing platforms. Average daily data input counts for December 2003 are shown in Tables 2 and 3. The "Percent Input" listed in the righthand column represents the percentage contribution of each category to the total data count.

Table 2. Average Daily Non-Satellite Data Counts for December 2003 for the Global Data Assimilation System

Category	Subcategory	Total Input	Percent Input
Land Surface	Synoptic	56018	
	METAR	119308	
	subtotal	175326	1.483
Marine Surface	Ship	2685	
	Drifting Buoy	12294	
	Moored Buoy	3187	
	CMAN	1699	
	Tide Gauge	3400	
	subtotal	23265	0.197
Land Soundings	Fixed RAOB	1367	
	Ship RAOB	18	
	Dropsonde	3	
	Pibal	258	
	Profiler	668	
	VAD Winds	6044	
	subtotal	8358	0.071
Aircraft	AIREP	3238	
	PIREP	894	
	AMDAR	16760	
	ACARS	80158	
	RECCO	7	
	subtotal	101057	0.855
Total	Non-Satellite	308006	2.607

Table 3. Average Daily Satellite Data Counts for December 2003 for the Global Data Assimilation System

Category	Subcategory	Total Input	Percent Input
Satellite Soundings	GOES Soundings	84084	
	Ozone	3217	
	subtotal	87301	0.739
Satellite Winds	US IR Imagery	66708	
	US Picture Triplet	1459	
	Japan	1569	
	Europe	74585	
	subtotal	144321	1.221
DMSP	SSMI Rainfall Rates	129999	
	SSMI Neural Net	129997	
	subtotal	259996	2.200
Satellite Surface	Quikscat Scatterometer Winds	167190	
	NASA TRMM Data	42375	
	subtotal	209565	1.774
ATOVS	HIRS2 1B	473423	
	HIRS3 1B	1918883	
	MSU 1B	25424	
	AMSUA 1B	563304	
	AMSUB 1B	7825111	
	subtotal	10806145	91.459
	Satellite	11507328	97.393

5.1.2 Next-day Monitoring

"Next-day" data assessment monitoring is accomplished by routinely running a variety of diagnostics on the previous day's data assimilation, the operational quality control programs, and the NWP analyses to detect problems. When problems are detected, steps are taken to determine the origin of the problem(s), to delineate possible solutions for improvement and to contact appropriate data providers if it is an encoding or instrument problem. Data quality is also reviewed to see if platforms should be added or removed from production reject lists.

5.1.3 Delayed-time Monitoring

"Delayed-time" monitoring includes a twice weekly automated review of the production reject list and monthly reports on the quantity and quality of data (in accordance with the WMO/CBS) which are shared with other Global Data Processing System (GDPS) centers. A monthly report is prepared showing the quality, quantity, and timeliness of U. S. supported sites. Monthly statistics on hydrostatic checks and guess values of station pressure are used to help find elevation or barometric problems at upper air sites. This monitoring system includes statistics on meteorological data which can be used for maintaining the reject list and for contacting sites with problems. For global surface marine data, monthly statistics are generated for those platforms which meet specific criteria (e.g. at least 20 observations in a given month). These statistics are forwarded to the UKMET Office in Bracknell, UK (the lead center for marine data monitoring) and are also uploaded to an NCO web site for use by U.S. Port Meteorological Officers. Separate monthly statistics for global drifting and moored buoys are forwarded to the Data Buoy Cooperation Panel, who may then contact the appropriate parties to have faulty buoy data removed from GTS distribution until data problems are fixed.

6. Forecasting System

6.1 Global Forecast System

6.1.1 Status of the Global Forecasting System at the End of 2003

Global Forecast System Configuration: The global forecasting system consists of:

a) The final (FNL) Global Data Assimilation System (GDAS), an assimilation cycle with 6-hourly updates and late data cut-off times. Minor changes to the GDAS were made to process more satellite data in 2003;

b) The Global Forecast System (GFS) initializes the 384-hour forecasts run at 0000, 0600, 1200 and 1800 UTC with a data cut-off of 2 hours and 45 minutes using the 6-hour forecast from the FNL as the first guess. The resolution of the model for the first 84 hours is T254L64, T170L42 from 84 to 180 hours, and T126L28 from 180 to 384 hours. A new Rapid Radiative Transfer Model (RRTM) longwave radiation package was implemented into the GFS model;

c) Ensembles of global 16-day forecasts from perturbed FNL initial conditions (ten forecasts from 1200 UTC, and ten forecasts from 0000 UTC). Ensembles are run at T126 for the first 84 hours and at T62 after that. There have been no changes to the global ensembles in 2003;

Global Data Assimilation System: Global data assimilation for the FNL and GFS is done by a multi-variate Spectral Statistical Interpolation (SSI) analysis scheme using a 3-dimensional variational technique which incorporates a linear balance constraint, eliminating the need for a separate initialization step. The analyzed variables are the associated Legendre spectral coefficients of temperature, vorticity, divergence, water vapor mixing ratio, and the natural logarithm of surface pressure (ln psfc). All global analyses are done on 64 sigma levels at a T254 spectral truncation. Data cut-off times are 0600, 1200, 2100 and 0000 UTC for the 0000, 0600, 1200, and 1800 UTC FNL analyses, respectively, and 0245, 0845, 1445, and 2045 UTC for the 0000, 0600, 1200, and 1800 UTC GFS analyses.

Global Forecast Model: The global forecast model has the associated Legendre coefficients of the natural logarithm of surface pressure, temperature, vorticity, divergence, water vapor mixing ratio and cloud water/ice as its prognostic variables. The vertical domain ranges from the surface to 0.266 mb and is discretized with 64 sigma layers. The associated Legendre spectral coefficients for all variables are trucated (triangular) at T254L64 for the FNL. A semi-implicit time integration is used. The model includes a full set of parameterizations for physical processes, including moist convection, cloud-radiation interactions, stability dependent vertical diffusion, evaporation of falling rain, similarity theory derived boundary layer processes, land surface vegetation effects, surface hydrology, and horizontal diffusion.

Global Forecast System Products: Products from the global system include:

a) Gridded (in GRIB format) Sea level pressure (SLP) and height (H), temperature (T), zonal wind component (U), meridional wind component (V), and relative humidity (R) at a number of constant pressure levels every 3 hours for the first 180 hours of all four runs of the GFS and every 6 hours from 180 hours to 384 hours;

b) Specialized aviation grids (in GRIB format) with tropopause H, T, and pressure as well as fields depicting the altitude and intensity of maximum winds and precipitation;

c) A large number of graphics products.

6.1.2 Future Plans for the Global Forecasting System

Changes planned for the production suite include:

a) Increase the global ensemble resolution to T126 for the first 180 hours;

b) Add 10-member ensemble runs for 0600 and 1800 UTC cycles;

c) Implement a sub-grid scale orographic drag to the GFS model;

d) Implement a global ocean-atmosphere coupled model suite for seasonal prediction;

e) Begin assimilating Aqua/AIRS 1B radiance data;

f) Begin assimilating Cooperative Agency (CAP) and Japanese wind profiler data (in addition to NOAA Profiler Network data now assimilated);

g) Begin assimilating E-EDAS (European AMDAR) aircraft data;

6.2 Regional Forecast System

6.2.1 Status of the Regional Forecasting Systems at the End of 2003
Regional Forecast System Configuration: The Regional Systems are:

a) The Mesoscale Eta Forecast Model (Meso Eta) provides high resolution (12 km and 60 levels) forecasts over North America out to 84 hours at 0000, 0600, 1200 and 1800 UTC. This modeling system and its data assimilation component (see below) was updated in July and September 2003. The July update included: modifications of the cloud microphysics and especially its interaction with atmospheric radiation, extension of 0600 and 1800 UTC runs out to 84-h; and extension of output frequency (hourly posting of model products) and content (more fields that describe different aspects of grid-resolved and parameterized (convective) clouds. Prior to the July update, 0600 and 1800 UTC runs went only to 48 hours. The Technical Procedures Bulletin describing the July changes in detail can be found at http://wwwt.emc.ncep.noaa.gov/mmb/tpb.spring03/tpb.htm ;

b) The Rapid Update Cycle (RUC) System, which generates (20 km and 50 level) analyses and 3-hour forecasts for the contiguous United States every hour with forecasts extended to 12-hr at 0000, 0300, 0600, 0900, 1200, 1500, 1800, and 2100 UTC. In May 2003 this system was updated: 3DVAR replaced Optimum Interpolation as method of analysis and a few minor changes to the RUC post-processing were made including addition of 0-1 km helicity and some small fixes. A Technical Procedures Bulletin describing these changes in detail can be found at http://ruc.fsl.noaa.gov/ppt_pres/RUC-3dvar-tpb-May03.htm ;

c) The Nested Grid Model (NGM), whose North American grid has approximately 90 km resolution on 16 layers, and which generates 48-hr forecasts for North America twice daily at 0000 and 1200 UTC; this model run is used for MOS;

d) A 15 member Short Range Ensemble Forecast (SREF) system runs twice daily at 0900 and 2100 UTC with the forecast period extending out to 63 hours. There are 10 Meso Eta Model members (5 using Betts-Miller-Janjic convection and 5 using Kain-Fritsch convection) and 5 Regional Spectral Model members. All runs are made at 48 km and cover the same North American continental domain as the Meso Eta. They are initialized from their own “breeding” cycles which are used to make two pairs of positive and negative perturbations of the initial conditions. Lateral boundary conditions come from the medium range ensemble forecasts made from 0600 and 1800 UTC. This produces SREF output that can be used together with the early guidance from the Meso Eta runs.

e) A set of High Resolution Window (HiRes) nested runs of the Nonhydrostatic Meso Model (NMM) are made following the completion of the GFS run suite whenever there is no GFDL hurricane model run. These runs get hourly boundary conditions and their initial conditions from the parent continental domain Meso Eta runs (see above). The HiRes version of the NMM uses 60 levels in the vertical, and runs to 48 hours. At 0000 UTC, a large 10 km nest is run for Alaska and a small 8 km domain is run for Hawaii. At 0600 UTC a large 8 km nest is run for the western portion of the CONUS and a small 8 km nest is run for Puerto Rico. At 1200 UTC, a large 8 km nest is run for the central portion of the CONUS and a small 8 km domain is run for Hawaii. At 1800 UTC a large 8 km nest is run for the eastern portion of the CONUS and a small 8 km domain is run for Puerto Rico.

f) A set of Fire Weather support runs of the 8 km 60 level NMM were implemented in May 2003. The Fire Weather control facility in Boise Idaho and NCEP's Storm Prediction Center decide which of 26 selectable 8 km domains are to be run. The choice for each of four runtimes (0000, 0600, 1200 and 1800 UTC) is communicated to the Senior Duty Meteorologist (SDM) who communicates the information to the production suite. These runs are made over the top of the Meso Eta with only a slight time-lag. In this way, they can be made more reliably and more timely than the HiRes runs, and are not subject to conflict with higher priority hurricane runs. The Fire Weather NMM is identical to the NMM in the HiRes runs, with hourly boundary conditions (to 48 hours) and initial conditions coming from the parent Meso Eta run. Gridded guidance is shipped to the TOC where regional centers pick it up and prepare fields for transmission to IMETs who have been dispatched to wild fire locations and who display the guidance on laptops equipped with satellite communications and FX-NET software (to emulate AWIPS). Fire Weather grids are 25% the size of the large CONUS HiRes nests. There are 21 overlapping grids covering CONUS plus 5 grids covering Hawaii, Juneau, Fairbanks, Anchorage and Puerto Rico.

Regional Forecast System Data Assimilation: Initial conditions for the four Meso Eta forecasts are produced by a multivariate 3-dimensional variational (3DVAR) analysis which uses as its first guess a 3 hour Meso Eta forecast from the Eta-based Data Assimilation System. The EDAS is a fully cycled, self-contained system using global fields only for lateral boundary conditions. Data cut-off is at 1 hour and 10 minutes past the nominal analysis times. In most cases, the data time window radius is 1.5 hours about the center of the data dump cycle time. No initialization is applied. Assimilation of observed precipitation (in the form of analyses of automatic gage and radar estimates) is performed within the model component of the EDAS. In 2003, the 3DVAR analysis was upgraded in July. Changes included perpetual cycling of the new cloud microphysical fields, assimilation (in model) of GOES-sounder cloud-top pressures, assimilation (in model) of the Stage IV precipitation, assimilation of super-observations of radial wind data from the NEXRAD 88D radar network, update radiance processing to keep current with global, eliminate thinning of radiances, improve qc of radiances, include NOAA-17 radiances. The Technical Procedures Bulletin describing the July changes in detail can be found at http://wwwt.emc.ncep.noaa.gov/mmb/tpb.spring03/tpb.htm . In September, the use of surface temperature observations over land and the use of all data more than 30 minutes off analysis time were stopped in the 3DVAR.

Since May 2003, initial conditions for the RUC runs come from a multivariate 3DVAR analysis of state variables and a univariate cloud analysis. After the RUC 3DVAR analysis, initialization is achieved using a digital filter which includes diabatic processes.

Since March 2000, initial conditions for the NGM runs come from a combination of Meso Eta analysis over North America and 6-hr forecast from the GDAS over the rest of the Northern Hemisphere. After the Meso Eta analysis is interpolated to the NGM grid, an implicit normal mode initialization is performed.

Regional Forecast System Models: The Meso Eta Model has surface pressure, temperature, u, v, turbulent kinetic energy, water vapor mixing ratio and total condensate as its prognostic variables. The vertical domain is discretized with 60 eta layers with the top of the model currently set at 25 mb. The horizontal domain is a 12 km semi-staggered Arakawa E-grid covering all of North America. An Euler-backward time integration scheme is used. The model is based on precise dynamics and numerics, a step-mountain silhouette terrain representation and includes a full set of parameterizations for physical processes, including Janjic modified Betts-Miller convection, Mellor-Yamada turbulent exchange, Fels-Schwartzkopf radiation, a land surface scheme with 4 soil layers and predictive cloud scheme. The lateral boundary conditions are derived from the prior global model forecast at a 3 hour frequency.

The Nonhydrostatic Mesoscale Model (used for HiRes and Fire Weather runs) has surface pressure, temperature, u, v, turbulent kinetic energy, water vapor mixing ratio, total condensate and the nonhydrostatic vertical velocity and pressure as its prognostic variables. The NMM uses a hybrid sigma-pressure vertical coordinate discretized to 60 layers where the lowest 42 layers are terrain following topped by 18 constant pressure surfaces above 420 mb with the top of the model currently set at 25 mb. The horizontal grid uses an Arakawa E-grid and covers nests of various sizes within the North American domain of the Meso Eta. An Adams-Bashforth time integration scheme is used. The model is based on precise dynamics and numerics, like the Meso Eta, but use of shape-preserving conservative schemes is more prevalent. The NMM includes a full set of parameterizations for physical processes similar to the Meso Eta, including Janjic modified Betts-Miller convection, Mellor-Yamada turbulent exchange, Fels-Schwartzkopf radiation, a land surface scheme with 4 soil layers and predictive cloud scheme. The lateral boundary conditions are derived from the current Meso Eta model forecast at a 1 hour frequency.

The RUC system was developed by the NOAA/Forecast Systems Laboratory. The RUC run provides high-frequency, short-term forecasts on a 20-km resolution domain covering the lower 48 United States and adjacent areas of Canada, Mexico, and ocean. Run with a data cut off of 26 minutes at 0100-1100 and 1300-2300 UTC and 58 minutes at 0000 and 1200 UTC, the analysis relies heavily on high temporal resolution asynoptic data from surface marine and METAR (and soon mesonet) reports, wind profiler, GOES satellite layer PW retrievals and derived wind data, and MDCRS ACARS aircraft data. In most cases, the data time window radius is 0.5 hours about the center data dump cycle time, but for some data types (i.e. satellite derived winds) it is extended further back in time. One of its unique aspects it is use of a hybrid vertical coordinate that is primarily isentropic. Most of its 50 levels are isentropic except for layers in the lowest 1-2 km of the atmosphere where terrain-following coordinates are used. The two types of surfaces change smoothly from one to another. A full package of physics is included with 5 cloud precipitation species carried as historic variables of the model.

The NGM model uses a flux formulation of the primitive equations, and has surface pressure, and sigma u, sigma v, and sigma q as prognostic variables where   is potential temperature and sigma q specific humidity. The finest of the nested grids has a resolution of 85 km at 45 N and covers North America and the surrounding oceans. The coarser hemispheric grid has resolution of 170 km. Fourth-order horizontal differencing and a Lax-Wendroff time integration scheme are used. Vertical discretization is done using 16 sigma levels. Parameterized physical processes include surface friction, grid-scale precipitation, dry and moist convection, and vertical diffusion.

Regional Forecast System Products: Products from the various regional systems include:

a) Heights, winds, temperature, winds aloft:
   (1) Meso Eta to 84 hr at 0000, 0600, 1200 and 1800 UTC every 25 hpa and every 3 hours;
   (2) RUC (to 3 hr or 12 hr) every 25 hpa and hourly; and
   (3) NGM (to 48 hr) every 50 hpa and every 6 hours;

b) 3, 6 and 12 hour accumulated precipitation (totals & convective);

c) Freezing level;

d) Relative humidity;

e) Tropopause information;

f) Many model fields in GRIB format;

g) Hourly soundings in BUFR at hundreds of sites (1400 for Meso Eta, 400 for RUC and NGM);

h) Hundreds of graphical output products and alphanumeric bulletins;

i) HiResWindow and FireWx grids include subset of above and are downloaded from servers and viewable on INTERNET webpages.

Operational Techniques for Application of Regional Forecast System Products: Model Output Statistics (MOS) forecasts of an assortment of weather parameters such as probability of precipitation, maximum and minimum temperatures, indicators of possible severe convection, etc. are generated from NGM and Meso Eta model output. These forecasts are made using regression techniques based on statistics from many years of NGM forecasts and several seasons for the Meso Eta.

6.2.2 Future Plans for the Regional Forecast Systems

Year 2004 Goals

a) Replace deterministic NMM runs in HiResWindow with 6 member ensemble made from equal number of runs based on NCEP developed WRF-NMM and NCAR developed WRF-Mass-core.

b) Implement Air Quality Forecasting (AQF) system for Northeast US domain consisting of EPA's CMAQ model for ozone driven by Meso Eta meteorological fields.

c) Upgrade Meso Eta Model, then freeze it until WRF replacement in late 2005 - final upgrades to include: radiation schemes used in GFS, refined treatmensts for deep and shallow convection, refined gridscale cloud & precip, upgrades to NOAH Land-Surface Model, improved assimilation via model nudging of GOES cloud & precipitation analyses and use of digital filter and move model top to 2 mb for assimilation of satellite radiances.

d) Upgrade Meso Eta 3DVAR analysis, then freeze it until WRF capability in late 2005 - final upgrades to include: assimilate doppler radar Level II.5 radial velocities, assimilate AIRS data, assimilate surface mesonet data, assimilate GPS IPW data and move top to 2 mb.

e) Improve Regional Ensembles in both resolution (from 48 km to 32 km) and diversity of members (eliminate control runs and reduce to single bred pairs to allow 7 model/physics versions instead of only 3), add hourly BUFR output, improve ensemble products and their display & dissemination.

f) Provide GFS grids downscaled to 12 km via Meso Eta model out to 8 days for NWS filed-office use in initializing thier medium range NDFD grids, for CONUS domain at 0600 and 1800 UTC and for Alaska, Hawaii and Puerto Rico at 0000 and 1200 UTC.

g) Upgrade RUC to higher resolution (13 km) and make physics refinements;

h) Upgrade NMM in FireWx & Homeland Security runs to the WRF (Weather Research and Forecasting (WRF) common modeling infrastructure version WRF-NMM.

i) Begin assimilating surface mesonet data;

j) Begin assimilating GPS Integrated Precipitable Water (GPS-IPW) data;

k) Begin assimilating GOES-12 radiance data (Eta/EDAS);

l) Begin assimilating Cooperative Agency (CAP) wind profiler data (an addition to NOAA Profiler Netword data now assimilated).

6.3 Specialized forecasts

a) The global ocean data assimilation system (GODAS) has been implemented into production. The GODAS uses the Modular Ocean Model version 3 (MOM3) together with a 3-D VAR to produce the state of the ocean on a daily basis. The atmospheric fluxes from GDAS are used to provide the surface forcing. The product of the GODAS will be used to initialize a global ocean-atmosphere coupled model to be implemented in 2004.

b) The East Coast-Regional Ocean Forecast System produces a nowcast and a 48-hour forecast once every day, after becoming fully operational at NCEP in March 2002. Graphic products and grid point numerical fields are made available to the user community over the internet at http://polar.ncep.noaa.gov..

c) NCEP is working on a common coupled ocean-atmosphere modeling framework that would make it possible to provide forecasts on regional, basin, and global scales, as needed, in a unified and efficient manner. This new effort is being undertaken using the Hybrid Coordinate Ocean Model (HYCOM) system. In NCEP's HYCOM system, regional and basin scale domains will be nested in the global domain as needed. All the models will be forced by the fluxes from NCEP's GFS and will include a full suite of ocean data assimilation procedures and tidal forcing. Initial HYCOM experiments focus on the Atlantic Basin between 25 S and 75 N latitude. The horizontal grid resolution varies from about 18 km in the south eastern Atlantic to 4-6 km in the west and has 26 vertical coordinate levels. Several experiments are currently underway to test the sensitivity of the model to different vertical mixing schemes, parameterization of atmospheric fluxes, boundary conditions at the two open boundaries at the northern andsouthern extremities, etc. The system will be capable of supporting NOAA's National Ocean Service's development of higher resolution bay and estuarine forecast models as well as NCEP's development of improved coupled ocean-atmosphere models for tropical hurricane prediction.

d) Currently, a simple quality control procedure based on JPL rain flag detectors for QuikSCAT winds is used in the operational NCEP global data assimilation system. We evaluated this quality control procedure and found that QuikSCAT wind data are not useful when the rain flag probability of a retrieval is greater than 10 %. A data assimilation experiment showed favorable results when QuikSCAT winds with a rain flag probability higher than 10% are not used in the NCEP GDAS. One problem with this approach is that any good quality QuikSCAT wind data over storm areas is discarded. To address this deficiency, NCEP tested a new QC procedure based on analyses of maximum likelihood estimates in a global data assimilation experiment using 53 days of 2003 QuikSCAT data. The results showed that this QC procedure may be useful in certain synoptic storm forecast situations, but is not viable in the overall performance of global forecasts.

e) A data assimilation experiment testing higher resolution (half degree longitude-latitude superobed) QuikSCAT winds showed a very significant improvement in the forecasts of wind and mass fields in the tropics. As a result, the half degree longitude-latitude superobed QuikSCAT wind data were implemented at NCEP GDAS on March 11, 2003. In anticipation of the arrival of ADEO-II SeaWind data (12 km spatial resolution), another data assimilation experiment was conducted using the full resolution (25 km) QuikSCAT wind data at NCEP, and results of this experiment are under active investigation.

f) Sea-state forecasts are issued daily at the conventional four synoptic time cycles. Apart from official forecasts issued by the NWS, wave predictions are made available via MMAB=s web site, which receives approximately 35,000 hits daily. Sea-state forecasts made at NCEP presently use several implementations of the numerical wave model WAVEWATCH III, version 2.22. Wave model forecasts are driven using operational products generated by other model and analysis fields at NCEP such as the global atmospheric wind and air temperature forecast fields SST and sea-ice analysis fields. The combination of the wave model and its inputs is known as the NOAA WAVEWATCH III (NWW3) wave model suite.

g) The operational NWW3 model suite consists of a global model and three Aall-year-round@ regional models covering the following domains: Gulf of Alaska and Bering Sea (AKW), Western North Atlantic (WNA) and Eastern North Pacific (ENP). During hurricane seasons, the NWW3 expands to include two other regional implementations: the North Atlantic and North Pacific hurricane models, NAH and NPH. The NPH became operational on August 12, 2003. Recent upgrades to NWW3 include the implementation of a regional model for the eastern North Pacific Ocean (ENP) model and the introduction of two specialized implementations of WAVEWATCH III (NAH, NPH) to predict waves generated by hurricanes. Wind input resolution was upgraded to match the T254L64 of the GFS, and the hurricane wave models now use hourly wind fields. All models now forecast out to 7 days (hurricane wave models maintain their 3-day forecast limit).

h) NCEP uses a separate atmospheric model developed by NOAA=s Geophysical Fluid Dynamics Laboratory (GFDL) during hurricane seasons. NCEP runs the GFDL model with a system of two nested moving grids whenever the development of a tropical storm is detected in the northern hemisphere. Model runs are made for individual storms, producing detailed wind fields for up to four individual co-existing tropical cyclones. NWW3 hurricane wave models use the GFDL winds as input. NAH and NPH wave model grids are identical to the WNA and ENP GFDL model domains, respectively. The WNA and ENP models are run even if hurricanes are present, providing an `early= forecast. The NAH and NPH model results are available later because they must wait for the completion of the GFDL runs. Even if no GFDL data is available, the NAH and NPH models are run throughout the hurricane season to assure that hurricane swells can be tracked for their entire life cycle in each model domain.

i) An assimilation scheme for wave heights (observed at buoys and with altimeters) for the NWW3 global wave model is being tested and validated through parallel runs. Preliminary results indicate a positive impact on wave height forecasts up to 24h, but little impact thereafter.

j) Future Directions for NWW3 (multi-year projects):

- Development of a multi-scale environment for WAVEWATCH III, with two-way information flow between different domains. Replace the present set of models with a single multi-scale model, using high resolution only where needed.
- Coupling the wave model with atmosphere and ocean models.
- Add products reflecting user demand, such as wave steepness measures.
- Expand the wave field separation scheme to produce spatial fields of individual wave systems.
- Start producing wave ensemble forecasts.
- Investigate alternative economical approximations of the well known exact interactions to improve the spectral shape.
- Improve the dissipation source term to include explicit steepness measures for later model applications including effects of mean currents and mean water level changes.
- Improve and retune wind input and wave-energy dissipation source terms to accommodate changes in the nonlinear interactions formulation.

k) The CMAQ model is under development by the Environmental Protection Agency (EPA) in response to a Congreesional amendment to the Clean Air Act (CAA) to bolster America's efforts to attain the National Ambient Air Quality Standards (NAAQS). The amendment required further reductions in the amount of permissible tailpipe emissions, initiated more stringent control measures in areas that still failed to attain the NAAQS (non-attainment areas), and provided for a stronger, more rigorous linkage between transportation and air quality planning. The initial phase of testing took place during the summer of 2003 and will continue through the summer of 2004, with a planned implementation in the fall of 2004.

l) A dynamical seasonal forecast system is currently operational at NCEP and provides guidance for generating seasonal forecast outlooks at the Climate Prediction Center (CPC). This production suite, based on an atmospheric model, is run once a month and consists of: (1) A 10-member ensemble of atmospheric hindcast runs for the 1979-1999 period. All the atmospheric simulations in the hindcast are forced with the observed SST's and are valid 7 months from the initial conditions. Hindcasts provide an assessment of the level of forecast skill and the impact of climatology on forecast anomalies. (2) The second componentof the production suite is a 20-member forecast ensemble. Forecast simulations are forced by the SST's predicted by the NCEP's coupled ocean-atmosphere model.

6.4 Verification of Forecast Products for Year 2003

Annual verification statistics are calculated for NCEP's global models by comparing the model forecasts to the verifying analysis and the model forecast interpolated to the position of verifying rawinsondes (see Tables 4 and 5). The parameter used in the verification is the inverse S1 score. The lower this score, the more accurate were the model predictions.

Table 4. Verification against the Global Analysis for 2003. Change from 2002 statistics in parentheses. Negative values (BLUE) indicate improvement; positive values (RED) indicate degradation.

Statistic	GFS 24 hr	GFS 72 hr	MRF 120 hr
500hPa Geopotential RMSE(m)			
Northern Hemisphere	12.3 (+1.1)	31.0 (-0.2)	56.1 (-0.4)
Southern Hemisphere	14.3 (+0.1)	37.0 (-3.4)	63.9 (-6.4)
250hPa Wind RMSVE(m/s)			
North Hemisphere	4.6 (+0.1)	9.8 (-0.2)	14.9 (-0.3)
Southern Hemisphere	4.8 (-0.1)	10.3 (-0.4)	15.5 (-0.6)
Tropics	4.1 (0.0)	6.9 (0.0)	8.5 (-0.1)
850hPa Wind RMSVE(m/s)			
Tropics	2.7 (-0.1)	4.1 (-0.1)	4.9 (0.0)

Table 5. Verification against rawinsondes for 2003. Change from 2002 statistics in parentheses.
Negative values (BLUE) indicate improvement; positive values (RED) indicate degradation.

Statistic	GFS 24 hr	GFS 72 hr	MRF 120 hr
500hPa Geopotential RMSE(m)			
North America	12.8 (0.0)	31.5 (-2.7)	54.8 (-2.0)
Europe	15.7 (+0.5)	33.7 (+1.3)	59.9 (0.0)
Asia	14.7 (-0.2)	28.3 (-1.9)	45.4 (-4.2)
Australia/New Zealand	12.4 (+0.5)	21.6 (-1.0)	36.1 (-1.1)
250hPa Wind RMSVE(m/s)			
North America	6.4 (0.0)	11.7 (-0.4)	16.9 (-0.7)
Europe	5.9 (-0.2)	11.3 (0.0)	17.5 (-0.1)
Asia	6.3 (-0.3)	10.6 (-0.3)	14.1 (-0.3)
Australia/New Zealand	6.2 (0.0)	9.5 (-0.3)	13.2 (-0.5)
Tropics	5.7 (-0.1)	7.5 (-0.1)	8.6 (-0.3)
850hPa Wind RMSVE(m/s)			
Tropics	4.3 (0.0)	5.2 (+0.1)	5.9 (+0.2)