The Development and Success of

NCEP's Global Forecast System

Development Division/Environmental Modeling Center

National Meteorological Center/National Centers for Environmental Prediction

U. S. Department of Commerce

National Oceanic and Atmospheric Administration

National Weather Service

To be submitted as an NCEP Office Note

LongGFSHist.docx

Dec. 30, 2018

Corresponding authors: Glenn White (gw15911@gmail.com), Fanglin.Yang@noaa.gov,

Vijay.Tallapragada@noaa.gov

Abstract

The invention of the first electronic computer in 1946 enabled a continuing revolution in weather forecasting driven by analyses and forecasts produced by integrating the equations of motion of the atmosphere. In the United States public forecasts have become increasingly based on computer analyses and forecasts produced by the National Centers for Environmental Prediction (NCEP) Production Suite, the foundation of which is the Global Forecast System (GFS). Most of the other NCEP forecast systems are dependent on fields from the GFS.

Following the introduction of spectral techniques in the 1970s that produced forecasts as skillful as and more computationally efficient than finite differencing methods, Joseph Sela began to develop spectral modeling at NCEP (then the National Meteorological Center (NMC)) in 1975. The first global NMC model with a hydrostatic spectral dynamic core became operational in Aug. 1980 with a limited physics package, 12 layers in the vertical and a horizontal resolution of 30 waves, equivalent to 375 km. Over the years vertical and horizontal resolution have increased as has the complexity of the physics and the dynamic core. Its current resolution is 64 layers in the vertical and 1534 waves, equivalent to 13 km. Of at least equal importance have been advances in data assimilation techniques and observations.

The GFS's resolution now approaches the atmospheric scales where non-hydrostatic effects are important. The American meteorological community joined together with NCEP to test several non-hydrostatic dynamic cores and choose a new dynamic core for the GFS. It is therefore fitting for this paper to review the evolution and increasing skill of the GFS over its 39-year history.

1. The emergence of NWP

A. The 19th Century

Humans have long tried to forecast the weather. For most of our history, forecasts have been based empirically on observations of the weather—such as numerous weather proverbs or a Vermont dairy farmer forecasting whether his hay would get wet by observing whether storm clouds passed over a nearby mountain or to either side. Technological revolutions have enabled revolutionary improvements in weather forecasting. In the 19th century the invention of the telegraph and the development of telegraphic networks enabled the development of operational analyses and forecasts at the Smithsonian under the leadership of Joseph Henry and in England under the leadership of Robert Fitzroy (Moore, 2015). Following the wreck of the *Royal Charter* in a terrible storm in 1858, Fitzroy wrote:" Man cannot still the raging of the wind, but he can predict it. He cannot appease the storm, but he can escape its violence …" (Slingo, 2017).

B. Richardson 1922

The idea of numerical weather prediction (NWP), predicting weather by integrating in time the equations governing the atmosphere, was first recorded by Bjerknes (1904) (Shuman, 1989; Tribbia and Anthes, 1987). During World War I Richardson (1922) made the first numerical weather forecast; the resulting 6-hour forecast was quite unrealistic (Shuman, 1989; Kalnay, 2003). Two obstacles he faced were the large number of computations required and the lack of observations; these obstacles were overcome by the development of electronic computers and the radiosonde network after World War II (Edwards, 2010). Richardson integrated the full

primitive equations with an unbalanced initial condition and too large a time step (Kalnay, 2003). Decades later initial progress in NWP required a simplified set of equations and considerable research and development.

C. The first computer forecast 1950

In May 1946 the U. S. Navy's Office of Research and Inventions funded the Meteorology Project at the Institute for Advanced Study at Princeton University (Lynch, 2008), leading to the first weather forecasts performed on an electronic computer in 1950 by a group led by Jule Charney and John von Neumann (Charney *et al*., 1950; Shuman, 1989). They integrated the non-divergent barotropic vorticity equation, a simplified form of the equations governing the atmosphere inspired by and very similar to a simplified linearized equation developed by Carl Gustav Rossby in 1939 (Rossby *et al*., 1939; Rossby, 1940), except for Charney's inclusion of non-linear advection of vorticity (Kalnay *et al*., 1998). Klara Dan von Neumann programmed the first computer forecasts and taught the scientists how to program (Edwards, 2010). Lynch (2008) reviewed the skill of these forecasts.

D. JNWPU 1954

The Joint Numerical Weather Prediction Unit (JNWPU), organized by the U.S. Weather Bureau (now the National Weather Service (NWS)) and the weather services of the U.S. Air Force and Navy, began July 1, 1954. In Feb. 1955 the Weather Bureau purchased an IBM 701 for JNWPU, a 1 kiloflop computer, IBM's first fully electronic computer; it was the last of 19 produced by IBM and was put together from spare parts (DiMego, 2010). JNWPU issued its

first operational forecast May 6, 1955 (Harper *et al*., 2007). Sweden had produced the first operational NWP forecast in late Sep. 1954 with a barotropic model (Persson, 2004).

At first JNWPU used what Shuman (1989) called "the most sophisticated model of the time", a three-level quasi-geostrophic model developed by Charney (1954). Of very limited use to forecasters, it was replaced by a two-level model (Thompson and Gates, 1956). The two-level model also failed to produce useful forecasts and was replaced by a single-level barotropic model (Charney, 1949). Even though the barotropic model as well initially failed to produce useful forecasts, its relative simplicity made it easier to diagnose and improve. In 1958 the barotropic model with advanced finite differences and initial conditions produced by objective analysis made timely forecasts that significantly improved 500 hPa height forecasts (Shuman, 1989; Kalnay, 2003); the persistent upward trend in skill continues today, as shown in Fig. 1. Fig. 1 is based on the S_1 skill score (Teweles and Wobus, 1954; Hirano, 1992) which measures the relative error in the pressure gradient in 500 hPa height over North America. 72h forecasts now have reached and 36h forecasts now exceed the skill level considered perfect in 1954, reflecting greater precision in NWP analyses and forecasts than in human analyses.

E. NMC 1958

In January 1958 the JNWPU divided into three units to transfer NWP from research to operational weather forecasting; the Air Force and Navy formed their own NWP centers, and NWS founded the National Meteorological Center (NMC) (McPherson, 1994). Following the success of the improved barotropic model, the complexity of operational NWP systems increased as computers became more powerful, enabling aggressive research and development. Table 1

lists the computers used by NMC/National Centers for Environmental Prediction (NCEP) over the years (DiMego, 2010; Kane, personal communication). The speeds listed represent the computers' theoretical limits on specific algorithms; even after considerable effort was devoted to redesigning codes, the actual speeds obtained in running NWP systems were considerably less (G. DiMego, personal communication). Mesinger *et al.* (2018) discussed the continuing evolution of numerical methods to exploit the increasing power and changing architecture of computers.

F. References

Glahn (2012a, b) reviewed the first 100 years of the NWS up to 1970, including JNWPU and NMC. Kalnay (2003) presented the history and science of NWP. Shuman (1989) reviewed the development of NMC numerical forecasts from 1954 to 1989; McPherson (1980) and Dey (1989) discussed the evolution of observations and data assimilation at NMC through 1988. As part of a 2004 Symposium on the 50th Anniversary of Operational NWP (Harper *et al*., 2007), (http://www.ncep.noaa.gov/nwp50/Presentations/), DiMego (2004) reviewed the history of supercomputers used by NMC/NCEP.

NMC/NCEP Office Notes detail many aspects of forecast system development from 1955 to 2012 (http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html). Acronyms used in this paper are listed in Appendix A. Many documents can also be found on NOAA's Institutional Repository (https://repository.library.noaa.gov/).

2. First global forecast systems

A. Global branch

The changes in the NCEP GFS from 1974 to 2018 are listed in Appendix B. Much of the GFS development work described below was done by the global modeling branch (since 1994, the Global Climate and Weather Modeling Branch) within the Development Division (DD) (now the Environmental Modeling Center (EMC)), but the work often involved people in EMC outside the global branch. Reorganization in 2017 established one modeling branch for all forecast systems.

B. First satellite observations

Satellites offered the first truly global observations. Launched into geosynchronous orbit June 6, 1966, the Applications Technology Satellite (ATS-1)'s television camera produced sequential images of cloud masses that could be tracked in time and converted into an atmospheric motion vector (AMV). AMVs were first used operationally in 1967 (Bristor, 1975; Dey, 1989). Menzel (2001) reviewed the development of AMVs. Remotely sensed temperature soundings first entered the upper air data base May 29, 1969 (Dey, 1989). Vertical Temperature Profile Radiometer (VTPR) instruments on board NOAA 2-5 satellites from 1972 to 1979 provided routine atmospheric sounding measurements by an infrared radiometer with 6 temperature sounding channels plus water vapor and atmospheric window channels (Dudhia, 2015). The National Environmental Satellite, Data and Information Service (NESDIS) retrieved vertical temperature profiles based on the radiometer measurements.

C. First GFS Sept. 18, 1974

Following the introduction of satellite observations (Dey, 1989), more sophisticated atmospheric models and more powerful computers, NCEP's first global system was implemented on Sept. 18, 1974 (Shuman, 1989), using the Hough analysis (Flattery, 1971) and a 2.5^{o} 9-layer sigma coordinate global primitive equation model (Stackpole, 1973, 1978). It was initially motivated by the need to provide initial conditions for an expanded Northern Hemisphere (NH) grid whose corners extended into the Southern Hemisphere (SH) and by the desire to use VTPR satellite profiles (R. Kistler, personal communication). Initially generating short range forecasts in the Global Data Assimilation System (GDAS), it was soon used for 5 and later 10-day forecasts (Shuman, 1989).

Global solutions to Laplace's tidal equations (Flattery, 1971), Hough functions represented some of the natural oscillatory modes of the atmosphere, introducing a physical basis for analysis (Shuman, 1989; McPherson, 1980). Since Hough functions related wind and pressure fields, they were well suited for analyzing globally a mix of wind and pressure observations (Shuman, 1989). The Hough analysis produced an essentially nondivergent wind, creating imbalances between the wind and mass fields in short range forecasts (White, 1983). Cressman (1980) and White (1983) reviewed problems in vertical motions from NMC 6 h forecasts during the 60s and 70s.

The GFS for many years had three roles: a) as GDAS to generate initial conditions for global forecasts and first guess fields and boundary conditions for regional forecast systems, b) as the AVN, a short-range system with a shorter data cutoff for aviation forecasts run initially

twice a day, and c) as the MRF, a medium-range system with a longer data cutoff run once a day. In the early years the three systems did not always use the same version of the GFS. Eventually the AVN and MRF were replaced by one medium-range forecast four times a day with a shorter data cutoff time; a final analysis with a longer data cutoff time incorporated later arriving data four times a day.

In 1975 NMC gained access to an offsite backup IBM computer at Patuxent River Naval Air Station, requiring driving 52 miles in the middle of the night with multiple boxes of punch cards of programs to be tested (D. Parrish, personal communication). In 1977 NMC had three IBM 360/195 computers, each with a single CPU. The global forecast took approximately 6 hours to run; no other jobs larger than 256 kilobytes or longer than 1 minute of CPU time were allowed on the IBM during the day.

Fixing delivery times began gradually. If the same start time was used with a new, faster computer, the run could easily finish earlier and be delivered sooner. Once accustomed to the new delivery time, forecasters did not want to return to a later delivery time when a more sophisticated model was implemented on the same computer and required more time. DD worked with the computer division to use less than the maximum compute to maintain the same delivery time (G. DiMego, personal communication).

D. Optimal interpolation Sept. 22, 1978

In 1977 the 12-h forecast used in GDAS as the background field in the analysis was changed to a 6-h forecast (Dey, 1989). An optimum interpolation (OI) analysis replaced the

Hough analysis in GDAS Sept. 22, 1978 (Gandin, 1963; Bergman, 1979; McPherson *et al*., 1979; Dey, 1989). The new system was preferred over the Hough analysis in handling the heterogeneous data base produced by the First Global Atmospheric Research Program (GARP) Global Experiment (FGGE) (R. Kistler, personal communication).

OI was designed from a carefully formulated mathematical/statistical theory (Gandin, 1963), guided by three considerations: spatial coherence, temporal continuity and dynamic consistency (Derber *et al.*, 1991). It tried to make the analysis an optimal combination of the observations and the 6-h forecast from the previous analysis. While OI theory required that all observations be combined simultaneously with the 6-h forecast to produce the new analysis, the practical application of OI had to be done for a particular grid-point or small volume of grid-points, requiring computationally expensive and logically complex data selection that did not always make the best decisions. In 1991 the implementation of the Spectral Statistical Interpolation (SSI) eliminated the need for spatially-based data selection (Derber *et al*., 1991).

E. TIROS TOVS 1978

The Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) first flew on TIROS-N in 1978 and then on NOAA 6-14 consisted of three instruments: a) High-resolution Infrared Radiation Sounder (HIRS) /2, a 20-channel infrared radiometer with 12 temperature sounding channels and water vapor, ozone and atmospheric window channels; b) Stratospheric Sounding Unit (SSU) for stratospheric temperature sounding; c) Microwave Sounding Unit (MSU), a four-channel microwave radiometer (Dudhia, 2015). Initially NESDIS

used statistical retrievals based on the comparison of radiance data to collocated radiosonde reports to produce temperature soundings.

3. Introduction of the spectral GFS 1975-1986

This paper reviews the history of the NCEP GFS, the foundation of the NCEP Production Suite since 1980. The model component of the GFS is a spectral model initially developed by Sela (1980). Its early development and the resulting improvements in forecast skill were discussed by Caplan and White (1989), Caplan *et al*. (1989), Dey (1989), Bonner (1989), Kalnay *et al*. (1990), Kalnay *et al*. (1998), and Kalnay (2003). Table 2 summarizes the increasing resolution of the GFS and the most important changes in dynamics, physics and data assimilation over the years.

A. Spectral GFS R30L12 May 27, 1980

In the 1970s interest developed in spectral techniques as a more computationally efficient alternative to finite differencing methods in NWP, following the introduction of the transform method (Orzag, 1970; Eliasen *et al*., 1970) and its application to a multilevel primitive equation model (Bourke, 1970). At NMC Joseph Sela began the development of spectral modeling in 1975; most of the physical parameterizations came from the then operational primitive equation grid point model (Shuman and Hovermale, 1968; Stackpole, 1973; Sela, 1980). A sigma vertical coordinate (Phillips, 1959) was used with an Arakawa quadratic conserving finite differencing in the vertical (Arakawa and Mintz, 1974).

The physics included large-scale precipitation, Kuo-type convection (Kuo, 1965, 1974; Phillips, 1979), and evaporation and sensible heat flux over the oceans but not land (Sela, 1980); to ease conversion between physical and spectral space, physical processes were calculated on a Gaussian grid. Sub grid-scale dissipation was parameterized by 4^{th}-order horizontal diffusion; orography was "passed through a 9-point filter twice to remove excessive ocean irregularities when expressed spectrally" (Sela, 1980). Over the history of the GFS considerable effort has gone into modeling mountains and their effects, adjusting diffusion and filtering fields to achieve the most accurate forecasts and address model problems.

The spectral model proved comparable in skill to the operational grid point model in global and hemispheric forecasts; a rhomboidal truncation 30-wave (R30) 12-layer (L12) (equivalent to 375 km) global spectral model was implemented May 27, 1980 in GDAS and in Aug. 1980 in AVN and MRF forecasts. At 48 h horizontal resolution was reduced to 24 waves; at 144 h vertical resolution was reduced to 6 layers; at 192 h the domain was reduced to the NH. Forecasts extended to 240 h at 000 UTC and 60 h at 1200 UTC. A non-linear normal mode initialization (Machenhauer, 1977; Ballish, 1981) reduced unwanted gravity waves. Kistler and Parrish (1982) discussed changes in GDAS from 1978 to 1981 and found the spectral GDAS superior to the gridded system used before. Initially the AVN and MRF forecasts were initialized by the Hough analysis (Flattery, 1971); in Aug. 1984 the multivariate OI analysis (Bergman, 1979) replaced the Hough analysis.

B. Multivariate OI on pressure surfaces Aug. 14, 1982

Dey and Morone (1985) reviewed updates to GDAS in 1982 and 1983: the most significant change was the replacement Aug. 14, 1982 of the quasi-univariate, sigma-coordinate OI analysis by a full multivariate analysis on mandatory pressure surfaces, in response to problems with wind corrections where satellite temperature soundings were the only observations (Dey ,1989; Kistler and Parrish, 1982).

C. R40L12 Oct. 19, 1983

Acquiring the Cyber 205 in 1981 and the use of highly optimized Fast Fourier Transforms developed by C. Temperton of the United Kingdom Meteorological Office (UKMO) allowed the implementation of a R40L12 (300 km) global spectral model Oct. 19, 1983 (Gerrity *et al*., 1985); the Cyber 205 was the first parallel computer with two long vector pipes, significantly increasing speed. Shuman (1989) reviewed the computers used over the years. One factor not always considered in computer acquisition was the time required to convert operational and experimental code from one computer to another; developers found the conversion to the Cyber 205 especially time-consuming (R. Kistler, personal communications; DiMego, 2010).

D. MRF85 E2 physics Apr. 17, 1985

The MRF85 upgraded physical parameterizations Apr. 17, 1985 in the MRF once daily forecast at 0000 UTC after a 6 h data cutoff. Vertical resolution increased to 18 equally spaced layers; only the bottom 12 included forecast moisture. The AVN continued to run the R40L12

spectral model with limited physics twice a day at 000 and 1200 UTC after a 3 h 30 m data cutoff (Bonner, 1989).

Extensive collaboration with the Geophysical Fluid Dynamics Laboratory (GFDL) led to the implementation of "E-2" physics (Miyakoda and Sirutis, 1986) in the MRF85, including shortwave (Lacis and Hansen, 1974) and longwave (Fels and Schwarzkopf, 1975) radiation, zonally symmetric climatological clouds, interaction with the surface over both land and ocean and Kuo cumulus convection up to 300 hPa. In the Kuo parameterization, convection occurs where there is conditionally unstable stratification and moisture convergence exceeds a set value; converged moisture is partitioned into moistening and heating depending on the mean relative humidity of the column; the vertical heating profile is determined by the difference between the predicted sounding and the moist adiabat constructed from the planetary boundary layer (PBL) temperature and humidity (Kanamitsu, 1989). The absence of convection above 300 hPa in the MRF85 led to a shallower Hadley cell in the forecasts than analyzed with sinking motion above the rising branch of the Hadley cell in the 1-day forecasts (White, personal communication).

A PBL parameterization based on Monin-Obukhov similarity theory governed the exchange of momentum and sensible and latent heat between the surface and the atmosphere (Miyakoda and Sirutis, 1977), although the 56 hPa thick lowest layer was too thick to compute vertical eddy diffusion accurately. Paul Long modified the parameterization for very stable and very unstable situations (Environmental Modeling Center, 2003). Within the PBL eddy diffusion mixed water vapor and momentum; the exchange coefficient, a function of vertical wind shear and a linearly varying mixing length, vanished at 2500 m above the surface. A Kessler-type

parameterization of the evaporation of non-convective rain and cooling below cloud base was also introduced (Kessler, 1969).

Since the coarse spatial resolution of the atmospheric models at this time underestimated the effect of mountains on the atmospheric flow, enhanced silhouette mountains (first suggested by Mintz and adapted by Mesinger *et al*., 1988) were implemented in the MRF85 (Caplan, 1985). A higher resolution (1/6^{o}) orographic data set defined the highest heights in each horizontal coordinate direction in each model Gaussian grid box; the highest heights were averaged together. Silhouette mountains were later removed in 1991 because they complicated the assimilation of surface pressure observations. Wallace, Tibaldi and Simmons (1983) discussed the motivation for enhanced envelope mountains in the European Centre for Medium Range Weather Forecasts (ECMWF) forecast model.

The experimental MRF85 ran in parallel with the operational spectral system and skill scores; synoptic fields and systematic errors and differences were scrutinized, compared and discussed (Caplan, 1985; Gerrity *et al*., 1985). In forecasts for Feb. and Mar. 1985, the new GFS surpassed 60% monthly mean anomaly correlation for 500 hPa height through 5 days in the extratropical NH, a day longer than the old model (Gerrity *et al*., 1985); much of the gain may have come from the introduction of silhouette mountains.

E. MRF86, new physics May 29, 1986

The MRF86 unified the GDAS and the MRF May 29, 1986, improving the assimilation, introducing a new version of GFDL physics, correcting problems and adding new physical

processes (Gerrity, 1986). Previously GDAS had used an older R30L12 version of the spectral model with more modest physics. Table 1 of Dey *et al*. (1987) compared the old and new GDAS. The new GDAS used the MRF86 forecast model, enabling the supply of initial conditions in sigma coordinates and a higher resolution grid compatible with the R40 model, and assimilated satellite soundings over land at 50 and 70 hPa. Further vectorizing the analysis code enabled a new data selection procedure that chose up to 33 observations based on all 9 mass/wind correlation functions rather than choosing the closest 20 (Dey *et al*., 1987). DiMego (1988) discussed the NMC global assimilation in detail.

While the number of vertical levels remained at 18, MRF86 increased vertical resolution near the surface, with the lowest layer now 10 hPa thick rather than 54, and decreased resolution in mid-troposphere. It had 6 layers below 700 hPa and two in the stratosphere. The new physics introduced stability-dependent vertical diffusion of heat, momentum and moisture and a parameterization of shallow convection (Tiedtke, 1983), while making horizontal diffusion quasi-isobaric, extending deep convection throughout the troposphere, and correcting sea surface temperature (SST) where the land-sea table and spectral orography disagreed. Monthly varying fields of albedo and sea ice and spatially variable surface roughness dependent on vegetation were introduced.

The components of the MRF86 were evaluated individually before evaluation of the complete system. These changes reduced the day 5 root mean square (RMS) error of 500 hPa height by 10% and increased 1000 hPa height anomaly correlation by more than 10%. Stronger initialized analyses of tropical divergence represented the tropical convective centers more

accurately; forecasts reduced excessive upper level divergence over Mexico present in the MRF85. Throughout the history of the GFS, EMC followed the philosophy articulated by Han *et al*. (2017) in improving model physics:

1) Identify forecast biases or problems.
2) Reduce bias by updating physics with more realistic parameterizations.
3) Test the physics in operational mode.
4) Implement only if forecast skill is improved or at least neutral.

A better parameterization may interact negatively with other parameterizations or with the model dynamics or may be incompatible with observations and the assimilation, requiring the assessment of forecast skill to test its compatibility with the GFS.

F. GDAS, AVN, MRF unified Nov. 1986

In Nov. 1986 the AVN forecast system was upgraded from the R40L12 spectral model with limited physics to the MRF86, unifying the GDAS, AVN and MRF and significantly improving AVN 48-hour forecasts of 250 hPa NH winds and sea level pressure (Bonner, 1989).

G. DERF I and II

Beginning in 1985 NMC explored the use of the GFS for making dynamic extended range forecasts (DERF). In phase I, nine 30-day forecasts produced results encouraging enough to initiate phase II, in which the MRF86 generated 108 30-day forecasts from initial conditions 24 hours apart from 14 Dec. 1986 to 31 March 1987 (Tracton *et al*., 1989). While the 30-day

forecast almost always outperformed persistence, the first 7 to 10 days provided the best estimate of the 30-day mean. Beyond 10 days the average skill was low, but varied widely, pointing to a need to predict skill; the study explored ways of predicting skill.

4. 1987-1991: Improvements in resolution and physics

A. MRF87 T80 Aug. 13, 1987

A second Cyber 205 was acquired in 1987. The MRF87 increased horizontal resolution from R40L18 (300 km) to triangularly truncated T80L18 (150km) on Aug. 13, 1987, doubling the resolution in the east-west direction while increasing resolution in the north-south direction by 20%. Physical processes were now calculated on a Gaussian grid of 243 by 122, radiation on a coarser 162 by 81 grid. Prorating shortwave radiative fluxes calculated every 12 hours by the cosine of the solar zenith angle introduced a simple diurnal cycle. Longwave radiative fluxes were fixed for twelve hours, although downward longwave flux at the surface was empirically adjusted for the diurnal change in the lowest layer temperature. Other changes were made in radiative and deep convective parameterizations and in the time-marching algorithm for surface temperatures, vertical diffusion and surface fluxes, improving the tropical moisture profile and eliminating spurious oscillations in skin temperature (McPherson, 1988). Table 2 of Dey and Ballish (1988) listed the characteristics of the MRF87.

B. Gravity wave drag Sept. 3, 1987

Flow over large mountain barriers generates unresolved vertically propagating gravity waves that transfer momentum from low wind regions near the surface to high wind regions near

the tropopause where the waves break, dissipate and slow down winds. Due to their neglect of this process most numerical models of the time increased the zonal flow with forecast length and decreased the amplitudes of waves. The parameterization of gravity wave drag based on the work of Palmer *et al.* (1986), Pierrehumbert (1987) and Helfand *et al.* (1987) was developed and tested by Alpert *et al.* (1988) and implemented into the GFS Sept. 3, 1987. It reduced the strength of upper level jets in the NH, especially beyond 5 days, to levels more consistent with observations.

C. 18 Levels of Moisture Jan. 20, 1988

The number of levels containing moisture increased from 12 to 18 Jan. 20, 1988, reducing an accumulation of moisture in layer 12 in the MRF86 and a systematic cooling error.

D. Penman-Monteith evapotranspiration May 18, 1988

Starting May 18, 1988, the Penman-Monteith parameterization of evapotranspiration over land (Pan, 1990) used a surface energy balance to determine potential evaporation and introduced resistance due to vegetation, reducing excessive evaporation over land and reducing relative humidity in the lower atmosphere to more realistic levels (McPherson, 1988).

E. Physical TOVS retrievals Sept. 20, 1988

On Sep. 20, 1988, after careful testing by the DD and NESDIS, physical retrievals of TOVS temperature soundings based on temperature-radiance relations derived from inversion of the radiative transfer equation replaced statistical retrievals based on the comparison of radiance

data to collocated radiosonde reports; the change produced slightly better 5-day forecasts (Dey *et al.*, 1989). Kistler *et al.* (2001) found that the use of NESDIS retrieved temperature soundings from TOVS extended the length of useful forecasts by 3 days in the SH but had little impact in the NH

F. MRF88 Interactive clouds Nov. 30, 1988

Interactive clouds diagnosed as a function of forecast relative humidity, convection depth and rainfall rate every time the radiation was calculated replaced zonally averaged climatological clouds Nov. 30, 1988 in the MRF88 (Slingo, 1987; Campana *et al.*, 1989). Zonal mean radiative heating rates became more realistic, top of the atmosphere radiative fluxes improved significantly and a low to middle tropospheric cold bias decreased (Campana *et al.*, 1990).

Horizontal diffusion was significantly reduced for most wavelengths but increased for very short wavelengths in the upper quarter of the atmosphere, reducing the total horizontal dissipation of energy (Caplan and White, 1989). In the analysis, correcting errors in satellite profiles improved the analysis of heights and winds above 300 hPa; observation errors were adjusted and forecast errors reduced to reflect the increased skill of the GFS; increasing the limit for large isolated residuals (observations minus first guess) retained more SH observations (Deaven *et al.*, 1990).

G. Comprehensive hydrostatic quality control Dec. 14, 1988

A quality control system Dec. 14, 1988 helped to ensure the quality of assimilated rawinsonde observations and increased analysis accuracy. Initially it performed hydrostatic

checks and, when possible, corrected rawinsonde heights and temperatures on mandatory pressure levels (Gandin, 1988; Collins and Gandin, 1990); the system was extended to significant levels Apr. 11, 1990. Half of the errors detected by the system were corrected (Gandin *et al.*, 1993). An expanded system incorporated additional increment, horizontal and vertical checks Nov. 15, 1991 (Collins, 1998).

H. MRF89 Improved surface parameters Dec. 13, 1989

The MRF89 improved the specification of surface parameters Dec. 13, 1989. Soil wetness had been initialized from climatology but was now updated by model forecasts of precipitation and evaporation; forecasts also updated soil and two-meter temperatures. Snow depth was now a combination of observed snow cover (updated once a week) and model forecast; sea ice boundaries came from the initial analysis by the Joint Ice Center. The changes in surface parameters improved near-surface analysis fields and the overall forecasts, primarily in the 6 to 10-day forecast range (Kalnay *et al.*, 1990). A nonlinear initialization of analysis increments replaced the old initialization; the new system did not zero out the gravity modes, but removed imbalances created by the analysis and improved the tidal oscillation in the tropics (Ballish *et al.*, 1992). Kanamitsu (1989) described the MRF89 in detail.

I. More accurate long wave Feb. 1990

In Feb. 1990 a more accurate longwave parameterization developed by GFDL incorporated improvements based on an international comparison of radiation algorithms and

new laboratory absorption data for water vapor and carbon dioxide (Kalnay *et al.*, 1990; Schwarzkopf and Fels, 1985, 1991).

J. Higher resolution SST Feb. 20, 1991

On Feb. 20, 1991 a new SST analysis (Reynolds, 1991) used OI (Gandin, 1963) and increased horizontal resolution to 2^o from 5^o and temporal resolution to one week from two weeks (Reynolds, 1988).

K. Coupled model project

In early 1991 the Climate Analysis Center and DD began the Coupled Model Project to develop a multi-season forecast capability using a coupled ocean-atmosphere general circulation model (GCM) (Ji *et al.*, 1994a, b). The atmospheric component was a T40 version of the MRF with modified physics. In extended runs the GFS spread tropical convection too widely and did not focus convection in the observed regions, producing too weak surface stresses. Extensive adjustments of vertical mixing, convection and cloud-radiation interactions led to much more intense convection and stronger, more realistic surface stresses in the equatorial Pacific.

The ocean component was a Pacific basin ocean model, extending from 45S to 55N and from 120E to 70W, developed at GFDL by Bryan (1969), Cox (1984) and Philander *et al.* (1987). Ocean initial conditions came from a 10-year ocean reanalysis for the Pacific covering 1982-1992. Nine to twelve-month hindcast experiments for 1984 to 1992 outperformed persistence for the first six months (Ji *et al.*, 1994a, b). Two later versions of the seasonal forecast system with modified heat flux coupling and model output statistics (MOS) correction

of surface stress anomalies between the atmosphere and ocean outperformed the original seasonal forecast system; the seasonal forecast system showed useful skill in forecasting low frequency SST variability but had less skill for short-lived warmings than longer-lasting events (Ji *et al.*, 1996).

L. MRF91 T126, marine stratus, mean orography Mar. 6, 1991

The installation of the Cray YMP 8-processor enabled an increase in horizontal resolution in the MRF91 to T126 (105 km) Mar. 6, 1991. Mean orography replaced silhouette mountains because T80 experiments no longer showed a systematic difference in 5-day forecast skill. The new SSI analysis system, implemented June 25, 1991, used surface pressure observations over land; fewer surface pressure observations were rejected if mean orography was used.

A new marine stratus parameterization based on relative humidity searched for clouds in the lowest 10% of the atmosphere only if the point was over ocean, had a sufficiently strong inversion and was capped by dry air, resulting in realistic stratus clouds. The previous cloud parameterization did not search for clouds in the lowest 10% of the atmosphere due to excessive longwave cooling with low level cloud over elevated terrain (Kanamitsu *et al.*, 1991).

To reduce a loss of wave energy with time, horizontal diffusion was reduced in the medium scales by switching from a del-4 horizontal diffusion formulation to a parameterization developed by Leith (1971) based on turbulence theory. Mass conservation was improved and the moisture convergence threshold for convective precipitation increased by a factor of three to prevent an unrealistic increase in precipitation due to the resolution of smaller scales.

In five months of parallel tests the MRF89 (T80) and MRF91 (T126) showed very similar skill outside the tropics at 1000 and 500 hPa; the MRF91 proved superior at 100 hPa and reduced a cold bias in tropical temperatures. Operational forecasters found the T126 forecasts much more useful in producing meteorologically realistic forecast fields such as precipitation, especially for higher amounts.

M. New albedo, SiB May 21, 1991

A new albedo, stomatal resistance and roughness length based on a simple interactive biosphere model (Dorman and Sellers, 1989) were introduced along with a correction to the Kuo convective parameterization May 21, 1991.

5. 1991-1995: SSI analyses and SAS convection

A. SSI assimilation June 25, 1991

A major advance in data assimilation, on June 25, 1991 the 3-D variational (3-D var) SSI (Parrish and Derber, 1992) was the first operational variational assimilation system used in NWP. It used 100 iterations to analyze the spectral coefficients directly on model sigma surfaces using the fundamental equations and theory of OI (Gandin, 1963). The major differences of the SSI from the previously operational grid point OI were (Derber *et al.*, 1991):

-The analysis variables were closely related to the spectral model coefficients.

-Temperature observations were used instead of heights, enabling the use of aircraft temperature observations.

-Nonstandard observations could easily be included.

-All observations were used simultaneously; no data selection was required.

-The more realistic dynamic linear balance constraint between the wind and mass fields was applied directly and included parameterized friction.

-The need for model initialization was eliminated.

The resulting SSI increments were smaller and smoother than grid point OI increments but fitted the observations as well as the OI; SSI analyses were more consistent with the forecast model. Operational aviation and medium range forecasters found that better maintained upper tropospheric jets evolved more realistically. The SSI also improved forecasts of the phasing of extratropical systems, major changes in trough/ridge patterns and subsequent development and had fewer erroneous small-scale features, but a greater tendency to overdevelop features. 1000 and 500 hPa height anomaly correlations and RMS errors were substantially improved in both hemispheres.

B. MRF92 Improved radiation Feb. 1992

In Feb. 1992 the MRF92 called radiation every three hours rather than every twelve and used twelve bands of shortwave gas absorption rather than nine.

C. Tropical storm bogussing Aug. 18, 1992

The bogussing of tropical storms in the GFS Aug. 18, 1992 generalized work by Andersson and Hollingsworth (1988) and inserted the storm position, size and intensity into the GFS, specified the size of the circulation and had an automatic tracking algorithm (Lord, 1991).

D. GEFS Dec. 7, 1992

Both ECMWF and NMC implemented their first ensemble forecasting systems in Dec. 1992. The two goals of ensemble forecasting were to produce an ensemble average that outperformed individual forecasts beyond the first few days and to provide an estimate of the reliability of the forecasts. NMC's first global ensemble forecast system (GEFS) on Dec. 7, 1992, replaced the MRF forecast at 0000 UTC with an ensemble of four 12-day forecasts at 0000 UTC and extended the AVN forecast to 12 days at 1200 UTC (Tracton and Kalnay, 1993). Since the Cray YMP then in use was already saturated, the GFS resolution was truncated from T126 to T62 beyond day 6. Experiments showed that the truncation had little effect on skill.

The remaining three forecasts at 0000 UTC were T62 forecasts from a T62 data assimilation, one a control run and the other two from a pair of positive and negative perturbed initial conditions generated by the Breeding of Growing Modes (Toth and Kalnay, 1993). This method added a small arbitrary perturbation to the analysis, integrated the model for 6 hours from both an unperturbed and perturbed initial condition, subtracted the two, scaled down the difference to the same size as the initial perturbation, and repeated the process. After three or four days, the growth rate of the perturbations reached a saturation value comparable to the short-range forecast error growth. The initial perturbations contained leading Lyapunov vectors, the fastest naturally growing dynamical perturbations in the atmosphere, also present in the analysis errors. The three-day AVN forecast at 1200 UTC was at T126; the extension to day 12 was at T62; Lagged-Average Forecasting over two days of forecasts generated an ensemble of 14

members for days 1-10 (Tracton and Kalnay, 1993). Table 3 lists the different GEFS configurations and improvements from 1992 through 2015.

E. The Mar. 12-14, 1993 Superstorm

The impact of all the GFS implementations was demonstrated when the Mar. 12-14, 1993 "Superstorm" struck the east coast of the United States. The GFS produced accurate forecasts of the large-scale pressure, temperature and precipitation fields at least five days in advance; its consistency gave forecasters the confidence to issue severe weather forecasts over a very large area well in advance, despite its having trouble at short range with smaller-scale features connected with the timing and intensity of development in the Gulf of Mexico, the passage of a wind shift line and storm surge (Caplan, 1993). The 1993 superstorm increased public awareness of the increased reliability of forecasts. In 1993 commercial airlines kept flying until their planes were grounded; it took days to dig them out. When the Blizzard of 1996 (Jan. 6, 1996) was forecast, the airlines moved planes out of the storm's path (Kalnay *et al.*, 1998), saving time and money on restoring service.

F. SSM/I wind Mar. 30, 1993

Special sensor microwave/imager (SSM/I) near-surface wind speeds over water were first assimilated Mar. 30, 1993 using the algorithm of Goodberlet *et al.* (1989); wind direction came from the GFS first guess. The algorithm assumed a linear dependence of the wind speed on brightness temperature, an assumption valid only for low levels of moisture; as a result, only SSM/I wind speeds up to 16-18 m/s in clear-sky areas were assimilated, missing active weather

regions (Yu *et al.*, 1997). The new 10 m winds substantially improved GFS analyses and forecasts. Other changes to the analysis included making low-level divergence proportional to vorticity. In 1998 a neural network algorithm (Krasnopolsky *et al.*, 1995) capable of retrieving wind speeds up to 20-21 m/s (after bias correction up to 25-26 m/s) and including much more of the oceans and areas of high moisture and active weather replaced the Goodberlet algorithm.

G. SAS convection, L28 Aug. 11, 1993

Using mass and thermodynamic budgets to determine cumulus vertical mass flux and adjust atmospheric temperatures and moisture profiles, the simplified Arakawa-Schubert (SAS) convection parameterization based on Grell (1993) replaced Kuo convection Aug. 11, 1993 in the MRF93 (Pan and Wu, 1995). The Arakawa and Schubert (1974) parameterization, first implemented by Lord (1978), was more physically realistic than profile adjustment parameterizations, but was more complex and required the assumption of an ensemble of cloud types at each grid point. It used a quasi-equilibrium assumption as a closure and a saturated downdraft parameterization analogous to the updraft parameterization. Pan and Wu (1995) adopted a simplified parameterization using the deepest cloud (Grell, 1993) with entrainment of the updraft and detrainment of the downdraft (Hong and Pan, 1996). Concurrently, the vertical resolution of the model increased from L18 to L28, with a bottom layer 10hPa thick, 5 levels in the PBL and 7 levels above 100 hPa. GDAS still used NESDIS-generated retrievals of temperatures and moisture from satellite soundings of the atmosphere, but now used 40 levels of TOVS data rather than 15.

The new SAS convection produced more spatially coherent, realistic precipitation patterns, and better continental U. S. (CONUS) precipitation forecasts except for very high amounts. It improved tropical storm forecasts in the west and east Pacific and the Atlantic, the tropical precipitation climatology and the precipitation-SST relationship. It also reduced a tropical troposphere cold bias. The 500 hPa 5d height anomaly correlation was slightly worse, usually 1-2 percentage points.

H. Interactive TOVS retrievals Dec. 7, 1993

The SSI data assimilation and NESDIS retrievals relied on different background fields and error statistics. To reduce this problem, on Dec.7, 1993, interactive retrievals were introduced in the NH that used the analysis background (6 h forecast) as the retrieval background (Baker *et al.*, 1995). The new method did not account for the 3-D structure of background errors; difficulties could result due to correlated errors between retrievals and analysis background errors (Caplan *et al*., 1997).

I. GEFS 17 forecasts/day Mar. 1994

In Mar. 1994 the GEFS expanded to 17 forecasts per day out to 16 days, 10 perturbed, the MRF run, and a T62 control at 000 UTC and 4 perturbed runs and an AVN control at 1200 UTC, enabling outlooks for week 2 as well as days 6-10. Perturbations came from seven independent breeding cycles.

J. MRF95 SSM/I precipitable water Jan. 10, 1995

The MRF95 (Global Modeling Branch, 1994) first assimilated SSM/I satellite-derived total column precipitable water Jan. 10, 1995 and introduced a constraint on the divergence tendency and enhancements to the balance constraint, eliminated "super obs" (averages of partially redundant observations) except for conventional satellite soundings, and improved the background error including a reduction in the background temperature error (Parrish *et al.*, 1997), which also briefly described the analysis process of the time). The diagnostic cloud parameterization, surface physics and sea ice were improved (Caplan *et al.*, 1997); a new 2-layer soil model based on Mahrt and Pan (1984) and Pan and Mahrt (1987) developed at Oregon State University (OSU) was also implemented.

Campana *et al.* (1994) reviewed revisions to the diagnostic clouds. Regional cloud-relative humidity relationships were introduced (Mitchell and Hahn, 1989); stratiform clouds were allowed to exist above and below convective clouds; and radiative properties were determined from a parameterization of optical depth and liquid water path in each cloud layer (Harshvardhan *et al.*, 1989). The new clouds produced more realistic outgoing longwave radiation at the top of the atmosphere (Kalnay *et al.*, 1996).

K. NCEP Oct. 1995

In Oct. 1995 Congress approved the long-planned restructuring of the national centers of the NWS combining NMC, the National Severe Storms Forecast Center and the National

Hurricane Center (NHC) into NCEP with an expanded mission including climate and ocean prediction as well as weather and hydrologic prediction; DD became EMC (McPherson, 1994).

L. Assimilation of radiances and improved physics Oct. 25, 1995

The direct use of satellite radiances replaced NESDIS retrieved temperature and moisture soundings and interactive retrievals in the SSI analysis Oct. 25, 1995, producing large increases in forecast skill at all SH levels and for the first time throughout the NH troposphere (Caplan *et al.*, 1997). Previously satellite data had not had a significant impact on GFS skill in the NH (G. DiMego, personal communication). It typically took several months to over a year after the launch of a satellite for its data to be assimilated operationally in the GFS.

The radiances underwent substantial preprocessing by NESDIS before use. Improvements to surface skin temperature and an ozone analysis improved the simulation of clear radiances from model fields for assimilating observed radiances. A constraint on the divergence increment at very large scales was also introduced (Caplan *et al.*, 1997; Derber and Wu, 1998).

Betts *et al.* (1996) reviewed the MRF95 (implemented Jan. 10, 1995) physical parameterizations and compared its seasonal and diurnal cycles of the surface energy budget and PBL to observations at a field site in Kansas in the summer of 1987. The study suggested several improvements, some of which were implemented in Oct. 1995. The MRF95 version of the GFS was used in the NCEP/National Center for Atmospheric Research (NCAR) reanalysis. Nine months of parallel testing found improved fits of the 6-h forecasts of both temperatures and

winds to rawinsondes, better forecasts of NH geopotential height, and reduced RMS vector wind error (Parrish *et al.*, 1997).

In the Oct. 25, 1995 implementation, SAS convective parameterization made more complete use of the convective available potential energy (CAPE); the PBL was improved by a bulk Richardson number approach with a prescribed cubic profile of diffusivity coefficient joined to the surface layer to produce a matching flux; shortwave radiation was called every hour instead of every three; and a weak nudging of the deep soil moisture to climatology was removed.

The original Arakawa-Schubert parameterization was based on the tropical oceanic atmosphere which is almost always conditionally unstable; therefore, the parameterization adjusted the atmosphere to a conditionally unstable state. However, this assumption produced weaker convection than observed over subtropical and midlatitude land masses. The new version allowed the atmosphere to approach neutral stability and consume all the CAPE over land, producing more active convection over the U.S. (Caplan *et al.*, 1997), enhanced connective heating in disturbed areas and decreased convective heating in suppressed regimes.

In the previous PBL parameterization turbulent mixing occurred through vertical diffusion with diffusivity coefficients related to the local Richardson number. In reality vertical transports in the PBL are accomplished mostly by the largest eddies; such eddies should be modelled by the bulk properties of the PBL, not local properties. The new parameterization used a nonlocal vertical diffusion approach proposed by Troen and Mahrt (1986) and replaced a local Richardson number with a bulk Richardson number for the entire PBL. The nonlocal vertical

diffusion parameterization simulated the daytime PBL structure more realistically than the local parameterization (Hong and Pan, 1998), produced stronger mixing of moisture in the PBL, supplied more moisture to the free atmosphere and improved U.S. precipitation.

The deep convection and PBL changes contributed nearly equally to the enhancement of surface evaporation and deep convective heating, producing a more realistic distribution of moisture and precipitation near the intertropical convergence zone and more sharply defined tropical convective systems. Together with other Oct. 25, 1995 changes, they decreased upper level tropical wind errors and improved tropical storm track forecasts in the GFS and the GFDL hurricane model that used the global analysis in its initialization (Surgi *et al.* (1998), part of a collaborative effort begun in the mid-1990s by EMC and NHC to improve tropical storm forecasts).

M. ERS-1 winds Nov. 20, 1995

European Remote Sensing Satellite (ERS-1) scatterometer winds were first assimilated Nov. 20, 1995. Since the scatterometer provided multiple solutions for wind direction, a background field was required to assist in selecting the best direction. Based on work at UKMO, Peters *et al.* (1994) produced a system to quality control and reprocess the data at NCEP; using GFS surface winds as a background field in the NCEP processing system improved the selection of directions.

N. NCEP-NCAR Reanalysis

From the early days of NWP, analyses produced to initialize forecasts were used to study the observed atmosphere. Frequent changes improved the analyses and included more and more observations, but inserted discontinuities in the climate record (Trenberth and Olson, 1988). To eliminate such discontinuities, it was proposed that a frozen analysis/forecast system reanalyze decades of past observations to create a more continuous and consistent record of the observed atmosphere and that the system continue to run in real time as a climate data assimilation system (CDAS). Several centers, including the Center for Ocean-Land-Atmospheric Interactions, ECMWF, and the National Aeronautics and Space Administration (NASA) Goddard Laboratory for Atmosphere, carried out reanalyses.

NCEP and NCAR worked together to assemble all available data and to carry out a reanalysis with the MRF95 at T62L28 rather than T126L28 resolution. The NCEP/NCAR reanalysis eventually extended back to 1948 and is still running today, a 70-year record affected only by changes in observations (Kalnay *et al.*, 1996; Kistler *et al.*, 2001). The production and distribution of several CD-ROMs of reanalysis output helped make the reanalysis easily available to the global community. The reanalysis used NESDIS retrieved temperature retrievals rather than satellite radiances; later reanalyses that assimilated radiances directly found discontinuities due to changes in satellites. Before the reanalysis began one advisor commented "They will criticize it but they will use it" (Kalnay, personal communication). According to Google Scholar, as of Aug. 15, 2018 Kalnay *et al.* (1996) has been cited 26,507 times, Kistler *et al.* (2001) 4410 times.

Fig. 1 of Kistler *et al.* (2001) displayed how the number of observations has increased over the years. The SH had very few observations until the late 1960s; in 1979 a truly global observational network was established. As part of the reanalysis, 8d forecasts were made every five days throughout the reanalysis with the frozen T62 reanalysis model. Fig. 7 of Kistler *et al.* (2001) displayed the annually averaged anomaly correlation of 5d forecasts with the reanalysis model for 1948-1998, displaying the effect of changes in observations, and the operational scores for 5d forecasts with the evolving operational GFS for 1984-1998. Before 1958 scores were higher in the SH, reflecting the lack of observations and the forecast model's dominance of the analysis. The reanalysis model showed continued increases in SH skill until the late 1990s, reflecting an improving observational network, while in the NH little change was seen after 1975, reflecting the better, earlier established observational network. The operational model in Fig.7 of Kistler *et al.* (2001) showed the effect of improvements in the GFS, especially in the NH in the later 1980s.

O. NCEP R-2 Reanalysis

Kanamitsu *et al.* (2002b) produced a second, updated reanalysis, the NCEP-Department of Energy (DOE) Atmospheric Model Inter-Comparison Project (AMIP)-II Reanalysis (R-2), correcting errors in the NCEP/NCAR reanalysis and updating the physics. One motivation was to provide a better comparison to long model integrations as part of AMIP; the reanalysis was done in partnership with the Program for Climate Model Diagnostics and Inter-comparison and was run at the DOE National Energy Research Supercomputing Center.

Error fixes included correctly locating SH bogus data, correcting errors in snow cover analysis, ocean albedo and snowmelt, and correcting humidity diffusion to remove "valley snow". Physics improvements included many implemented in the GFS during 1992-1998; fixed fields were also improved. A smoothed orography reduced noisy Gibbs phenomena-like features in precipitation and sensible and latent heat near orography. A simple assimilation of rainfall over land surfaces dramatically improved soil moisture. Near surface land temperatures, the land surface hydrology budget, snow cover and radiation fluxes over the ocean also improved. The R-2 reanalysis extends from 1979 to the present. According to Google Scholar Kanamitsu *et al.* (2002b) has been cited 4389 times as of Aug. 15, 2018.

6. 1997-2000: Improving tropical precipitation, TOVS1b radiances, T170L42

A. Improving tropical precipitation Feb. 4, 1997

While the GFS steadily improved over the years, problems occasionally developed, sometimes from small imbalances that grew with time. One such problem developed with the assimilation of SSM/I precipitable water over the oceans since Jan. 1995. SSM/I precipitable water estimates were 10% higher than collocated values from the GFS first guess over the tropical North Pacific; the differences were greatest for large values of precipitable water. Precipitable water from radiosondes was 2% higher than the GFS, indicating a GFS dry bias, but also implying that SSM/I values were too high and clearly incompatible with the GFS climate. During the first 48 hours of the forecast, global precipitation decreased drastically, mostly over the tropical ocean, as the GFS quickly rained out the precipitable water incompatible with its climate.

Short range forecasts produced more rainfall over the tropical ocean and less over Africa and western South America than satellite estimates and a slightly older version of the GFS that did not use SSM/I precipitable water. Raining out the excess precipitable water over the tropical oceans changed the vertical circulation in the tropics, decreasing rainfall and soil moisture over the tropical continents. Removing a weak nudging of soil moisture to climatology from GDAS in Oct. 1995 permitted soil moisture to drift to extreme values, decreasing rainfall over tropical continents further. Experiments with not using or with bias correcting SSM/I precipitable water and resetting soil moisture to a new climatology from the Climate Prediction Center (CPC) showed considerable improvement (G. White, personal communication). Tropical rainfall and soil moisture were restored Feb. 4, 1997 by stopping the assimilation of SSM/I precipitable water and resetting soil moisture to climatological values.

B. AVN forecasts to 78h 4 times/day June 1997

Beginning in June 1997 the AVN after a data cutoff time of 2 hours and 45 minutes ran every 6 hours at 0000, 0600, 1200 and 1800 UTC; its forecasts extended to 78 h for use in tropical storm forecasting; previously, the AVN had run twice a day at 0000 and 1200 UTC to 72 h. Four-times-daily global forecasts were made by insertion of the "off-hour" AVN runs into the cycle. In spite of the paucity of upper-air data available at 0600 and 1800 UTC compared to the standard launch times of 0000 and 1200 UTC, the resulting forecasts still displayed only slightly less skill and were still superior to the 6-hour-longer forecasts initialized from the previous 0000 or 1200 UTC runs and verifying at the same time (Iredell and Caplan 1997).

C. Minor changes Nov. 5, 1997

A number of minor changes Nov. 5, 1997 (Wu *et al.*, 1997) included reduced observation errors in the SSI based on Stofflen *et al.* (1994), changes in TOVS radiance assimilation, inclusion of several additional data sources (including Profiler winds, more satellite winds and aircraft observations), soil moisture nudging towards a recent CPC climatology, and conservation of dry mass. Spurious "valley snow" in wintertime was eliminated by removing a partial correction to the conversion of horizontal diffusion of moisture from sigma to pressure surfaces that assumed a constant vertical gradient of specific humidity globally.

D. TOVS1b radiances Jan. 13, 1998

GDAS had assimilated TOVS cloud-cleared radiances since October 25, 1995, a clear improvement over using NESDIS retrieved temperature retrievals. Using the data further back in the NESDIS processing chain yielded better results. On Jan. 13, 1998, GDAS began using TOVS level-1b radiances that had not been preprocessed; Rabier (2005) discussed the advantages of using “raw” radiances. Tony McNally visiting from ECMWF worked with EMC personnel to develop algorithms for data pre-selection, cloud detection, bias correction, quality control and observation weights. TOVS level-1b radiances were available earlier and more reliably than the pre-processed cloud-cleared radiances; their assimilation had a large positive impact in the SH and above 250 hPa everywhere.

E. Increased vertical diffusion Jan. 14, 1998

Vertical diffusion was increased in the free atmosphere Jan. 14, 1998, reducing spurious tropical storms and producing smoother tropical circulations (McNally *et al.*, 1998; McNally *et al.*, 2000).

F. T170L42 June 15, 1998 The Implementation from Hell

An extensive package of changes June 15, 1998 increased the resolution to T170L42 in GDAS, AVN and the first 84 hours of the MRF forecast and changed model physics and data assimilation (Derber *et al.*, 1998). The shortest resolvable wave was reduced to just over 2 degrees from almost 3 degrees, roughly equivalent to increasing grid resolution to 80 km from 105 km. The model nonlinear dynamics was calculated on a 512 by 256 Gaussian grid; the analysis interpolation to observations and the model physics used the same 384 by 190 Gaussian grid as before. Both orography and diabatic physics terms were lightly filtered near the end of the spectrum (up to 40% at T170) to reduce the direct forcing of the shortest waves and preserve the shortest waves in a spectral model for dynamical forcing (Lander and Hoskins, 1997).

3 new levels were inserted in the stratosphere, 8 in the middle to upper troposphere, and 3 in the PBL. Near 250 hPa in the jet region vertical resolution was now less than 1 km to improve the accuracy and sharpness of jet wind forecasts for aviation; it had been greater than 1.5 km. Resolution in the stratosphere was still coarse but minimally adequate for prognostic ozone. Resolution in the GFS changed at 84 h to T126L28 and at 168 h to T62L28; a 3h digital

filter initialization reduced any imbalances from the change in resolution. Considerable work went into fitting the higher resolution model into the available time windows on the computer.

Upgrades to the land surface parameterization included new climatological data sets for vegetative cover based on satellite observations and soil and vegetation types, reduction in maximum canopy water, cessation of transpiration after dark, partial snow cover in a grid box, frozen soil, reduced snow depth estimate, increased roughness and changes in the thermal roughness length over land. SAS was changed to initiate convection earlier and to a CAPE elimination parameterization over the ocean to reduce precipitation bullseyes seen in the summer of 1997. Adding a dependence on wind speed to the evaporation of convective rain over the oceans (Fritsch and Chappell, 1980) reduced a GFS tendency to dry out convective regions, leading to an improved global humidity model climatology.

Gravity wave breaking in the lower troposphere was enhanced to improve the modeling of leeside mountain cyclogenesis and accompanying cold air outbreaks (Alpert *et al.*, 1996). The parameterization now considered orographic asymmetry, orographic convexity, and the direction of the wind impinging the face of the mountains, based on the work of Kim and Arakawa (1995).

A more accurate shortwave radiation parameterization (Chou, 1992; Chou and Lee, 1996) replaced an older parameterization (Lacis and Hansen, 1974). Including a multi-spectral band technique, an improved calculation in cloudy atmospheres, the addition of climatological aerosol effects and a new surface albedo, a function of 14 distinct surface types, it enhanced atmospheric shortwave radiative heating, resulting in a warmer troposphere especially in the tropics, and improved the radiation budget at the top of the atmosphere and at the surface. The

old model had adjusted radiative heating rates in regions of multi-layer cloud to the same value for all cloud layers, eliminating strong longwave cooling at the cloud top. This adjustment was now absent from the new shortwave radiation and the longwave radiation as well. The diagnostic cloud was re-tuned to more recent synoptic (not daily mean) Real-Time Nephanalysis (RTNEPH) data and modified to include cloud/humidity relationships for stratus over land as well as ocean, producing more realistic albedo, radiation and clouds.

Previously, the GFS had used a zonal and seasonal mean ozone climatology. More accurate ozone was required for long-range prediction and the prediction of the spatial variability of harmful ultra-violet radiation reaching the surface (UV index) and as first-guess fields for the direct assimilation of radiances. This implementation included ozone as a 3-D prognostic variable with stratospheric photochemical production and destruction parameterized following the NASA/GSFC approach (Rood *et al.*, 1991). All L42 forecasts used prognostic ozone, L28 forecasts used climatological ozone.

Increased horizontal resolution in the analysis reduced differences between the observed and modelled elevations, improving the resolution of local orographic effects and the fit between the background and the observations; increased vertical resolution improved the forward calculations for radiances. The previous analysis had had only the 6-h forecast available as the background field for observations; the new analysis had 3 and 9-h forecasts available as well to linearly interpolate to analysis time, reducing the largest interpolation to 1.5 hours, slightly improving the fit of observations to the background, especially for satellite radiances.

The new analysis made the minimization algorithm slightly non-linear; it included an external iteration, a loop around the entire analysis procedure (which involved calculation of observation increments, linearization, and 37 minimization iterations) which allowed weak nonlinearities in the observation and balance constraint. Experiments indicated that 5 external iterations would account well for nonlinearities, but a requirement not to delay delivery time with the new system permitted only two external iterations (Derber *et al.*, 1998). The external iterations significantly enhanced the inclusion of nonlinearities, making negative and supersaturated values of moisture much more unlikely in the analysis, and increased the weight given moisture channels for the satellite radiances.

The new analysis also included a new formulation of background error covariance allowing the specification of a spatial variance field, reformulation of the balance equation, definition of the spectral statistics as a function of total wave number and much greater flexibility. The formulation was similar to that documented in Derber and Bouttier (1999) and Boutier *et al.* (1997) and applied to the ECMWF assimilation. The new background error covariance improved tropical analyses and forecasts and better modelled the latitudinal variation of moisture. In this implementation the spatial variance varied only with latitude; in the future variance fields could vary with synoptic situation. In addition, a 3-D ozone analysis was combined with the temperature, moisture and wind analysis, incorporating Solar Backscatter Ultraviolet Radiometer (SBUV) data and no longer using NESDIS total ozone retrievals; the basic features of ozone analyses and forecasts were reasonable.

Radiances from NOAA-11 were added to radiances from two other polar orbiters, NOAA-12 and NOAA-14; less thinning was applied to the radiances. Radiative transfer calculations for the radiances also changed significantly, becoming more computationally efficient, more accurate and applicable to different satellites (McMillin *et al.*, 1995); surface emissivity calculations for infrared and micro-wave channels were also improved (Masuda *et al.*, 1988). These changes in radiative transfer calculations made it possible to use Geostationary Operational Environment Satellite GOES-8 and GOES-9 radiances over the oceans, improving moisture fields over the eastern Pacific and western Atlantic. The new analysis was also largely Y2K compliant.

The new GFS reduced a wet bias in precipitation over the United States and improved the precipitation threat score and other scores. It made better wind forecasts when verified against its own analyses but worse wind forecasts when verified against observations. The new system reduced a cold bias in much of the atmosphere and created warmer low-level temperatures over land, correcting a cold bias in the old system but creating a warm bias over land against observations.

G. Corrections to T170L42 June 24 and July 21, 1998

Following implementation, forecast quality declined, as assessed objectively by statistics and subjectively by forecasters. The analysis did not draw well for the data. An analysis change June 24, 1998 increased the rate of convergence of the iterations. Convective adjustment was added to the grid-scale precipitation code July 21, 1998 to remove instability not eliminated by convection and reduce the strength of spurious tropical vortices; horizontal diffusion of small

scales was substantially increased. To reduce a warm bias over land, shortwave absorption by aerosols was eliminated, surface albedo corrected and evaporation from vegetation increased (Derber *et al.*, 1998).

H. Reduction of resolution to T126L28 Oct. 5, 1998

Computer limitations did not allow the testing of potential corrections at full T170L42 resolution. Following a shortened testing period, the higher resolution was replaced with T126L28 to 168 h Oct. 5, 1998. In addition, the iterations in the analysis were increased from 75 to 200 and observational error variances corrected to draw more closely to the data. The physics was kept the same except for correction of an error in ice albedo and setting equal roughness lengths over land for heat and momentum. The changes produced better fits to observations and improved forecasts verified against own analyses. The T170L42 model had misplaced tropical storms and over-intensified them; the T126L28 system performed better (Caplan and Saha, 1998). The experience emphasized the importance of including analysis and forecast fits to observations in the evaluation of proposed changes; EMC made future implementations more rigorous.

I. AMSU-A and HIRS-3 satellite observations Mar. 8, 1999

Radiances from the Advanced Microwave Sounding Unit (AMSU)-A and HIRS-3 instruments on NOAA-15 were first assimilated Mar. 8, 1999. On NOAA-15 and following satellites the Advanced TIROS Operational Vertical Sounder (ATOVS) replaced the older TOVS suite of instruments used on NOAA-14 and older satellites. English *et al.* (2000) discussed the

positive effect of ATOVS retrievals on numerical weather forecasts. NOAA-15 had been launched on May 13, 1998; the AMSU-A replaced the MSU and SSU instruments.

J. The first MPP computer, asbestos and fire

A new IBM Scalable Performance Optimized With Enhanced Reduced Instruction Set Computer (RISC) (POWER) parallel 2 (SP2) computer capable of 700 billion floating point operations (Gflops) per second (DiMego, 2010), the first massively parallel processing (MPP) computer at NCEP, was installed in an old federal office building in Suitland, Maryland in 1999. After installation the power supply fell through the floor and asbestos was found under the raised tiles, requiring a new floor. Asbestos was then found above the ceiling tiles in the air movers and the IBM was moved to a more modern Census Bureau computer facility in Bowie, Maryland. On the day of the move, Sep. 27, 1999, a fire started in an internal power supply in the primary operational computer for generating weather analyses and forecasts, a Cray C9016 capable of 15.3 Gflops/second, and burned through the top of the computer; firefighters extinguished the fire by using dry chemicals that corroded other components of the computer. The computer could not be repaired (DiMego, 2010).

NCEP turned to two backup computers with 40% of the capacity of the Cray C9016, ran some forecasts on computers at other NWP centers, increased the use of forecasts from other centers and reduced the frequency and lengths of some model runs. GFS forecasts, normally run out to 16 days, were now terminated at 10 days. Some forecast runs finished later than before the fire.

In Nov. 1999 the IBM SP2 with 896 processors and 256 Gbytes of memory provided a five-fold increase in computer power over the Cray C9016. It was dedicated on Jan. 17, 2000; the director of the NWS commented that it would put the NWS closer to its "..goal of becoming America's no-surprise weather service".

K. T170L42 restored Jan. 24, 2000

On Jan. 24, 2000, T170L42 resolution was restored (Caplan and Pan, 2000)); no other changes were made. The T170L42 GFS made forecasts out to 7 days; the 6.5 d forecast was truncated to T62L28 and the reduced resolution GFS ran to 16 days. Truncating 12 hours before day 7 reduced convective precipitation spin-up. The analysis continued to use 200 iterations. Fits against observations showed improved wind forecasts for the NH and North America at T170L42, probably due to increased vertical resolution; little overall change was seen in precipitation forecasts. Tropical storms were stronger at T170L42, with an increase in false alarms, but somewhat better track forecasts; the use of T170L42 analyses to initialize the GFDL hurricane model produced better track forecasts.

L. Surprise snow Jan. 25, 2000

On Jan. 24, 2000 at 4 PM the NWS forecast less than an inch of snow for Washington, D.C, based on the GFS and other models; that evening satellite, radar and surface observations and new model forecasts showed the storm further west than had been forecast; at 10 PM the new NWS forecast predicted 4-8 inches for the DC area. On Jan. 25 up to 12-15 inches of snow fell in the DC area. The forecast failure was widely criticized in the media, including an editorial

cartoon by Pat Oliphant. Bosart (2003) reviewed the forecasts and suggested that forecasters had become too reliant on deterministic thinking and computer models and had disregarded satellite and radar observations. Experimental forecasts with a short-range ensemble predicted at least a 40% chance of significant snow on Jan. 25 24 hours in advance (Tracton, 2010).

7. 2000-2004: Cloud condensate, RRTM, mountain blocking

A. *Two-tiered seasonal prediction system Apr. 2000*

A second-generation dynamical seasonal prediction system (Kanamitsu *et al.*, 2002a) used a "two tiered" approach in Apr. 2000; SST anomalies over the tropical Pacific were predicted using a coupled ocean-atmosphere model, combined with observed SST anomalies outside the tropical Pacific damped during the forecast, and used to force an atmospheric model. The atmospheric model was run at T42L28 and later at T62L28 instead of T170L42 then used in the operational GFS and used a Relaxed Arakawa-Schubert (RAS) parameterization of convection (Moorthi and Suarez, 1992), the longwave radiation parameterization of Chou and Lee (1996), the cloud parameterization of Slingo (1987), and a smoothed mean orography. RAS convection allowed clouds with different tops but had no downdraft; RAS produced different tropical convection and vertical heating profiles and a more realistic Pacific-North American response than SAS convection in long 10-year atmospheric integrations. The longwave and cloudiness parameterizations and smoothed mean orography reduced a warm bias over the NH continents.

The new seasonal prediction system forecast both soil wetness and temperature to enhance summertime predictability and included forcing by snow and ice. Each month a 20-member forecast ensemble predicted anomalies from the model climatology, effectively correcting the forecast systematic error (Kanamitsu *et al.*, 2002a). The new system performed much better than the previous seasonal forecast system; predictions were better for years with larger tropical Pacific SST anomalies, but high skill scores did occur in years without large tropical anomalies. The new system also showed useful skill for near-surface temperature and precipitation over North America for 1997-2001, except during the NH summer.

B. Higher resolution GEFS, new tropical storm relocation May-July 2000

The new computer permitted the extension of AVN forecasts at both 000 and 1200 UTC to 126 h at full T170L42 resolution May 17, 2000, and an increase in resolution of GEFS forecasts from T62L18 to T126L28 out to 60 hours on June 27, 2000. A new relocation procedure to initialize hurricanes, implemented July 6, 2000 (Liu *et al.*, 2000), drawn from the GFDL hurricane model suite (Kurihara *et al.*, 1993, 1995) moved the hurricane vortex center in the guess field to the official NHC location and added a bogus vortex in the analysis if the guess field's vortex was too weak. Hurricane track forecasts in the GFDL hurricane model (which used GFS initial conditions), as well as in the MRF and AVN, improved dramatically.

C. *Minor changes Oct. 1, 2000*

Minor changes Oct. 1, 2000 included an improved ozone analysis, better ozone forecasts, a new observation error diagnosis, rawinsonde radiation corrections, inclusion of the

effects of balloon drift in space and time, higher resolution orography with filtering in the highest wave number to limit "ringing", and a reduced grid over the polar regions. A new albedo climatology, aerosols in the shortwave radiation and a refined single scattering albedo parameter reduced a summertime continental near-surface warm bias.

In Nov. 2000 the IBM SP2 was upgraded to 2048 processors and 1024 Gb of memory, a further six-fold increase in computer power. The GFS model operational on Jan. 9, 2001 is documented at http://www.emc.ncep.noaa.gov/gmb/mrf.html. Thermal roughness over the ocean was based on a formulation derived from the Tropical Ocean Global Atmospheres /Coupled Ocean Atmosphere Response Experiment of 1992/1993 (Zeng *et al.*, 1998).

D. Cloud condensate May 21, 2001

On May 15, 2001, cloud condensate was added to the GFS as a prognostic variable and used in the radiative transfer calculation; cumulus momentum mixing was also introduced (Moorthi *et al.*, 2001). The evolution of cloud condensate was determined by 3-D advection, a convective source from cloud top detrainment in the convective parameterization, a large-scale condensation source based on Zhao and Carr (1997) and Sundqvist *et l* (1989), sinks due to precipitation and evaporation, and 3-D diffusion. The fractional area of the grid box covered by cloud was diagnosed following Xu and Randall (1996). The new parameterization had only one type of cloud cover in each box: primarily convective anvils in the tropics and grid-scale condensation clouds in the extra-tropics. Clouds in all layers were assumed to be randomly overlapped. In the old GFS tropical cloud top detrainment had been unrealistically confined to a

few layers; the convective cloud top was now randomly chosen to be a level between the level of neutral buoyancy and that of minimum moist static energy.

The new shortwave radiation calculated optical thickness from the extinction coefficient and predicted cloud condensate path (Slingo, 1989; Chou *et al.*, 1998; Kiehl *et al.*, 1998; Hou *et al.*, 2002); longwave radiation calculated cloud emissivity from the predicted cloud condensate (Kiehl *et al.*, 1998; Stephens, 1984). The fit of the GDAS 6h forecast to mid- and upper-atmospheric moisture satellite channels improved (Treadon *et al.*, 2003), as were rain rates in the initialization compared to satellite microwave estimates (Kuligowski *et al.*, 2002).

The GFS of the time intensified initially weak tropical disturbances into stronger storms too often leading to false alarms; convection did not alter momentum in the GFS while in nature the growth of storms is limited by the interaction between convection and vertical wind shear. This implementation introduced the transport of momentum by mass fluxes in convective updrafts and downdrafts. A new United States Geological Survey (USGS) orographic dataset was used to define model orography. Microwave radiances from NOAA-16 AMSU-A/B and NOAA-15 AMSU-B, an additional quality control test for AMSU-A radiances, improved rawinsonde data quality control and a refinement of the tropical storm relocation algorithm were also implemented along with modification of the saturation vapor pressure calculation and changes to the moisture constraint.

The major components of the package were tested separately and together from Oct. 2000 to May 2001 in real time and in retrospective forecasts from May to mid-June 1999 and over the 2000 hurricane season; the Hydrometeorological Prediction Center (HPC) including its

International Desks and NHC evaluated the forecasts. Many false alarms in tropical storms were eliminated, cut-off circulations were reduced in intensity and a near-surface warm bias was replaced by a cold bias. 500 hPa anomaly correlations improved in the extra-tropics and RMS wind errors at 200 and 850 hPa decreased in the tropics. A low-level cold bias replaced a weaker warm bias in the NH; the skill of precipitation forecasts over the continental U. S. (CONUS) was very similar but the new model had excessive amounts of light precipitation (Treadon *et al.*, 2003).

E. Assimilation of satellite rain estimates 2001

The assimilation of rain rate estimates from SSM/I began Feb. 13, 2001; estimates from the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) followed Oct. 9, 2001. The impact on assimilation and forecast fields was small compared to the impact of prognostic cloud condensate and cumulus momentum mixing, although excessive precipitation was reduced. The non-linear nature of precipitation physics was difficult to include in the assimilation of satellite estimates of rain rate; rain rate assimilation was also complicated by the lack of error statistics (Treadon *et al.*, 2003; Kuligowski *et al.*, 2002).

F. GFS 4 times/day to 384 h 2002

Beginning Mar. 5, 2002, the AVN ran 4 times a day with a data cutoff of 2 hours and 45 minutes after synoptic time to 384 hours, at T170L42 to 180 h and thereafter at T62L28, replacing on Apr. 23, the MRF forecast at 000 UTC, and in the fall of 2002 was renamed the GFS. The final (FNL) analysis was run at least 6 hours after synoptic time to include data

arriving after the GFS data cutoff; it made a global analysis and 3, 6 and 9 h forecasts to serve as first guess fields for the next GFS analysis 4 times a day.

G. T254L64 Oct. 29, 2002

On Oct. 29, 2002 the GFS ran at T254L64 (55 km) to 84 hours, at T170L42 to 180h and at T62L28 levels to 384 h (Iredell *et al.*, 2002). Table 4 compares the differing vertical structure in the three resolutions. Sixth order horizontal diffusion was modified for the T254L64 part of the forecast. The analysis changes included an updated background error, increased use of satellite data, assimilation of Meteorological Terminal Aviation Routine Weather Report (METAR) surface pressure observations and termination of the divergence tendency constraint.

Five NCEP centers and the Meteorological Development Laboratory (MDL) evaluated the new model; HPC found small improvements in mass fields and precipitation, its South American/Tropical Desk found precipitation "bombs" in cases of weak synoptic forcing in the new system, the Marine Prediction Center preferred the deeper cyclones in the new system, NHC found larger track errors in the Atlantic beyond 72 h in the limited sample available, the Aviation Weather Center (AWC) found improvements in 1d forecasts of clouds, precipitation and icing, and the Storm Prediction Center and MDL found little difference between the old and the new GFS. The new GFS had higher 500 hPa height anomaly correlations out to 7-8 d, a reduced cold bias over land in the cold season, and higher threat scores for heavy amounts of precipitation (over 1.5 inches) over the U.S. (Iredell *et al.*, 2002).

H. Summaries of the GFS

Hou *et al.* (2002) described in detail the parameterization of solar radiation in the GFS as of Oct. 2002, including aerosols, cloud optical properties and albedo. A summary of the GFS atmospheric model as of Nov. 2003 was presented in Environmental Modeling Center (2003) and the SSI analysis system as of Apr. 2004 was summarized in Global Climate and Weather Modeling Center (2004), with a list of observations used in the GFS. Derber *et al.* (2003) discussed enhancements to the use of radiance data, including updates to radiative transfer, modifications to data selection and quality control and enhancements to data assimilation and forecast systems, implemented by Nov. 20, 2003.

I. RRTM Aug. 28, 2003

On Aug. 28, 2003, the Rapid Radiative Transfer Model (RRTM) for longwave radiation developed at Atmospheric and Environmental Research, Inc. (AER) by Mlawer *et al.* (1997) replaced the longwave parameterization developed by Schwarzkopf and Fels (1991). Highly accurate and very computationally efficient, RRTM ran in half the time of the old longwave parameterization in the L64 GFS. Besides carbon dioxide, water vapor and ozone, RRTM contained nitrous oxide, methane and halocarbons; for water vapor continuum absorption calculations RRTM used a model by Clough *et al.* (1992). Clear-sky cooling rates for typical atmospheric profiles were very similar for the two longwave parameterizations except in the upper stratosphere where RRTM produced more cooling; RRTM improved tropospheric temperature bias but enhanced a stratospheric cool bias. In 5y climate runs RRTM produced a more accurate top of the atmosphere outgoing longwave radiation.

J. Minor changes Nov. 20, 2003

On Nov. 20, 2003, a package of minor analysis changes included upgrades in radiance assimilation, time interpolation of the guess and assimilation of GOES-12 sounder radiances, Next Generation Radar (NEXRAD) radial winds and lidar line of sight winds.

K. Mountain blocking Feb. 24, 2004

An enhanced parameterization of sub-grid scale orographic forcing similar to Lott and Miller (1997) introduced mountain blocking Feb. 24, 2004 (Alpert, 2004). Wind flow around sub-grid scale orography is in contact with orography longer and therefore encounters more friction. The parameterization defined a dividing streamline below which flow goes around the mountain and above which flow goes over the mountain and can form vertically propagating gravity waves. The level of the dividing streamline was found by comparing the potential and kinetic energies of upstream large-scale wind and sub-grid scale air parcel movements as defined by the wind and stability. Mountain blocking led to improved NH skill scores and a reduction in very poor 5-day forecasts (Alpert, 2004). Alpert (2013) reviewed the parameterization of gravity wave drag and mountain blocking.

L. CFS Aug. 2004

In Aug. 2004 the Climate Forecast System (CFS), a fully coupled ocean-land-atmosphere dynamical seasonal prediction system, replaced the two-tier seasonal prediction system of Kanamitsu *et al.* (2002a). It featured full daily coupling with no flux correction of the ocean and atmosphere between 50N and 65S and relied on observed climatological SST poleward of 74S

and 64N, with a weighted average of SST from climatology and from the ocean component of the CFS in between. 15 retrospective 9 month forecasts per calendar month over 1981-2004 provided the basis for bias correction and a priori estimates of the usefulness of the operational seasonal forecasts (Saha *et al.*, 2006).

The CFS used a T62L64 (200 km) version of the GFS operational as of Feb. 2003 for the atmosphere, the GFDL Modular Ocean Model Version 3 (MOM3) (Pacanowski and Griffies, 1998) for the ocean, and the OSU 2-layer model of soil hydrology for land (Mahrt and Pan, 1984). Initial conditions for the atmosphere and for heat and buoyancy fluxes and wind stress came from the NCEP-2 reanalysis (Kanamitsu *et al.*, 2002b); oceanic initial conditions came from the Global Ocean Data Assimilation System (GODAS) (Behringer and Xue, 2004; Behringer, 2007). The CFS demonstrated comparable, complementary skill in seasonal forecasts to statistical methods used by CPC and superior skill to the "two-tier" seasonal prediction system it replaced (Saha *et al.*, 2006).

M. NAEFS 2004

In 2004 NWS joined with the Meteorological Service of Canada and the National Meteorological Service of Mexico to produce the North American Ensemble Forecast System (NAEFS), the combination of the GEFS and the Canadian ensemble, 40 16d forecasts twice a day (Zhu and Toth, 2006).

8. 2005-2009: Noah, GSI, GPS RO

A. T384L64

On May 31, 2005, the horizontal resolution of the GFS increased to T382L64 (35 km) from T254L64 (50 km). T382L64 forecasts extended to 180 h, followed by T190L64 (70 km) from 180 to 384 h.

The addition of Atmospheric Infrared Sounder (AIRS) and AMSU-A data from the polar-orbiting satellite Aqua launched in 2002 substantially increased the number of radiances assimilated; the intelligent thinning algorithm used on radiance data was altered to group data based on sensor type to spread data use more evenly (Campana and Caplan, 2005). Every six hours 200 million AIRS radiances were input to the analysis system, 2.1 million were selected for possible use, and 850,000 radiances free of cloud effects were used in the analysis; experiments showed AIRS significantly improved forecast skill in both hemispheres (LeMarshall *et al.*, 2006). A new cloudiness algorithm enhanced the quality control of infrared radiances and a new microwave emissivity model over snow and ice increased the use of microwave data over snow and ice. Rabier (2005) reviewed developments in data assimilation at numerical weather centers around the world up to 2005.

The introduction of the Noah Land Surface Model (LSM) (Chen *et al.*, 1996; Koren *et al.*, 1999; Ek *et al.*, 2003) with 4 sub-surface layers replaced the OSU LSM of the mid-1990s with 2 sub-surface layers. The Noah LSM was developed by the EMC land team and many collaborators; it contained improved treatment of frozen soil, ground heat flux, and energy/water balance at the surface, reformulated infiltration and runoff and an upgraded vegetation fraction.

The Noah LSM cycled continuously on itself in GDAS, initializing soil temperature and moisture; forecast land-surface forcing, including model precipitation, updated the LSM every time step. A 60d relaxation nudging of soil moisture towards an external global moisture monthly climatology constrained the LSM. The new LSM reduced a tendency for overly rapid snow-pack depletion in winter (Campana and Caplan, 2005).

A new 3-layer (two for ice and one for snow) sea-ice model based on Winton (2000) predicted sea-ice/snow thickness, surface temperature and ice temperature structure. Heat and moisture fluxes and albedo were treated separately for ice and open water in each grid box (Wu *et al.*, 1997).

Results from a number of 1d experiments with the T382 GFS showed that enhancing orographic heights by 10% of the mountain variance during the calculation of mountain blocking dissipative forces improved results, as did reducing both the background diffusion in the free atmosphere and the turbulent diffusion length scale from 150 to 30 meters in stable cases. Reducing diffusion did not eliminate a weak bias in the Asian jet, but did improve the stratosphere, making it warmer and drier.

All NCEP centers participated in the evaluation of real-time parallel forecasts for Apr.-May 2005 and retrospective forecasts for Aug-Sep. 2004 and Dec. 2004-Feb. 2005. Verification against analyses showed general improvements in a number of forecast products. The Ocean Prediction Center (OPC) found that the new GFS produced more realistically consolidated developing storms; HPC found improved mass fields, but less run-to-run continuity in the medium range. NHC found a significant improvement in Atlantic tropical cyclone tracks, but

little skill in the East Pacific in either GFS; AWC found improved jet stream guidance and greatly improved turbulence guidance.

CONUS precipitation scores showed clear improvement in winter, but were somewhat worse in summer, reflecting a larger precipitation bias due to too much summertime surface evaporation in regions of non-sparse vegetation cover in the Noah LSM. Changes to canopy resistance properties implemented on June 14, 2005 reduced the evaporation and the precipitation bias. New GFS temperature analyses above 5 hPa were not very good after implementation, reflecting problems in the SSI analysis handling over 100 AIRS channels in the two topmost model layers until sub-layers were added on July 7, 2005 (Campana and Caplan, 2005).

B. GEFS tropical storm relocation Aug. 25, 2005

On Aug. 25, 2005, tropical storm relocation was added to the GEFS. If a tropical storm was present in the initial conditions, tropical storm relocation was applied to each ensemble member to adjust the initial central location to the observed (Liu *et al.*, 2006). A hurricane perturbation was added to the ensemble perturbation in each member. T126L28 resolution was extended to 16d with 10 ensemble members 4 times a day. The new system was tested over 38 days and produced better hurricane track forecasts than the GFS and the old GEFS.

C. GEFS Ensemble Transform method May 30, 2006

On May 30, 2006, the method used in the GEFS to generate initial perturbations was updated to the Ensemble Transform method with Rescaling (ETR) (Wei and al., 2008). A

generalization of the breeding method used previously, ETR was consistent with the data assimilation unlike the previous method and used analysis error covariances to restrain the initial perturbations. ETR improved most probabilistic scores; its perturbations explained more of the forecast error. The number of members every 6 h increased to 14 and, on Mar. 27, 2007, to 20.

D. Minor changes Aug. 22, 2006

Minor changes on Aug. 22, 2006 to address problems pointed out by GFS users included a new orography upgrading Antarctic orography and land sea mask identifying shelf ice as land, an improved snow analysis, particularly around the Alps, and land property quality control. An update of the ozone production and destruction terms from the Naval Research Laboratory (NRL) removed a low tropical bias in total ozone by using monthly mean photochemical coefficients (McCormack *et al.*, 2006), a new land model algorithm dealt with polar ice shelves and an error in surface downward longwave radiation was corrected. Time filtering increased, achieving greater stability without affecting the forecast, and the code infrastructure was upgraded. These changes had negligible impact on hemispheric scores, CONUS precipitation and tropical storms and were approved by four NCEP centers.

E. GSI May 1, 2007

On May 1, 2007, the Grid-point Statistical Interpolation (GSI) replaced the SSI (Kleist *et al.*, 2009b). Formulating the analysis in grid point rather than spectral space permitted the development of inhomogeneous and anisotropic background error covariances that could more accurately represent terrain and weather features, preparing the GFS for future advances in data

assimilation; it also simplified the unification of global and regional assimilations (Wu *et al.*, 2002). An important component was an improved method of incremental balance in the analysis, the tangent-linear normal-mode constraint (TLNMC) (Kleist *et al.*, 2009a).

Other new features in the GSI included changes in observation selection, improved quality control of radiances, a better minimization algorithm and updated radiative transfer for the forward model. The Community Radiative Transfer Model (CRTM) developed by the Joint Center for Satellite Data Assimilation (JCSDA) simulated satellite microwave (MW) and infrared (IR) radiances and computed radiance derivatives (Jacobians) with respect to the state variables (Han *et al.*, 2006).

Global positioning system (GPS) radio occultation (RO) data from the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) mission and Microwave Humidity Sounder (MHS) brightness temperatures from NOAA-18 were also added to the assimilation. Launched in Apr. 2006 six small satellites with GPS RO receivers in their final orbit provided 2500-3000 RO soundings/day evenly distributed around the globe (Cucurull *et al.*, 2007). The GSI could assimilate refractivity and bending angle from GPS RO.

A hybrid sigma-pressure vertical coordinate replaced the sigma coordinate; it transitioned from terrain-following sigma surfaces in the lower troposphere to constant pressure surfaces in the stratosphere to improve the midlatitude jet stream (Sela, 2009; Juang, 2005) and other upper tropospheric fields. A modularized radiation package with a few upgrades was also introduced.

The new GSI code proved more computationally efficient than the old, requiring two outer loops of 100 iterations each rather than the SSI's two loops of 150 iterations each. The new model was tested retrospectively over two hurricane seasons and one NH winter as well as in real time for two months. Forecast skill in the tropics improved and short-term forecast errors in the extra-tropics were substantially reduced. GFS forecasts and GFDL forecasts initialized by the GFS of tropical cyclones both showed improvement in the Atlantic, but results were more mixed in the east Pacific (Kleist *et al.*, 2009b); 200 hPa winds over East Asia substantially improved.

F. Minor implementations 2007

A minor implementation May 26, 2007 improved quality control and observation errors for satellite-derived winds and introduced HIRS, MSU and AMSU-A brightness temperatures from the European Organization for the Exploitation of Meteorological Satellites' Meteorological Operation Satellite A (METOP-A) and higher resolution non-averaged observations from GEOS (Kleist *et al.*, 2009b). An upgrade Sept. 25, 2007 unified the postprocessors of the GFS and the North American Mesoscale (NAM) forecast system and changed the precipitation type algorithm. Problems reported with this implementation, including interpolation problems and spurious high pressure over mountainous coasts, were corrected Oct. 10, 2007. On Dec. 4, 2007, the GSI was updated and modified to assimilate SBUV-8 and JMA wind data. Postprocessing of the GEFS and NAEFS were upgraded.

G. Cancelled implementation 2008

A proposed 2008 implementation was cancelled after a limited test period. Statistical evaluation showed little difference between the old and new GFS; two NCEP centers found evidence of forecast deterioration. The components of the proposed implementation became operational in later implementations.

H. Assimilation changes Feb. 24, 2009

On Feb. 24, 2009, Infrared Atmospheric Sounding Interferometer (IASI) observations from the European Space Agency (ESA)'s METOP series of polar orbiting satellites were first assimilated. Other changes included a flow dependent reweighting of background error variances and a new version of CRTM.

I. GFS changes Dec. 15, 2009

The GFS software structure was upgraded, low cloud was redefined to include PBL cloud and the GFS and GDAS postprocessors were unified and upgraded Dec. 15, 2009. Tropical cyclone advisory minimum sea level pressure observations were assimilated to address forecaster concerns that some tropical systems were much too weak in GFS analyses; the initial intensity bias of tropical systems was substantially reduced and track and intensity forecasts improved out to five days (Kleist, 2011). Several new observations such as NOAA 19 HIRS/4 (with a field of view of 10 km rather than 20 km in HIRS/3), AMSU-A brightness temperatures and NOAA-18 SBUV/2vs and Aura Ozone Monitoring Instrument (OMI) ozone observations

were also assimilated. Improving quality control, observation error structure and forward modeling improved GPS RO profile assimilation (Cucurull, 2010).

9. 2010-2014: 23 km, Hybrid 3D EnVar, Sandy

A. GEFS STTP Feb. 23, 2010

A Stochastic Total Tendency Perturbation (STTP) parameterization was implemented in the GEFS Feb. 23, 2010 to represent combined dynamics/physics uncertainties in the forecast model, based on the hypothesis that ensemble perturbation tendencies are a representative sample of the random total model errors (Hou *et al.*, 2006, 2008, 2012). Stochastic forcing terms were formed by randomly combining the total tendencies of ensemble perturbations and rescaling them to appropriate sizes. STTP significantly increased ensemble spread, reduced outliers and systematic errors and improved ensemble-based probabilistic forecasts and the ensemble forecast distribution. The horizontal resolution of the ensemble forecasts increased to T190.

B. T574 July 28, 2010

A major upgrade in GFS resolution, physics and postprocessing July 28, 2010 increased resolution out to 192h to T574L64 (23 km); beyond forecast hour 192h, resolution remained at T190L64 (70 km). The GFS had two long-standing problems: too little stratus, particularly near the coast in the eastern subtropical oceans, and too heavy non-convective grid scale precipitation. To address these problems, a mass flux parameterization of shallow convection replaced a turbulent diffusion-based approach; stronger and deeper cumulus convection depleted more

instability in the atmospheric column and suppressed excessive grid-scale precipitation (Han and Pan, 2011).

The revised PBL model enhanced turbulent diffusion in the stratocumulus areas. The old shallow convection had assumed equal areas of updraft and downdraft and produced cooling above and heating below in the lower atmosphere; the new assumed the environment was dominated by subsidence producing environmental warming and drying and produced heating throughout the shallow convection layers. The revised deep convection had a larger cloud-base mass flux, higher cloud tops, and detrainment from all levels. To avoid excessive erosion of stratus along the coast, background diffusivity of heat and moisture in the lower inversion layers was reduced to 30% of the surface value.

Besides producing a much more realistic distribution of stratus and reducing the bias in non-convective grid scale precipitation, the new physics improved 500 hPa heights, vector winds, CONUS precipitation and hurricane track and intensity forecasts. To reduce wind forecast errors, background diffusivity for momentum was substantially increased everywhere. As Figs. 8a and 8b show subsequent to 2010, Root Mean Square (RMS) vector wind error decreased substantially, but wind speeds decreased slightly, increasing a negative bias against observations; after the implementation, a more negative wind bias resulted in weaker jet stream maxima, too weak a quasi-biennial oscillation (QBO) and weaker easterlies in the tropical stratosphere and too weak trade wind bursts in the Caribbean, where the previous GFS had produced more realistic trade wind bursts before. Trade wind bursts are an important aviation forecast concern in the Caribbean.

A new parameterization of shortwave radiation from AER, RRTM Version 2 (RRTM2) (Clough *et al.*, 2005) contained 14 spectral bands, a fast two-stream radiative transfer parameterization and many absorbing gases. The frequency of computing longwave radiation was increased from every 3h to 1h; changes were made to aerosols; time varying global mean rather than constant carbon dioxide was introduced. Shortwave cloud overlap was changed from random to maximum random overlap. A new treatment of the dependence of albedo on solar zenith angle based on Atmospheric Radiation Measurement Surface Radiation Budget Network (ARM-SURFRAD) observations was introduced (Yang *et al.*, 2008), as was a positive-definite tracer transport in the vertical to remove negative water vapor (Yang, 2009).

Compared to the previous T382L64 GFS, the new T574L64 GFS used four times stronger mountain blocking and half the strength of gravity wave drag; the best tuning of orographic forcing was determined by a large number of tests of different forcing strengths. The new gravity wave drag parameterization included an automatic scaling of sub-grid-scale orographic forcing with resolution. Changes were also made to hurricane relocation.

C. CFSv2 and CFSR Mar. 2011

A new seasonal forecast system (the second version of the NCEP CFS (CFSv2)) based on the GFS was implemented in Mar. 2011, although the previous seasonal forecast system, CFSv1, continued to run in operations until late Sept. 2012 to give users a chance to adapt to CFSv2 (Saha *et al.*, 2014). An essential component of CFSv2 was a comprehensive reforecast over 29 years (1982-2010) to obtain consistent and stable calibrations and skill estimates for operational real time sub-seasonal and seasonal CFSv2 predictions. The initial conditions for the reforecast

were generated by a coupled reanalysis (CFSR) over a 32-year period (1979-2010) (Saha *et al.*, 2010) with a system very close to that used in CFSv2 and very similar to the 2010 GFS.

Saha *et al.* (2010) (according to Google Scholar cited 2830 times as of Aug.15, 2008) discussed the advances in the atmospheric and oceanic models from CFSv1 and the NCEP/NCAR reanalysis to CFSv2 and CFS Reanalysis (CFSR) and reviewed the data used in CFSR. The resolution of the atmospheric system used in the NCEP/NCAR reanalysis had been T62 28 levels (200 km), the resolution of CFSR was T382L64 (38 km). The MOM4 ocean model used had roughly double the resolution of the MOM3 ocean model used in CFSv1 and was fully global (Griffies *et al.*, 2004). A 3-layer sea ice model (2 layers for ice and 1 for snow) originally from GFDL (Wu and Grumbine, 2009) and the Noah 4-layer LSM (Ek *et al*., 2003) were also used. Atmospheric, ocean and sea ice analyses were made every 6 hours, using the same 9h first guess forecast.

The three models ran independently but the atmospheric and sea ice models exchanged data every 3 minutes of model integration and the ocean, sea ice and atmospheric models exchanged data every 30 minutes. The land analysis used observed precipitation and ran daily. The CFSR assimilated satellite radiances directly rather than the NESDIS retrieved temperature soundings assimilated in the NCEP/NCAR reanalysis. The NCEP/NCAR reanalysis had used a single value of carbon dioxide for the entire period of analysis; the CFSR used observed monthly mean carbon dioxide.

A 120-h forecast was generated once a day from the CFSR, using the CFSv2 prediction resolution of T126L64, over the period 1979-2008; Fig. 20 of Saha *et al*. (2010) compared the

skill of these 500 hPa height forecasts to other versions of the GFS. Saha *et al*. (2010) discussed possible sources for improvements in skill. NH forecasts with a T382 assimilation system proved considerably more skillful than forecasts with a T62 assimilation system; increasing forecast model resolution from T126 to T384 may have resulted in a slight improvement in skill. A large increase in SH skill appeared to have come from the direct assimilation of radiances; another substantial increase seemed to reflect the replacement of TOVS by ATOVS, especially when assimilated by a higher resolution assimilation system. By 2009 SH scores were nearly comparable to NH scores, demonstrating that medium-range prediction of synoptic scale features was now nearly comparable over oceans and continents.

Compared to the NCEP-1 and NCEP-2 reanalyses, CFSR was considerably more accurate due to higher resolution in space and time and several years of improvements in observations, assimilation and modeling, demonstrated by the much higher skill of forecasts made from CFSR analyses. CFSR was also much more comprehensive, producing analyses of the ocean and sea ice as well as the atmosphere and land surface. Partly due to including increases in carbon dioxide, CFSR had a much more realistic climate trend in 2-meter (2-m) temperatures over land than NCEP-1 and NCEP-2 reanalyses.

The enhanced vertical structure of CFSR allowed it to better resolve the upper atmosphere, but limitations in the number and quality of radiosondes in the upper atmosphere prior to 1998 limited its ability to portray the quasi-biennial oscillation and semi-annual oscillation. Problems in bias-correcting satellite observations introduced false trends in the upper atmosphere; the use of six parallel streams to produce the CFSR inserted discontinuities in

the upper atmosphere. The CFSR had a much better relationship between tropical SST and precipitation than earlier reanalyses and better conservation of dry mass, but had problems with the atmospheric water balance, which became more imbalanced in 1998, possible due to new data sources such as AMSU. As a result of such problems in the CFSR, CPC required the continuation of NCEP-1 and NCEP-2 reanalyses. CFSR continues to run in near real-time, but its implementation of new satellite observations has lagged the GFS over the past few years.

Saha *et al*. (2014) reviewed minor differences between the system used in the CFSR, the system used in 29 years of retrospective forecasts (1982-2010) to obtain consistent and stable calibrations and skill estimates for operational seasonal and sub-seasonal forecasts and the system used operationally beginning Mar. 2011 to produce real time forecasts. The CFSv2 increased the length of skillful Madden-Julian Oscillation (MJO) forecasts from 6 to 17d, nearly doubled the skill of seasonal forecasts of 2-m temperatures over the United States (reflecting the inclusion of increases in carbon dioxide with time that was not incorporated in CFSv1) and significantly improved global SST forecasts over the CFSv1. As with other seasonal forecast systems, the CFSv2 demonstrated considerable skill in forecasting tropical SST, but small and often negligible skill for 2-m temperature and especially precipitation over land.

Barnston and Tippett (2013) discussed a discontinuity in the CFSv2 and CFSR. CFSv2 SST forecasts for the Nino 3.4 area (120W-150W, 5S-5N) demonstrated a cold bias for 1982-98 and a warm bias for 1999-2009, reflecting a discontinuity in CFSR in 1999 reflecting a change in the observing system. The change in bias led to a dual climatology bias correction.

D. Annual GFS reviews 2011-2015

Annual reviews of GFS implementations and forecast skill from 2011 to 2015 can be found at http://www.emc.ncep.noaa.gov/gmb/STATS_vsdb/longterm/, as well as graphs of long-term skill in the GFS and other global NWP systems.

E. Minor changes May 9, 2011

On May 9, 2011 version 2.0.2 of CRTM, GPS RO data from new satellites, uniform (higher resolution) thinning of satellite radiances, revised quality control and observation and background errors, new observations and an improved GSI code were introduced; other observations were discontinued.

A new thermal roughness length was introduced to reduce a cold bias in land surface skin temperature and low-level summer warm bias over arid land areas; background diffusion was reduced in the stratosphere to strengthen winds. The new thermal roughness also strengthened low level winds. Since this was not considered a major implementation, MDL did not recalculate MOS used to produce bias-corrected fields from GFS output. The existing MOS corrected for too weak low-level winds; following the implementation MOS produced more false alarms of high winds until MDL recalculated the bias correction. The incident highlighted the importance of coordinating changes in the GFS with products produced by downstream users of the GFS.

F. GSI hybrid 3D-EnVar May 22, 2012

A major improvement in data assimilation May 22, 2012, the GSI hybrid 3D Ensemble Variational (EnVar) assimilation used an ensemble of short-range forecasts to improve the representation of flow-dependent background errors (Whitaker *et al*., 2008; Wang *et al*., 2013). The new ensemble data assimilation resulted from collaboration between scientists at EMC, NOAA's Earth Systems Research Laboratory (ESRL) and the University of Oklahoma. Eighty ensemble members at T254L64 were run out to 9 h; a control component was run at T574L64 as was the deterministic forecast. The ensemble was re-centered every cycle about the control analysis; 75% of the background used in the hybrid analysis came from the ensemble, 25% from a static background error.

The new assimilation significantly improved tropospheric fields in week 1 and in the NH week 2; it extended the range of useful forecasts of 500 hPa height in the NH by 2.6 h in summer and 5.3 h in winter. GSI hybrid forecasts fit observations better, but GSI hybrid analyses did not. The use of GPS RO bending angle rather than refractivity was also implemented (Cucurull *et al*., 2012), as was new satellite data and quality control. Descriptions of the June 2012 GFS and the Oct.2014 GEFS can be found at https://www.wmo.int/pages/prog/www/DPFS/ProgressReports/2013/documents/2013_USA.pdf.

G. MEG 2012

In 2012 EMC established a Model Evaluation Group (MEG) (White *et al*., 2017) to increase synoptic evaluation and awareness within EMC and to reduce the number of surprise

problems in EMC forecast systems after implementation. MEG reviewed model performance weekly, initially for EMC; its audience gradually expanded via teleconference to many NWS regional centers and forecast offices and others in government, academia and private industry. As a result of MEG, EMC paid more attention to the GFS's synoptic and sensible weather forecasts and less to systematic errors in the GFS, especially outside the U.S.; NWS forecasters engaged more with EMC; EMC developers responded quickly to complaints about model performance and devised corrections in many instances. During the summer of 2012 MEG found a GFS cold moist bias most evident in hot air masses in the late afternoon over the central U. S. where drought conditions existed. On Sept. 5, 2012 the implementation of an updated look-up table for minimum canopy resistance and root zone depth reduced excessive evapotranspiration and the cold moist bias.

H. Sandy Oct. 30, 2012

On Oct. 30, 2012, after making an unusual turn to the west Hurricane Sandy struck the New Jersey coast as a post-tropical cyclone with hurricane force winds and caused $69 billion dollars in damages and 150 deaths. The ECMWF high resolution model forecast an east coast landfall 8 days in advance; from seven days before landfall ECMWF was consistent in its forecast of landfall; the GFS and GEFS were slower to forecast landfall. All the model guidance agreed on a high-impact land falling storm more than four days in advance. Beyond 72 hours ECMWF forecasts outperformed the GFS. Global ensembles 5 days and more before landfall all had members going out to sea rather than making landfall; these members tended to have a weaker subtropical ridge to the east of Sandy.

In short-range forecasts (72 hours or less) of Sandy's track, the performance of the GFS was comparable to that of the ECMWF high resolution model; the GEFS produced the highest probability of landfall within 300 km radius of the landfall point of any ensemble system (however, the GEFS tended to be less dispersive than other ensemble systems) (Magnusson *et al*., 2014). Torn *et al*. (2015) diagnosed the sources of GFS errors in forecasting Sandy. While all models produced skillful and useful forecasts of Sandy, demonstrating how skillful NWP had become, public attention and debate focused on ECMWF forecasting Sandy's westward turn and landfall earlier than the GFS. ECMWF did not perform better with all tropical systems: ECMWF had forecast an earlier 2012 tropical storm Debbie to come ashore in Texas; the GFS correctly forecast its Florida landfall. For the 2012 season as a whole, GFS forecasts of Atlantic tropical storm tracks were equally or more skillful than ECMWF out to 96 hours (Yang, 2013).

I. Winter storms

One is of the most important weather events in the U. S. is the east coast winter storm, potentially disrupting major population centers. The skill of the GFS in forecasting such storms has greatly improved over the decades, but problems remaining, especially in forecasting the exact placements of snow-rain line and heavy snow. The success of the GFS in forecasting the 1993 "Superstorm" and its failure to forecast the heavy snow of Jan. 25, 2000 more than 12 hours in advance have already been discussed.

In 2010 the GFS successfully forecast the Feb. 5-6 "Snowmageddon" 6 days in advance. In Mar. 2013 several models including the GFS forecast heavy snow for Washington DC on Mar. 6. Several inches of snow fell in the western suburbs of DC, but the city got mainly rain,

reflecting warmer temperatures than forecast. In Jan. 2015, based on guidance from the GFS and other models, the NWS issued heavy snow forecasts for the east coast of the U.S., forecasting over a foot of snow in Philadelphia and up to 3 feet in New York City. Philadelphia got about 2 inches of snow; New York City got less than 1 foot. Forecasts of heavy snow for Long Island, southern Connecticut and Boston were much more accurate. The GFS forecast the Blizzard of 2016 on Feb. 5 and 6 seven days in advance and made excellent predictions for Washington, D. C. and Philadelphia, but had problems with the tight gradient on the northern side of the storm and under-forecast snow totals in New York City (G. Manikin, personal communication).

Model forecasts for the Mar. 2013 "noquester" and the Jan. 2015 blizzard were quite skillful in many respects but had serious problems in major urban areas where offices and schools were closed in anticipation of severe weather conditions that did not occur due to model uncertainty about the exact placement of the snow edge of heavy snow. These two episodes demonstrated the faith placed in numerical weather forecasts and the need for better understanding and communication of forecast uncertainties. The GFS has a problem with the sleet/freezing rain/transition zone and often incorrectly forecasts a pure rain/snow transition, resulting in snow over-forecasts for New York City and Philadelphia in the Mar. 2017 winter storm (G. Manikin, personal communication).

J. New satellite data Aug. 20, 2013

New satellite data from three satellites was introduced Aug. 20, 2013. The Suomi National Polar-orbiting Partnership (NPP) satellite, launched on Oct. 28, 2011, provided 3-D temperature and moisture soundings from the Cross-track Infrared Sounder (CrIS) with 1305

spectral channels; AMSU-A, MSU and GRAS (Global Navigation Satellite System Receiver for Atmospheric Sounding) instruments on METOP-B replaced METOP-A; observations from the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) on Meterosat-10 replaced observations from Meteosat-9. Tests of the new observations showed no significant impact.

10. 2015-2017: 13 km, hybrid 4D EnVAR, NEMS

A. T_L1534 Jan. 14, 2015

On Jan. 14, 2015, the GFS resolution increased from T574L64 (23 km) for the first 192 h to T1534L64 (13km) for the first 240 h; from 240 h to 384 h the new resolution was T574L64. The increase in resolution was enabled in part by a change from Eulerian dynamics to a two-time level semi-implicit semi-Lagrangian advection on a linear grid, allowing larger time steps (Sela, 2010; Kar, 2013) and by more efficient numerical methods (Juang, 2004, 2016; Iredell, 2015). The new model used a 3-D Hermite interpolation to reduce a stratospheric cold bias; semi-Lagrangian advection calculations and physics were done on a linear reduced Gaussian grid. Second order divergence damping increasing with altitude was introduced above 100 hPa to reduce noise (Global Climate and Weather Modeling Branch, 2016).

Higher resolution orography and land-sea mask were introduced; lake temperatures and SST changed from a 5-d running mean at 40 km to daily values at 7 km, the ice analysis similarly became daily values at 7 km. Real-Time Global (RTG) SST replaced the Reynolds SST. Initial snow depth became more dependent on the GFS first guess, reflecting problems uncovered in the Air Force Weather Agency (AFWA) snow depth data. Ice and water cloud

conversion rates, orographic gravity-wave forcing and mountain blocking were retuned; the background diffusion of momentum and the drag coefficient at high wind speeds were reduced.

A 1º bucket soil moisture climatology was replaced with the 27 km CFS/Global Land Data Assimilation (GLDAS) climatology, 1º momentum roughness length climatology was replaced by a look-up table based on vegetation type and a dependence of the ratio of thermal and momentum roughness on vegetation type was added. The change in soil moisture climatology land surface reduced a GFS bias towards "socialist" rain, too much drizzle, but also created a summertime warm, dry bias over the Great Plains. A reduction in the warm, dry bias was developed, but could not be tested in time to include in this implementation.

Stationary convective gravity wave drag was added, reflecting momentum flux induced by subgrid-scale diabatic forcing. The parameterization was proposed by Chun and Baik (1998), tested by Chun *et al*. (2001, 2004) and applied by Johansson (2008) to the GFS, producing modest improvements in the tropical upper troposphere and lower stratosphere. To deal with unresolved sub-grid cloud variability within multi-layered clouds, a Monte-Carlo Independent Column Approximation (McICA) method was introduced in the RRTMG radiative transfer calculation (Iacano *et al*., 2006).

A new hybrid eddy-diffusivity counter gradient (EDCG) and eddy-diffusivity mass-flux (EDMF) PBL parameterization replaced EDCG which underpredicted the daytime convective PBL development (Han *et al*., 2016). EDMF used a mass-flux approach for nonlocal sub-grid scale transport due to strong updrafts and for nonlocal momentum transport and an eddy diffusivity parameterization for transport by smaller eddies. EDMF had too much vertical

mixing in wind and moisture over the tropics producing too large low cloud amounts and degraded tropical wind forecasts; as a result, EDMF was applied to strongly unstable PBL, EDCG to weakly unstable PBL. The local parameterization in the stable PBL was modified to use an eddy-diffusivity profile to strengthen vertical turbulent mixing for weakly and moderately stable conditions. Introducing a parameterization of turbulent kinetic energy dissipative heat significantly increased hurricane intensity, generally much too weak in the GFS. The new parameterization improved 500 hPa height scores, did not significantly change CONUS precipitation scores, and had mixed effects on hurricane track and intensity errors.

The resolution of the ensemble Kalman filter (EnKF) data assimilation increased to T574 from T254, three stochastic physics parameterizations replaced additive inflation in the EnKF perturbation methodology to account for system uncertainty (Pegion *et al*., 2016; Zhou *et al*., 2016), and a new version 2.1.3 of the CRTM improved the specification of microwave sea surface emissivities, improving the analysis of near surface temperatures over water, especially over the southern oceans. An updated radiance bias correction was introduced and new radiances and hourly GOES and European Organization for the Exploitation of Meteorological Satellite (EUMETSAT) atmospheric motion vectors were assimilated, improving moisture and surface winds.

The 13 km GFS was tested in parallel to the old GFS for 32 months and four hurricane seasons. The new GFS resolved smaller scales significantly better and forecast jet level winds, SH fields, extra-tropical synoptic forecasts, CONUS precipitation and tropical storm tracks and

intensity in the Atlantic and West Pacific through day 5 substantially better. Many gains were incremental.

The 13 km GFS was thoroughly evaluated by other NCEP Centers and NWS regional headquarters and forecast offices. They noted particular problems in the new system: worse day 6 and day 7 tropical storm forecasts, no improvement in severe weather forecasts, and much too low 2-m dewpoints in dry regimes were noted. EMC decided to increase collaboration with SPC, communicate with model users earlier for the next implementation and investigate particular case studies of interest to model users as part of future implementations.

B. GEFS T_L574

In version 11 of the GEFS horizontal resolution increased to semi-Lagrangian $T_L574L64$ (34 km) for the first 8d and $T_L382L64$ (52 km) for d 8-16 Dec. 2, 2015. GEFS initial perturbations were selected from the operational hybrid GDAS 80-member EnKF 6h forecast and included tropical storm relocation and centralization of the initial perturbations (Zhou *et al.*, 2016, 2017). The new GEFS significantly outperformed the old out to d 8-10 except for an increased warm bias over land in the extra-tropics; tropical cyclone tracks were significantly degraded at d 6-7 during the 2012-2014 seasons but were improved for the 2015 and 2016 hurricane seasons (Zhou *et al*., 2016, 2017).

C. 4D hybrid EnVar assimilation May 11, 2016

On May 11, 2016, 4D hybrid EnVar data assimilation (Wang and Lei, 2014; Kleist and Ide, 2015) replaced the 3D hybrid EnVar. The ensemble provided an updated estimate of

situation dependent background error every hour as it evolved through the ensemble window. The propagation of background error covariances in time was approximated by an hourly ensemble of forecasts rather than a tangent linear and adjoint model as used in 4D variational formulations; the 4D hybrid EnVar was more scalable, less computationally expensive and more easily applied to other models. The hourly ensemble covariances were combined with a time-invariant error estimate from the GFS's 24-48 h forecast climatology; the ensemble covariances received a weight of 87.5%, the climatology 12.5%.

A variational bias correction to aircraft temperatures reduced a warm bias in the upper troposphere (Zhu *et al*., 2015); Advanced Very High-Resolution Radiometer (AVHRR) winds were introduced. Zhu *et al*. (2016) discussed the assimilation of all-sky AMSU-A microwave radiance in the GSI, limited in this implementation to non-precipitating clouds over ocean surfaces; previously only clear-sky radiances had been assimilated. CRTM was updated to better process cloudy radiances. The all-sky approach produced more realistic simulated brightness temperatures and cloud water analysis increments and improved analysis of oceanic stratus regions.

Land surface characteristics for grassland and cropland in the Noah LSM were changed to reduce summertime warm and dry biases over the Great Plains; convective gravity wave drag was updated. An hourly forecast product out to 120 hours was introduced. A comprehensive web site for the implementation can be found at http://www.emc.ncep.noaa.gov/gmb/noor/4dGFS/synergy%20announcementjan08.htm. The site includes presentations detailing the components of the implementation, evaluations by NCEP

centers, NWS regional headquarters and forecast offices and others, objective evaluation by EMC and case studies requested by NWS regional headquarters and forecast offices and produced by EMC. Maps of three months of real time operational and experimental forecasts were available to model users. Dorian *et al*. (2017) reviewed the implementations of 2016 and 2017 and the interaction of EMC with operational forecasters in developing and evaluating model changes.

Forecasts of 2-m temperatures, dew points and 10 m winds verified against surface observations improved significantly over CONUS and Alaska; forecasts of 2-25 mm/day precipitation thresholds improved significantly over CONUS, as did forecasts of CAPE over the United States, tropical storm genesis, track and intensity and jet streams. Week 1 forecasts of atmospheric fields were significantly improved except in the upper stratosphere. Synoptic evaluation of the new GFS yielded no major new concerns.

A documentation of the GFS as of May 2016 is in Global Climate and Weather Modeling Branch (2016); a description of 2016 GFS physics can be found at: https://dtcenter.org/GMTB/gfs_phys_doc_dev/index.html.

D. NEMS July 19, 2017

The GFS2017 introduced the NOAA Environmental Modeling System (NEMS) (Iredell, 2013) July19, 2017, a shared, portable, high performance software superstructure and infrastructure with a modular structure for all NCEP models. A comprehensive web site for this implementation is: http://www.emc.ncep.noaa.gov/gmb/noor/GFS2017/GFS2017.htm. Also

implemented were upgrades to the land surface, updates to the cumulus convection parameterization and the introduction of Near-Surface Sea Temperature (NSST).

Land surface changes included the replacement of 1^{o} land and soil classifications by 1 km classifications to correct patchy near surface fields, new Moderate Resolution Imaging Spectroradiometer (MODIS)-based albedo fields and diurnal albedo treatment, increased ground heat flux under snow cover and introduction of a stability parameter constraint to prevent the land-atmosphere system from fully decoupling (Zheng *et al*., 2017). Before the change, decoupling sometimes produced excessive near-surface cooling in the late afternoon and nighttime that persisted for hours. The stability constraint improved 2-m air temperature forecasts, temperature forecasts throughout the lower atmosphere and forecasts of light and medium amounts of precipitation. The excessive cooling had been noted in Russia as well as in the U. S. Problems with a near-surface warm bias in winter over Siberia were subsequently reported.

Several changes updated the GFS convective parameterization (Han *et al*., 2017). SAS assumed that the fractional areas of convective updraft over the grid box were negligibly small; however, for resolutions of 10 km or less the assumption was not valid and a scale-aware parameterization was introduced, as was the effect of aerosol on convection. Other changes were made to deal with GFS problems with precipitation: too much light rain, too much convective precipitation, and unrealistically noisy "popcorn" rainfall, especially over high terrain. Convective triggering was made more difficult. Changes to the convection in the GFS17 included using convective turnover time as the convective adjustment time rather than a function

of grid-scale vertical velocity, an additional trigger condition with convective inhibition in the sub-cloud layer, increasing lateral entrainment in lower relative humidity to more strongly suppress convection in a drier environment, allowing shallow convection in the stable PBL, enhanced convective cloudiness as a function of suspended cloud condensate, and rain conversion rate decreasing with decreasing air temperature above the freezing level.

NSST described the oceanic vertical temperature structure and its diurnal cycle near the surface due to diurnal warming and sub-surface cooling processes. The GFS2017 GSI analyzed SST together with atmospheric variables. Data assimilation was also improved by additional satellite observations, minor bug fixes mostly related to cloud water, preparation for future satellite data and changes to land surface type specification in CRTM.

This upgraded system was tested for 749 days over three summers and two winters of forecasts and evaluated in coordination with other NCEP centers and NWS regional headquarters and forecast offices. Maps of several months of real time operational and upgraded forecasts were available to operational forecasters; selected case studies recommended by forecasters were conducted by EMC. The evaluation is detailed on http://www.emc.ncep.noaa.gov/gmb/noor/GFS2017/GFS2017.htm. Objective verification against observations and the model's own analyses showed small changes in the troposphere and improvements in the stratosphere, stronger and more realistic winds, smaller analysis increments outside the tropics, and improved fits to radiosondes and aircraft observations overall.

These changes significantly improved CONUS precipitation forecasts, especially in summer, and reduced the excessive drizzle seen in the GFS and a dry bias for moderate rain.

CONUS forecasts significantly improved for thresholds of 0.2 to 15 mm/day over forecast lengths of 0-24 to 72-96 hrs. Precipitation patterns averaged over several weeks showed improvements; AWC and Weather Prediction Center noted improvements in tropical convection. NEMS forecasts better maintained convection over the western tropical Pacific. Biases in U. S. 2-m temperatures and dew points against observations were reduced overall, RMS errors in 2-m dew points improved at all time of days, and RMS error in 2 m temperatures improved at 00 UTC (reflecting a reduced excessive cold bias at sunset). NHC found an 8-9% degradation in Atlantic tropical storm forecasts at 24-72 h. Correcting problems in the hurricane relocation algorithm in the NEMS GFS and including Global Hawk dropsonde data improved tropical storm forecasts.

E. Downstream problem in the GEFS

A downstream problem emerged in the GEFS that used NSSTs as initial conditions but did not include a diurnal cycle in SST in their forecasts. (Changes to the GFS are implemented later in the GEFS due to insufficient resources to change and test both at the same time.) Forecasts originating at 1800 UTC used warm early afternoon SSTs as initial conditions in the west Atlantic; since the GEFS model lacked a diurnal cycle in SST, these warm SSTs continued in the 1800 UTC forecasts, distorting the tropical divergence pattern and generating too many tropical storms. The GEFS switched to an initial SST without a diurnal cycle.

F. Development of NWP Internationally

Over the years EMC has worked with acientists in China, Taiwan, South Korea, India and Brazil to develop NWP in those countries. For example, the Indian Institute of Tropical Meteorology with assistance from NCEP has transformed the CFSv2 into a long-term climate model (Swapna *et al*., 2015) and has developed a 21-member version of the GEFS at T_L1534 forecasting out to 240 h that the Indian Meteorological Department is running in operations since Jun 1, 2018. NCEP has worked routinely with weather services in Africa, the Caribbean, and Central and South America to improve their use of NWP.

11. Forecast skill

Nearly 40 years of developing the spectral GFS has produced steady increases in forecast skill that have transformed weather prediction. This section demonstrates the effect of the many changes in the GFS over the decades on the skill of its forecasts. Long-term statistics for the performance of the GFS and other global forecast systems are available at http://www.emc.ncep.noaa.gov/gmb/STATS_vsdb/longterm/; a much more extensive evaluation of GFS and other global forecast systems over the most recent 30 days can be found at http://www.emc.ncep.noaa.gov/gmb/STATS_vsdb/. Yang (2012) reviewed GFS forecast skill and major upgrades.

A. 500 hPa heights

Fig. 5a compares annual mean anomaly correlations for 5d forecasts of 500 hPa geopotential height by the operational GFS and by the CFSR in the NH and SH; Fig. 5b displays

the difference in 5d anomaly correlations between the CFSR and GFS. The CFSR forecasts were generated with a T382L 64 assimilation/T126L64 forecast model similar to the 2010 GFS (Saha *et al*., 2010). The CFSR indicates the effect on skill of improvements in observations (during the last few years the CFSR 's implementation of new satellite observations has lagged the GFS); the GFS reflects improvements in the GFS as well as observations. Improvements in observations have increased skill by a few points in the NH and significantly more in the SH; however, improvements in the GFS have raised forecast skill much more than improvements in observations. EMC has devoted considerable effort to adapting the GFS to take advantage of new observations as they became available; the increasing skill of the CFSR over the years reflects extensive EMC efforts as well as the effects of new observations.

GFS anomaly correlations have increased from .55 in 1984 to .89 in 2017 in the NH, increasing the variance explained from 30% in 1984 to 79% in 2017. The SH shows similar gains since 1987; SH skill levels are now only slightly less than the NH, demonstrating the ability of satellite observations to capture synoptic flow patterns nearly as well as conventional observations. The implementation of GFDL "E-2" physics and silhouette mountains in 1985 dramatically increased NH skill; SH skill dramatically increased in the late 1990s. The CFSR shows increased skill in the late 1990s in Fig. 5a, indicating that new satellite data played an important role, but the GFS shows a substantially greater increase than the CFSR, suggesting that improving the use of satellite observations contributed substantially more to increasing skill than the new observations by themselves. The assimilation of satellite radiances began in 1995 and the use of un-preprocessed radiances in 1998. While the CFSR uses the 2010 GFS, the GFS began to outperform the CFSR some years before 2010, perhaps reflecting its higher resolution.

In Fig. 5a, the GFS has lower skill in 2017 than 2016, but the CFSR shows more of a decline, suggesting a decline in atmospheric predictability in 2017. Compared to the CFSR (Fig. 5b) the GFS had its most skillful year in 2017 in both hemispheres.

Table 5 displays the mean anomaly correlation of 500 hPa heights for 5-day and 8-day forecasts averaged from March 2017 to February 2018 for the GFS and other deterministic high resolution global forecast systems. CDAS, forecasts produced by the T62L28 GFS of the mid-1990s used in the NCEP/NCAR reanalysis, is included to show how far NWP has come in the past two decades. ECMWF is the most skillful; UKMO's global system performs somewhat better than the GFS; and the Canadian global system (CMC)'s scores are slightly lower than the GFS. Skill varies with spatial scale; as can be seen in die-off curves accessible on http://www.emc.ncep.noaa.gov/gmb/STATS_vsdb/, if one filters 500 hPa height by zonal wave number, zonal wavenumbers 0-3 can exhibit useful skill out to 9-10 days or more while zonal wavenumbers 10-20 exhibit useful skill only out to 5-6 days.

Figs. 6a (NH) and 6b (SH) display the forecast lengths at which 500 hPa height anomaly correlations reach .8 (regarded as the limit of good forecasts) and .6 (the limit of useful forecasts) for 5 NWP centers' global deterministic forecasts over the past two decades. A 13-month running mean has been applied. All the centers have greatly extended the range of good and useful forecasts over the past two decades; ECMWF forecasts clearly have the most skill, but the difference between ECMWF and GFS has not varied greatly over the period shown. The GFS drew closer to ECMWF in the NH in 2010; since 2010 ECMWF has slightly increased its lead in

the NH; these trends are most obvious for longer range forecasts. CMC has moved a little closer to the GFS in good forecast skill since 2006 but still displays slightly less skill than the GFS.

Over the years higher frequency changes in skill often occur simultaneously in different centers, suggesting changes in atmospheric predictability. In the past few years the centers showed an increase in skill followed by a decline, more pronounced for anomaly correlations of .6 than .8. This may reflect a change in the predictability of the atmospheric circulation, possibly related to ENSO; an earlier peak in 2010 in the NH reflected larger than normal anomalies from

climatology (van den Dool, personal communication). In the SH ECMWF has a somewhat larger lead; ECMWF has very similar skill in the NH and SH while GFS has slightly less skill in the SH than in the NH. Yang *et al*. (2013) found GFS problems with observational errors for certain AMSU-A channels; such a problem would affect the SH more since it lacks as extensive a network of conventional observations as the NH. Alpert *et al*. (2009) noted the greater occurrence of poor forecasts or "dropouts" in the GFS. ECMWF gridded analyses used as pseudo-observations in the GFS improved skill scores in most cases and reduced "dropouts", suggesting that a substantial part of ECMWF's advantage came from their assimilation.

Table 6 displays the 500 hPa height forecast length when the global deterministic forecast systems and global ensemble means reach .8 and .6 during Feb. 2017-Feb. 2018 in the NH. Good forecasts by the deterministic models extend out to 5.8-6.6 days; useful forecasts extend out to 7.8-8.5 days. Ensembles extend the range of useful forecasts by 1.4-1.9 days. The GFS and Canadian ensembles have quite similar skill; NAEFS, the combination of the GFS and Canadian ensembles, is .3 days better; European Centre Ensemble Prediction System (EPS) is

half a day better than NAEFS. Verification of the GEFS and other global ensembles can be found at http://www.emc.ncep.noaa.gov/GEFS/verif.php.

Fig. 7a shows the frequency distribution of daily anomaly correlations for 5-day GFS forecasts of 500 hPa height in the NH for each year from 1996 to 2017. The improvement in forecasts over the years is obvious; the number of forecasts above .9 is dramatically higher in 2015-2017 than earlier. Fig. 7b shows the percent of anomaly correlations below .7 for GFS 000 UTC 5-day forecasts of 500 hPa heights; Fig. 7c shows the percent of anomaly correlations above .9. In the late 1990s over half of SH forecasts and more than a quarter of NH forecasts were below .7; since 2008 less than 5% in either hemisphere are below .7. Until 2003 less than 5% of forecasts were above .9; since 2015 over 30% of SH forecasts and 40% of NH forecasts are above .9. Over the past 2 decades the quality and reliability of GFS forecasts have greatly improved.

B. Tropical wind fits to rawinsondes

In the tropics significant uncertainty remains in forecast system analyses; verification against observations is therefore more important in the tropics than in higher latitudes. Figs. 8a and b display the RMS fit of 200 and 850 hPa tropical wind analyses and forecasts to rawinsonde data over the tropics (20S-20N) from http://www.emc.ncep.noaa.gov/gmb/ssaha/. At both levels the forecast errors have declined over the past 19 years; the fit of 48-h forecasts now is as good as the fit of the 6-h forecast first guess in 1999.

The RMS fit declined markedly in 2010 at both levels; this reflects at least partly an increase in background diffusivity for momentum in the GFS introduced on July 28, 2010, which led to an increased negative wind bias. Reduction in the background diffusivity in later implementations does not appear to have caused a noticeable increase in the RMS fit. The number of observations has increased by a factor of three over the two decades.

C. Precipitation

Fig. 9a displays the annual mean equitable threat score (ETS) over the U. S. for 2003-2017 for thresholds of 0.25,1, and 2 in/day for 24, 48 and 72-hr GFS forecasts. This metric adjusts for the expected number of hits in a random forecast (Mesinger, 1996). The scores generally increase with time, although it is least evident for thresholds of 1 in/day; the increase appears greatest for heavier amounts. For thresholds of 0.25 in/day, 2017 day 3 forecasts are as skillful as day 2 forecasts in 2003; for thresholds of 2 in/day, 2017 day 3 forecasts exhibit as much skill as day 1 in 2003. Fig. 9b displays the annual mean fractional skill (FSS) score for 10 mm/day and 25 mm/day for 24, 48 and 72-hr GFS forecasts over 2003-2017 for a 62-km spatial scale. This score gives credit for near misses (Mittenmaier *et al*., 2011 and shows more consistent year-to-year increases in skill; 2017 day 2 forecasts are as skillful as 2003 day 1 forecasts.

It has been argued that the contingency table-based ETS is less appropriate than the neighborhood-based FSS in gauging the performance of higher resolution models because it tends to overly-penalize even a small mismatch between forecast and observation. Looking at two very different scores (e.g. ETS and FSS) provides a more complete picture. GFS QPF skill

has improved 10-15% overall from 2003 to 2017 as measured by these particular scores. There are far fewer events greater than 2"/day than at lower thresholds, therefore scores at higher thresholds are more volatile; some years have more predictable weather systems than others (Ying Lin, personal communication).

D. Tropical storms

Yamaguchi *et al*. (2017) conducted an inter-comparison of tropical cyclone track forecasts by operational global NWP models over the period 1991-2014 for the World Meteorological Organization's Working Group on Numerical Experimentation. While year-to-year variability is substantial, especially for individual ocean basins, for the globe as a whole track errors in ECMWF, NCEP and UKMO global model forecasts decreased from 2004 to 2014 by 4.2% per year for day 1 forecasts and by decreasing percentages with increasing forecast length out to 2.6% per year for day 5. ECMWF forecasts displayed the smallest errors, especially in the SH. During 2012-2014 the GFS was second out of 12 global models, slightly better than UKMO. Leonardo and Colle (2017) evaluated the skill of deterministic and ensemble forecasts of North Atlantic tropical cyclones by global forecast systems for 2008-2015 and found that deterministic ECMWF forecasts had the smallest track errors at all lead times but were not significantly different from EPS and GEFS ensemble mean forecasts. The GEFS' s weak intensity bias limited its skill at forecasting intensity.

In the 1990s EMC and NHC began collaborating to improve GFS tropical storm forecasts (Surgi *et al*., 1998). Fig. 10 displays the skill of GFS tropical cyclone intensity and track forecasts for the Atlantic (Fig. 10a) and east Pacific (Fig. 10b) for 1995-2017 relative to the NHC

climatology-persistence skill baseline (OCD5) (https://www.nhc.noaa.gov/modelsummary.shtml). While year to year skill varies, reflecting the small sample size in some years, the gain in skill relative to the statistical forecasts from the late 1990s to the 2010s is quite large. In the late 1990s intensity forecasts were quite poor compared to the statistical model; now intensity forecasts tend to equal or exceed the statistical model beyond the first 36 h. GFS track forecasts had less skill than OCD5 in the late 1990s; now they have much more skill than OCD5. While skill in the east Pacific clearly improved in the last few years, it still slightly trails skill in the Atlantic.

Fig. 11 displays the annual mean track error in GFS tropical cyclone forecasts from 2001 to 2017 in the Atlantic (Fig. 11a) and east Pacific (Fig. 11b). Year to year variability in error is clear, but long-term improvement in forecasts is obvious. Day 4 track error in 2016-2017 is clearly lower than day 3 error in 2001-2002 and nearly as low as day 2 error in 2001-2002.

A manifestation of the GFS's increasing skill in the tropics is the recent emphasis on forecasting tropical storm genesis. Fig. 12 shows the verification of GFS tropical cyclone genesis forecasts from 2004 to 2017 over the North Atlantic (Fig. 12a) and the east Pacific (Fig. 12b). The methodology of the verification is detailed in Halperin *et al*. (2016); the paper verified tropical cyclogenesis forecasts by four global systems from 2004 to 2014. The GFS significantly improved from the first to the second decade of the 21st century, although year to year performance varied considerably, reflecting the limited sample size and differences in the location and type of cyclogenesis. In the Atlantic there is reduced skill in 2017 compared to 2016; a quarter of the false alarms in 2017 were associated with one particular system that never

intensified in reality to become a tropical cyclone, lowering the overall skill of 2017 genesis forecasts (Halperin, personal communication). Wang *et al*. (2018) found that the predictability of tropical cyclogenesis over the North Atlantic varied considerably with synoptic regime.

12. The FV3

The GFS is now close to a resolution requiring non-hydrostatic dynamics. In the near future the forecast system may well require scaling across hundreds of thousands of processors. Review committees and others have recommended a simplified NCEP production suite, greater unification of modeling efforts within EMC, greater community involvement and accelerated forecast improvements.

To meet these objectives, the Next Generation Global Prediction System (NGGPS) program selected and is testing and implementing a new dynamic core for the GFS. Phase 1 identified and evaluated qualified dynamic cores for scalability, stability and characteristics. GFDL's Finite Volume Model on the Cubed Sphere Grid Version 3 (FV3) (Lin, 2004; Putnam and Lin, 2007). (https://www.gfdl.noaa.gov/fv3/) and NCAR's Model for Prediction Across Scales (MPAS) (https://mpas-dev.github.io/) dynamic cores were the most qualified. Mesinger *et al.* (2018) reviewed the NGGPS assessment of scalability, the cubed sphere grid and the FV3. Phase 2 involved testing the two dynamic cores with operational GFS physics and evaluating their performance (Toepfer and Schneider, 2016; Michalakes and Whitaker, 2016). The FV3 core was found to be the lower risk, lower cost choice.

In phase 3, the FV3 has been integrated into NEMS and GFS data assimilation and coupled to GFS physics and GFDL microphysics. The FV3GFS ran in parallel to the GFS in 2018 and in retrospectives for evaluation and has been approved to replace the GFS in early 2019. Results of the evaluation can be seen at ***https://www.emc.ncep.noaa.gov/users/Alicia.Bentley/fv3gfs/.*** Future implementations will increase the resolution and upgrade the physics of the FV3GFS.

13. Conclusion

The NWP revolution has been made possible by rapid advances in electronic computers, data processing and observations, and especially numerical modeling of physical knowledge and modern data assimilation techniques, since World War II. Mass (2014) called NWP one of the highest achievements of the human race and perhaps our most cooperative activity. In 2013 Superstorm Sandy caused 150 deaths. In 1938 a similar storm struck a more sparsely populated area of Long Island; over a thousand perished. Based on a survey, Lazo *et al*. (2009) estimated that Americans valued forecasts at $286/household/year, or $31.5 billon/year for the whole country; producing weather forecasts cost $5.1 billion/year. Bauer *et al*. (2015) reviewed "the quiet revolution" of NWP and found the impact of NWP among the greatest of any area of physical science. They also discussed future challenges for NWP.

The success of NWP and the steadily increasing skill of the GFS have multiplied enormously the range of uses of the GFS. It is used not only for short-range North American forecasts, but for global forecasts out to 16 days; it produces believable storm forecasts more than 10 days in advance and displays useful skill beyond 8 days on average, beyond 9 days in

many months, particularly in the NH winter. In 1985 it took the addition of the GFDL E-2 physics package and enhanced orography for the MRF85 to achieve useful skill out to 5 days in the NH winter. EMC development efforts have extended the range of skillful GFS forecasts by roughly 1 day every decade.

The GFS is used for global aviation forecasts, tropical forecasts, severe weather forecasts up to a week in advance, tropical storm genesis, track and intensity forecasts out to 7 days and beyond, and stratospheric warming forecasts. It is used widely around the world; forecasters in Argentina and Russia have contacted EMC with concerns about 2-m temperature forecasts; users in Israel and Spain have requested access to output from experimental versions of the GFS during implementation tests. Many national weather services outside the U. S., universities and private companies use GFS outputs as boundary conditions/initial conditions for their limited-area numerical weather forecasts (R. Garreaud, personal communication). The success of NWP and the GFS has caused a proliferation of commercial weather ventures; people now access constantly updated 10-day forecasts for any global location on their phones. A version of the GFS, the CFSv2, makes long-range forecasts of great interest to many users weeks, months and seasons in advance.

The progress of NWP has been driven by the translation of our imperfect and changing understanding of the science onto frequently changing and improving computer technology that allowed more and more detailed representations of weather and climate with more and more complex representations of physical processes. Inevitably mistakes occurred. Problems sometimes emerged during testing or following implementation; many of these were quickly

solved or ameliorated by EMC. The discovery, investigation and correction of these mistakes has been a significant part of the progress of NWP. The director of GFDL once told a modeler: "The only mistakes you should apologize for are the ones you haven't found" (J. Smagorinsky, personal communication).

The GFS is the foundation of the modern NCEP Production Suite; most of the rest of the Production Suite are dependent on fields from the GFS, as is evident in Fig. 13. GFS forecasts are available 4 times a day out to 16 days, more frequently, longer and earlier than ECMWF forecasts; they are available without charge to all users. The operational GFS computer code has long been available to the meteorological community and workshops have been held on running the GFS; FV3 codes are already available accompanied by analysis data for particular cases. As shown above, EMC has collaborated with many people and organizations. The extensive cooperation between NWP centers around the world deserves more recognition than it has received.

EMC interacts extensively with forecast users, responding to complaints and asking forecasters to evaluate proposed changes. Users have helped EMC set priorities. Recently EMC has involved users more in developing and evaluating forecast system changes, examining synoptic cases proposed by NWS forecasters, establishing MEG, opening its meetings to all, and sending MEG members to visit local forecast offices. In the last two GFS implementations, EMC introduced a longer scientific evaluation process and sought greater user access to parallel tests and increased user involvement in evaluating proposed changes.

The GFS's success is due to the dedication, hard work, knowledge and intelligence of hundreds of individuals over decades, a diverse group from around the world, most of whom are at most identified here only by their papers and reports. Many were civil servants working directly for the federal government; many others worked for private companies such as General Sciences Corporation, Science Applications International Corporation and I.M. Systems Group, Inc. Its success has depended on many outside EMC, particularly NCEP Central Operations' ensuring that the GFS runs on time every 6 hours every day and new versions run smoothly from their first operational forecast.

It has been said that prediction is the ultimate test of understanding. The steady increase in GFS skill shows increased understanding of the weather and improved application of that understanding over decades by extensive trials of new observations, techniques and parameterizations and discovering, investigating and correcting errors. Increasingly powerful computer technology enabled the increase in skill but only after EMC adapted to strict, unforgiving and frequently changing programming requirements. EMC's persistence in developing and improving the GFS over the decades has revolutionized our ability to safeguard life and property.

Acknowledgments Ken Campana, Pete Caplan and Fanglin Yang compiled an invaluable list of changes in the operational GFS over the decades. Glenn White organized and wrote the paper. Steve Lord, John Derber, Henry Juang, Shrinivas Moorthi, Mark Iredell, Ron McPherson, Geoff DiMego, Geoff Manikin, Jim Yoe, Bob Kistler, Ying Lin, Mike Kane, Eugenia Kalnay and Lou Uccellini made valuable comments on the manuscript. Fanglin Yang, Ying Lin, Dan Halperin, Andrew Penny, Geoff DiMego, Mike Kane, Suru Saha, Jack Woollen, Yuejian Zhu, Michiko Masutani and Bill Lapenta contributed figures and tables to the manuscript. Vijay Tallapragada proposed writing the paper. The development of the GFS was documented over the decades by many scientists who did the work and documented it in papers, office notes, technical bulletins and presentations without which this paper could not be written. Jan Thomas helped find many of these documents. Eric Rogers established an online web site for relevant documents.

Appendix A

List of Acronyms

AER Atmospheric and Environmental Research, Inc.
AFWA Air Force Weather Agency
AIRS Atmospheric Infrared Sounder
AMIP Atmospheric Model Inter-comparison Project
AMSU Advanced microwave sounding unit
AMV Atmospheric Motion Vector
ARM-SURFRAD Atmospheric Radiation Measurement-Surface Radiation Budget Network
ASCAT Advanced SCATterometer
ATMS Advanced Technology Microwave Sounder
ATOVS Advanced TOVS
AVHRR Advanced very high-resolution radiometer
AVN version of GFS for short-range aviation forecasts
AWC Aviation Weather Center
CAPE Convective available potential energy
CDAS Climate Data Assimilation System
CFC ChloroFluoroCarbon
CFS Climate forecast system
CFSR CFS coupled reanalysis
CMC Canadian Meteorological Centre
C/NOFS Communications/Navigation Outage Forecasting System satellite
CONUS Continental U. S.
COSMIC Constellation Observing System for Meteorology, Ionosphere and Climate
CPC Climate Prediction Center
CrIS Cross-track Infrared Sounder
CRTM Community Radiative Transfer Model
CSBT Clear Sky Brightness Temperature
DD Development Division
DOE Department of Energy
ECMWF European Centre for Medium-Range Weather Forecasts
EDCG Eddy-Diffusivity Counter-Gradient
EDMF Eddy-Diffusivity Mass-Flux
EMC Environmental Modeling Center
EnKF Ensemble Kalman Filter
ENSO El Nino-Southern Oscillation
EnVar Ensemble Variational
EPS European Centre Ensemble Prediction System
ERS Earth Remote-sensing Satellite
ESA European Space Agency

ESMF Earth System Modeling Framework
ESRL Earth Systems Research Laboratory
ETR Ensemble Transform with Rescaling
ETS Equitable Threat Score
EUMETSAT European Organization for the Exploitation of Meteorological Satellites
FGGE First GARP Global Experiment
FNL Final analysis
FSS Fractional Skill Score
FV3 Finite Volume Model Version 3
GARP Global Atmospheric Research Program
GCM General Circulation Model
GDAS Global Data Assimilation System
GEFS Global Ensemble Forecast System
GFDL Geophysical Fluid Dynamics Laboratory
GFS Global Forecast System
GLDAS Global Land Data Assimilation
GNSS Global Navigation Satellite System
GODAS Global Ocean Data Assimilation System
GOES Geostationary Operational Environmental Satellite
GPS Global Positioning System
GRAS Global Navigation Satellite System Receiver for Atmospheric Sounding
GSI Grid-point Statistical Interpolation
HIRS High-resolution Infrared Radiation Sounder
HPC Hydrometeorological Prediction System
IASI Infrared Atmospheric Sounding Interferometer
IBM International Business Machine Corporation
IOAS-AOLS Integrated Observing and Assimilation Systems for Atmosphere, Oceans and Land Surface
JCSDA Joint Center for Satellite Data Assimilation
JMA Japan Meteorological Agency
JNWPU Joint Numerical Weather Prediction Unit
LAS Low-Atmosphere temperature Sounding
LSM Land Surface Model
McICA Monte-Carlo Independent Column Approximation
MDL Meteorological Development Laboratory
MEG Model Evaluation Group
METAR MEteorological Terminal Aviation Routine weather report
MetOp Meteorological Operational Satellite
MHS Microwave Humidity Sounder
MJO Madden-Julian Oscillation
MODIS MODerate- resolution Imaging Spectroradiometer
MOM Modular Ocean Model
MOS Model Output Statistics
MPAS Model for Prediction Across Scales

MPP Massively parallel processing
MRF version of GFS for Medium-Range Forecasts
MSU Microwave Sounding Unit
NAEFS North American Ensemble Forecast System
NAM North American Mesoscale forecast system
NCAR National Center for Atmospheric Research
NCEP National Centers for Environmental Prediction
NCO NCEP Central Operations
NCWCP National Centers for Weather and Climate Prediction
NDGD National Digital Guidance Database
NEMS NOAA Environmental Modeling System
NESDIS National Environmental Satellite Data and Information Service
NEXRAD NEXt-generation RADar
NGGPS Next-Generation Global Prediction System
NH Northern Hemisphere
NHC National Hurricane Center
NMC National Meteorological Center
NOAA National Oceanic and Atmospheric Administration
NPP National Polar-orbiting Partnership
NRL Naval Research Laboratory
NSST Near-Surface Sea Temperature
NWP Numerical Weather Prediction
NWS National Weather Service
OCD5 NHC climatology-persistence skill baseline
OFCL Official
OI Optimal Interpolation
OMI Ozone Monitoring Instrument
OPC Ocean Prediction Center
OPTRANS Optical Path TRANSmittance algorithm
OSU Oregon State University
PBL Planetary Boundary Layer
POWER Performance Optimized With Enhanced RISC
QBO Quasi-Biennial Oscillation
QC Quality Control
QuikSCAT Quick SCATterometer
RARS Regional ATOVS Retransmission Services
RAS Relaxed Arakawa-Schubert
RISC Reduced Instruction Set Computer
RMS Root Mean Square
RO Radio Occultation
RRTM Rapid Radiative Transfer Model
RRTMG RRTM for GCM applications
RTG SST Real-Time Global SST
RTMA Real-Time Mesoscale Analysis

RTNeph Real-Time Nephanalysis
R-2 NCEP-DOE AMIP-II Reanalysis
SAC-C Scientific Application Satellite-C
SAPHIR Soundeur Atmospherique du Profil d'Humidite Intertropical par Radiometrie
SAR-X Synthetic Aperture Radar-XBand sensor
SAS Simplified Arakawa-Schubert
SBUV Solar Backscatter UltraViolet radiometer
SEVIRI Spinning Enhanced Visible and InfraRed Imager
SH Southern Hemisphere
SLP Sea Level Pressure
SP2 Scalable POWER parallel 2 system
SSI Spectral Statistical Interpolation
SSM/I Special Sensor Microwave/Imager
SST Sea Surface Temperature
SSU Stratospheric Sounding Unit
STTP Stochastic Total Tendency Perturbation
THORPEX THe Observing system Research and Predictability Experiment
TIGGE THORPEX Interactive Grand Global Ensemble
TIROS Television InfraRed Observation Satellite
TLNMC Tangent-Linear Normal-Mode Constraint
TMI TRIMM Microwave Imager
TOVS TIROS Operational Vertical Sounder
TOVS level1-b un-preprocessed TOVS radiances
TPB Technical Procedures Bulletin
TRMM Tropical Rainfall Measuring Mission
UKMO Meteorological Office (MO) of the United Kingdom
UPP Unified Preprocessor Package
USGS U. S. Geological Survey
UTC Coordinated Universal Time
UV UltraViolet
VAD Velocity Azimuth Display
VIIRS Visible/Infrared Imaging Radiometer Suite
VTPR Vertical Temperature Profile Radiometer
Y2K Year 2000

Appendix B

History of modifications to the operational global forecast/analysis system

09/18/1974: First global forecast system (**GFS**) in **GDAS**, 12 h forecast first guess
Hough analysis, 9-layer sigma-coordinate, 2.5^{o} grid primitive equation
1977: GDAS "Final" global analysis, 12 h forecast converted to 6 h analysis forecast cycle
09/22/1978: Quasi-univariate optimum interpolation on sigma surfaces in GDAS
05/27/1980: **SPECTRAL** R30L12 in GDAS
Non-linear normal mode initialization, limited physics
08/1980: Spectral in MRF, AVN
R24 at 48 h, L6 at 144 h, NH at 8 days
AVN, MRF Hough analysis
08/14/1982: full multivariate optimum interpolation on mandatory pressure surfaces in GDAS
10/18/1983: R40L12 layers 240 h forecast 00 UTC
04/17/1985: **MRF85** R40, 18 equally spaced layers, 12 layers moisture in MRF
GFDL physics including radiation, vertical diffusion
Silhouette mountains
Zonally symmetric climatological clouds
Kuo convection up to 300 hPa
Evaporation of non-convective rainfall and cooling below cloud base
05/29/1986: **MRF86** higher vertical resolution near surface GDAS, MRF
Shallow convection, Kuo convection throughout troposphere
Stability-dependent vertical diffusion, including sensible heat
Horizontal diffusion quasi-isobaric
SST adjusted to be more consistent with orography
Spatially variable surface roughness based on vegetation
Monthly fields of albedo, sea ice
11/1986: AVN upgraded to MRF86
08/13/1987: **MRF87** T80 simple diurnal cycle
Improved surface fluxes, vertical diffusion
Changes to radiation, Kuo convection
09/03/1987: Gravity wave drag
01/20/1988: 18 layers of moisture
05/18/1988: Penman-Monteith evapotranspiration
09/20/1988: NESDIS TOVS physical retrievals of satellite temperature soundings replace statistical retrievals
11/30/1988: **MRF88** diagnostic interactive clouds
Horizontal diffusion reduced, semi-implicit zonal advection of vorticity, moisture
New observational and forecast error statistics in data assimilation
12/14/1988: Implementation of Hydrostatic Complex Quality Control
01/1989: Advanced start times of AVN forecasts by 45 minutes to 0245 and 1445 UTC
08/10/1989: Physics and hydrodynamics separated
12/13/1989: **MRF89** surface parameters updated in analysis cycle

Observed snow, ice
Nonlinear initialization of analysis increments
02/1990: More accurate GFDL longwave radiation
04/11/1990: Hydrostatic Complex Quality Control extended to significant levels
05/1990: Horizontal resolution in analysis increased from 4.29^{o} long. X 7.2^{o} lat. to 2.14^{o} long. X 1.08^{o} lat.
02/20/1991: New SST 2^{o} horizontal resolution and 1-week temporal resolution replacing 5^{o} horizontal resolution and 2-week temporal resolution
03/06/1991: **MRF91** T126 (105 km), mean mountains
Diagnosed oceanic stratus clouds
Reduced horizontal diffusion in medium scales based on turbulence theory
More accurate mass conservation
05/21/1991: New fixed fields, correction to Kuo convection, new albedo, stomatal resistance, roughness length
06/25/1991: New spectral statistical interpolation analysis (3D variational), no initialization
11/15/1991: Addition of increments, horizontal and vertical OI checks to Hydrostatic Complex Quality Control
02/1992: 3 hourly radiation, new shortwave (12 bands instead of 9)
08/18/1992: Bogussing of tropical storms
12/07/1992: First GEFS: one pair of bred forecasts at 0000 UTC to 12 days; resolution T62L18. AVN extended to 10 days
03/30/1993: SSM/I wind speeds over water assimilated, analysis changes including low level divergence proportional to vorticity
08/11/1993: **MRF93** Simplified Arakawa-Schubert convection, 28 layers, 40 levels of TOVS data instead of 15
12/07/1993: Interactive satellite retrievals using first guess forecast from MRF
03/1994: Second GEFS: five pairs of bred forecasts at 000 UTC, two pairs at 1200 UTC, extension of AVN, 17 global forecasts every day to 16 days
01/10/1995: **MRF95** SSM/I precipitable water over sea; analysis changes: constraint on divergence tendency, no “super-obs” except for conventional satellite soundings, improved background error; various physics changes: 2-layer OSU soil model, improved surface physics, diagnostic clouds (TPB 417)
10/25/1995: Direct assimilation of TOVS cloud-cleared radiances; physics changes: New PBL based on nonlocal diffusion, improved Hydrostatic Complex Quality Control
11/20/1995: ERS-1 winds using wind direction as first guess from GFS
02/04/1997: SSM/I precipitable water discontinued
06/1997: AVN run four times a day to 78 h with a 2-hour 45-minute data cutoff
11/05/1997: Elimination of "valley snow" and other small changes (TPB 443)
New observational error statistics; change to assimilation of TOVS radiances and addition of other data sources
01/13/1998: Assimilation of non-cloud cleared radiances TOVS-1b radiances;
01/14/1998: Vertical diffusion increased in free atmosphere (TPB 445); improved physics
06/24/1998: T170L42 to 3.5d; physics changes; 3D ozone assimilation and changes; analysis changes (See TPB 449) including nonlinear increments in 3D-VAR

07/21/1998: First emergency implementation: fix to convection, horizontal diffusion, plant evaporation: see last section of TPB 449 above.

10/05/1998: Second emergency implementation: Back to T126L28, better fits to data, more iterations in analysis, physics changes (TPB 450)

12/01/1998: Snow resolution increased from 2.0 to 0.5 degrees; snow depth field added; snow depth no longer estimated by model

01/08/1999: Use of VAD winds eliminated due to problems with light wind speeds

03/08/1999: Introduction of high-resolution radiances from AMSU-A and HIRS-3 instruments on NOAA-15. NOAA-11 no longer providing AMSU data, will soon be unable to provide HIRS data.

01/24/2000: Resolution upgraded from T126L28 to T170L42, restoring the resolution used from June 15 through Oct 5, 1998. MRF T170L42 through day 7, then at T62L28 through day 16. AVN T170L42 out to 84 hours four times a day. (See TPB 452)

05/17/2000: AVN available out to 126h at full (T170) resolution at 00Z and 12Z.

06/27/2000: Resolution of GEFS increased from T62 to T126 for first 60 hours of forecast. T62 out to 16 days, vertical resolution increased to 28 levels out to 16 days. 10 members at 000 and 1200 UTC.

7/06/2000: Hurricanes and tropical storms in the model's guess field relocated to official NHC position in each 6-h analysis cycle. (TPB 472)

08/29/2000: **GDAS** Data cutoff time for the 0600 and 1800 UTC final analysis (FNL) extended from 4 to 6 hours.

10/01/2000: **GDAS GFS** Package of minor changes:

GDAS New observation error diagnosis, rawinsonde radiation correction, effects of balloon drift in time and space included, sea ice used in SSI

Improved ozone analysis and forecast

GFS Higher resolution orography filtered in highest wave number to prevent ringing, reduced gaussian grid over polar regions, new surface albedo climatology, aerosols included in shortwave, single scattering albedo parameter adjusted in shortwave

11/01/2000: **GDAS** ERS wind data erratic; turned off.

01/2001: GEFS T126 out to 3.5 d.

02/13/2001: Satellite radiance and moisture analysis changes plus smoothed output MSLP.

Assimilation of rain rate estimates from SSM/I

(1) improved satellite radiances, (2) improved moisture analysis, and (3) preparation for use of future satellite radiances/products

OBSERVATIONAL AND ANALYSIS CHANGES

Inclusion of GOES-10 and AMSU-B radiances

Updated radiative transfer coefficients for OPTRANS

Enhanced AMSU-A quality control, bias correction of radiance data

Changes in satellite data thinning and channel selection

Modification of near surface observation errors for surface pressure, temperature, moisture, and wind observations

Generalization of moisture constraint to include temperature and pressure sensitivities.

Reformulation of constraint in terms of vapor pressure.

Retuning of weights for upper and lower constraint limits.

Option to modify moisture background error to enhance local sensitivity
Preparation for use of NOAA-16 AMSU-A, AMSU-B, and HIRS radiance data, SSM/I and TRMM precipitation data
Code simplification and optimization changes

05/15/2001: (See TPB 484).
Saturation vapor pressure calculation now made with respect to water, ice or mixed phase depending on temperature.
Weight of moisture constraint limit reduced by one order of magnitude (i.e., from 1.0 to 0.1).
Upper level moisture constraint limit set to 100% relative humidity.
New satellite data-NOAA-16 AMSU-A/B and NOAA-15 AMSU-B microwave radiances
Stronger quality control for AMSU radiances
Improved rawinsonde quality control
Inclusion of cloud condensate as a history variable
Use of cloud condensate in calculation of radiative transfer
New aerosol parameterization added to shortwave radiation with monthly mean climatological aerosol distribution.
Momentum mixing included in deep convection
Refinement of hurricane relocation algorithm
SST anomaly damped toward climatology during forecast, with 90-day relaxation time
New USGS orography dataset to define model orography

06/27/2001: Minor increase in vegetation fraction.

07/24/2001: Data source for the daily 1^{o} SST analysis changed from NOAA-14 to NOAA-16 due to instrument drift.

08/14/2001: Package of minor changes:
Observations and analysis: Higher resolution sea ice mask
Forecast and post processing: Bug corrections in gravity wave drag, randomization of convective cloud tops, land surface evaporation with trace snow cover; minor adjustments in effective radius for ice crystals, auto-conversion rate for ice, evaporation of falling precipitation, and critical RH for condensation

10/09/2001: Snow depth updated daily at 000 UTC from observations; at 0600, 1200 and 1800 UTC model guess used. Formerly 000 UTC update reinserted in each cycle. Assimilation of rain rate estimate from TRMM TMI.

01/15/2002: **GDAS** QuikSCAT surface winds included.

03/05/2002 **AVN** four times a day out to 384 hours. Resolution T170L42 to 180h, thereafter T62L28.

04/23/2002: **AVN** MRF replaced by 000 UTC AVN.

09/-10/2002: **GFS** Name changes: AVN referred to as the Global Forecast System model (GFS).

10/29/2002: (See draft TPB) **GFS** Resolution change old: T170L42 to 180h, T62L28 to 384h new: T254L64 to 84h, T170L42 to 180h, T126L28 to 384h
Modified 6^{th} order horizontal diffusion for T254
GDAS Analysis and observation changes: background error recomputed, AMSU-A channels 12 and 13 from NOAA-15 and NOAA-16 and HIRS from NOAA-16 used,

METAR surface pressure observations used, divergence tendency constraint in tropics turned off.

03/11/2003: **GDAS** NOAA 17 1B radiances assimilated, NOAA-16 AMSU-A radiances restored, QuikSCAT winds superobbed at 0.5 degrees.

08/28/2003: **GFS** RRTM longwave radiation from AER installed: More trace gases (CH_4, N_2O, CFC's) and better tropospheric water vapor absorption.

10/28/2003: **GDAS** NOAA-17 AMSU-A radiances turned off

11/20/2003: **GDAS** Package of minor analysis changes (see description)
More compact fast radiative transfer
Sea surface emissivity model for IR radiances
Cloud top detection for IR radiances
Assimilation of GOES-12 sounder radiances
Time interpolation of the guess
NEXRAD radial winds and lidar line of sight winds

12/09/2003: **GFS** Vertical diffusion added to ozone

02/24/2004: **GFS** Mountain blocking: Parameterization of the separation of airflow in vertical with passage over mountainous terrain.

03/09/2004: **GEFS** Ensemble 4 times daily. Resolution T126 0-180 h, then T62 to 384 h.

05/25/2004: **GDAS** Turn off NOAA-16 HIRS/3 observations

01/31/2005: **GEFS** Radiation code corrected to prevent blowups in ensemble forecasts

05/31/2005: **GFS** (See TPB - includes changes made 06/14/05 and 07/07/05)
Resolution change: old: T254L64 to 84h, T170L42 to 180h, T126L28 to 384h
new: T382L64 to 180h, T190L64 to 384h
Mountain blocking increased, vertical diffusion decreased, sea ice modified.
1D 4-layer sea ice thermodynamics model attached to GFS
Land Surface Model: Upgraded from OSU 2-L LSM to 4-L Noah LSM.
GDAS AIRS radiances assimilated, new cloudiness algorithm, enhanced QC of IR radiances, microwave emissivity model over snow and ice increases use of microwave radiances over snow and ice

06/14/2005: **GFS** Canopy resistance of vegetation increased

07/07/2005: **GFS** Temperature error near top of model corrected

08/25/2005: **GEFS** T126L28 to 384 h. Tropical storm relocation applied to each member.

05/30/2006: **GEFS** 14 members, 4 times daily. Each cycle now has own control member. Resolution T126L28 to384 h. New ensemble transform method generates members.

08/22/2006: **GFS** (See Briefing Notes): New orography and land-sea mask, improved snow analysis, and new ozone physics. New glacial ice and fix error in downward LW radiation at earth's surface. Infrastructure changes (*e.g.* Full ESMF compliance).

03/12/2007: **GEFS** 20 ensemble members, 4 times daily

05/01/2007: **GFS, GDAS**: (See Briefing Notes): Vertical coordinate changed to hybrid sigma-pressure, reducing some upper air model errors. NCEP 3DVAR assimilation system unified under GSI, improving some performance metrics without affecting others and preparing for future analysis improvements. GPS RO global positioning system radio occultation data assimilated. TLNMC tangent-linear normal-mode constraint. CRTM updated radiative transfer to simulate radiances for forward model.

New observing systems added. Radiation package modernized.
Output increased, particularly for hydrology.

05/26/2007: **GDAS** Improved QC and obs errors for satellite-derived winds, include HIRS, MSU, and AMSU-A brightness temperatures from METOP-A, higher resolution, non-averaged GEOS observations

09/25/2007: **POST** New processor unifying postprocessors for GFS, NAM systems.
Change in precipitation type algorithm.

10/10/2007: **GFS** Spurious high pressure over mountainous coasts and interpolation problems corrected.

12/04/2007: **GDAS** GSI code modified to assimilate SBUV-8 and JMA wind data, upgraded to latest version.
GEFS (NAEFS/TIGGE) UPGRADE
GFS deterministic model bias corrected (bc) output through F180
GFS bc output used in GEFS processing
NCEP/GEFS (20 member) and CMC/GEFS (20 member) combined to produce:
Ensemble mean and anomaly, ensemble spread and mode along with 10%,50%,90% probability forecast at 1x1 degree resolution.
Statistical downscaling by using RTMA as a reference: a. At NDGD (5km) resolution (CONUS), generate mean and mode along with 10%, 50% and 90% probability forecast

12/09/2008: **Post** Formulation change in freezing level calculation.
Several new parameters added to output files. Fix to the calculation of precipitation type.

02/24/2009: **GDAS** Inclusion of MetOp IASI data; variational qc; Addition of background error covariance input file; reduced number of AIRS water vapor channels used
Change in land/snow/ice skin temperature variance
Flow dependent reweighting of background error variances
New version and coefficients for community radiative transfer model
Height assignment for height-based wind observations modified
Surface land use file--few permanent (~12) glacial points removed to improve surface temperature forecasts over those points.

12/15/2009: **GFS, GDAS and Post** see **NWS TIN** for the detail.
GFS ESMF upgraded to Version 3.1.0rp2. Output definition of low cloud changed to combine previously separately defined boundary-layer cloud and low cloud.
Baseline/housekeeping version of the GFS established for future implementations.
GDAS new data sources, improved numerical techniques. The analysis changes include: Tropical storm official advisory minimum sea-level pressure observations, NOAA-19 HIRS/4 and AMSU-A brightness temperature, RARS 1b data, NOAA-18 SBUV/2 and Aura OMI ozone, EUMETSAT-9 atmospheric motion vectors assimilated.
Uniform thinning mesh for brightness temperature data. Assimilation of GPS RO data improved. Observation errors retuned. Dry mass pressure constraint.
Goddard Modeling and Assimilation Office/EMC merged code including hooks for 4dvar and observation sensitivity. Update background error covariance file updated.
Post Post-processors used for GFS and GDAS unified; two minor changes introduced into GFS. Unification of the post processing codes resulted in slightly less smooth fields

in GDAS. A number of diagnostic variables improved; additional parameters added to model output files.

02/23/2010: **GEFS** T190L28 to 16 days, model uncertainty using Stochastic Total Tendency Perturbation included

07/28/2010: **GFS** Q3FY2010 Resolution and Physics Major Upgrade, see **NWS TIN.**

Resolution and ESMF Eulerian T574L64 for 0-192h, T190L64 for 192-384 h. ESMF 3.1.0rp2.

Radiation and cloud SW routine changed from ncep0 to RRTM2.
Longwave computation frequency changed from 3h to 1h.
Stratospheric aerosol added to SW and LW, tropospheric aerosol to LW.
Aerosol SW single scattering albedo changed from 0.90 in the operation to 0.99.
SW aerosol asymmetry factor changed. New aerosol climatology.
SW cloud overlap changed from random to maximum-random overlap.
Time varying global mean CO2 instead of constant CO2.
Yang *et al*. (2008) parameterization to treat the dependence of direct-beam surface albedo on solar zenith angle over snow-free land surface

Gravity-Wave Drag Parameterization
GWD modified to automatically scale mountain blocking and GWD with resolution.
Compared to the T382L64 GFS, the T574L64 GFS used four times stronger mountain block and one half the strength of GWD.

Removal of negative water vapor
Positive-definite tracer transport parameterization (Yang *et al*., 2009) in vertical replaced operational central-differencing parameterization to eliminate computationally-induced negative tracers.
GSI factqmin and factqmax parameters changed to reduce negative water vapor and supersaturation points from analysis step.
Cloud physics modified to limit the borrowing of water vapor used to fill negative cloud water to maximum amount of available water vapor to prevent the model from producing negative water vapor.
Minimum value of water vapor mass mixing ratio in radiation changed from 1.0e-5 to 1.0e-20. Otherwise, the model artificially injects water vapor in the upper atmosphere where water vapor mixing ratio is often below 1.0e-5.

Hurricane relocation
Hurricane relocation at 1760x880 forecast grid instead of 1152x576 analysis grid
GDAS pgb files posted first on Gaussian grid (1760x880), then converted to 0.5-deg for hurricane relocation.

Post processing and Utility
GFS forecast master pressure files posted on Gaussian grid, then interpolated to 1-deg for post-processing and archive. 20-bit and faster interpolation instead of 16-bit.
New interpolation double precision, fix in dry air mass (pdryini2=0)

Snow analysis T574 compatible high-resolution snow analysis

Upgraded boundary layer parameterization
Stratocumulus-top driven turbulence mixing.
Stratocumulus top driven diffusion enhanced when condition for cloud top entrainment

instability met. Local diffusion for the nighttime stable PBL.
Background diffusion in inversion layers below 2.5km over ocean reduced by 70% to decrease the erosion of stratocumulus along the coastal area.
New mass flux shallow convection parameterization
Cloud water detrained from every updraft layer
Convection starting level defined as level of maximum moist static energy within PBL
Cloud top limited to 700 hPa.
Entrainment rate inversely proportional to height and detrainment rate set to be a constant as entrainment rate at the cloud base.
Mass flux at cloud base function of convective boundary layer velocity scale.
Updated deep convection parameterization
Random cloud top eliminated, cloud water detrained from every layer of single cloud.
Finite entrainment and detrainment rates for heat, moisture, and momentum specified.
Similar to shallow convection parameterization, entrainment rate inversely proportional to height in sub-cloud layers and detrainment rate constant as entrainment rate at the cloud base.
Above cloud base, organized entrainment added, function of environmental relative humidity
Increased momentum background diffusivity for winds only
Convective overshooting: increased cloud water detrainment in upper cloud layers

05/09/2011: **GDAS/GFS** Q2FY2011 Minor Upgrade.
GDAS Improved OMI QC
Redundant SBUV/2 total ozone removed
SBUV/2 ozone ob errors retuned
AMSU-A Channel 5 QC relaxed
New version of CRTM 2.0.2
Field of View Size/Shape/Power for Radiative transfer
GPS RO data from SAC-C, C/NOFS and TerraS AR-X satellites
Down weighting of collocated radiances removed
Moisture limited >= 1.e-10 in each outer iteration and at end of analysis
Uniform (higher resolution) thinning for satellite radiances
Location of buoys improved in vertical (moved from 20 to 10m)
GSI code improved with optimization and additional options
Background errors recomputed
SBUV, MHS from NOAA-19; of AMSU-A NOAA-15 removed
Ambiguous vector quality control for ASCAT (type 290) data
GFS New Thermal Roughness Length -- Reduced land surface skin temperature cold bias and low-level summer warm bias over arid land areas.
Minimum moisture Value in stratosphere set to 1.0E-7
Background diffusion reduced in stratosphere
Post Error in 192 h, 12 h precipitation bucket corrected
New membrane sea level pressure (SLP) field to eliminate smoothing of MSLP over water to prevent misalignment of pressure and wind fields.

02/2012 **GEFS** T254L42 to 8d, T190L42 to 16d

05/22/2012: **GDAS** Q3FY2012 Major Upgrade. Hybrid variational/ensemble assimilation system. Background error projecting information in observations into analysis created by combination of static background error (as in the prior system) and -new background error produced from lower resolution (T254) Ensemble Kalman Filter. see **NWS TIN** .
Other Analysis Changes: GPS RO bending angle rather than refractivity compressibility factors for atmosphere.
SBUV observation errors retuned, bug fixed at top
Radiance usage flags updated
Monitoring NPP and METOP-B satellite data prepared
NPP ATMS satellite data, GOES-13/15 radiance data, SEVIRI CSBT radiance product, new satellite wind data and quality control added
Satellite monitoring statistics code

09/05/2012: **GFS** Minor Change.
Look-up table in land surface parameterization for minimum canopy resistance and root zone depth updated to reduce excessive evapotranspiration, mitigating GFS cold and moist biases found in late afternoon over central United States when drought conditions existed in summer of 2012. see **NWS TIN** .

08/20/2013: **GDAS** Suomi NPP CrIS, METOP-B AMSU-A, MHS, GRAS, Meteosat-10 SEVIRI assimilated

07/03/2013 **GDAS** HIRS4 radiances on NOAA-19 removed due to increased noise

01/14/2015: T1534 Semi-Lagrangian **GFS** Major Upgrade (NWS TIN)
GFS Upgrade from T574 Eulerian (~23km) to T1534 Semi-Lagrangian (~13km)
High resolution daily RTG SST instead of weekly OI SST, use daily sea ice analysis
High resolution forecast extended from 8 days to 10 days.
McICA radiation approximation, reduced drag coefficient at high wind speeds
Hybrid EDMF PBL parameterization and turbulent kinetic energy dissipative heating
Ice and water cloud conversion rates, background diffusion of momentum and heat, orographic gravity-wave forcing and mountain block retuned
Change from Lagrangian to Hermite interpolation in the vertical to reduce stratospheric temperature cold bias
Restructured physics and dynamics restart fields and updated sigio library
Consistent diagnosis of snow accumulation in post and model
Frozen precipitation fraction computed and outputed
Divergence damping in the stratosphere to reduce noise
Tracer fixer added for maintaining global column ozone mass
Stationary convective gravity wave drag
New blended snow analysis to reduce reliance on AFWA snow
Changes to treatment of lake ice to remove unfrozen lake in winter
Initialization modified to reduce sharp decrease in cloud water in first model time step
Bug corrected in the condensation calculation after digital filter applied
Bucket soil moisture climatology replaced by CFS/GLDAS
Vegetation dependence added to ratio of thermal and momentum roughness
Momentum roughness issue fixed
Accumulation bucket changed from 12 hour to 6 hour between day 8 and day 10

GDAS Convert GFS GSI to vertical structure
Increase horizontal resolution of ensemble from T254 to T574
T574 analysis for T1534 deterministic forecast
Number of second outer loop iterations reduced from 150 to 100.
Stochastic physics parameterizations replaced additive inflation to account for system uncertainty
Changes in radiance assimilation: upgrade to CRTM v2.1.3
Enhanced radiance bias correction parameterization
Bug in AMSU-A cloud liquid water bias correction term corrected
New radiances: F17 and F18 SSM/IS, METOP-B IASI
Known bad channels turned off: Aqua AIRS channels 321, NOAA-19 AMSUA channel 7, NOAA-19 MHS channel 3
ATMS observation errors increased: increase channels 6 - 10 from 0.3 K to 0.4 K, increase channels 11 - 12 from 0.4 K to 0.45 K
Cloud detection channels turned on for monitored instruments: NOAA-17, -19 HIRS, GOES-13 and -14 sounders
Changes in assimilation of atmospheric motion vectors (AMV): assimilate EUMETSAT, NESDIS GOES hourly AMVs, improve AMV quality control
GPS RO Improve quality control, use METOP-B
SSM/IS UPP LAS data assimilated, modified bias correction LAS data

12/02/2015: **GEFS** IC perturbations selected from operational GDAS 80-member Ensemble Kalman Filter 6-hr fcsts, T574L64 to 8 days, T382L64 to 16 days semi-Lagrangian

05/11/2016: (NWS TIN) **GDAS** Hybrid 3D Ensemble-Variational upgraded to hybrid 4D Ensemble-Variational Data Assimilation System, Multivariate Ozone update
All-sky (clear and cloudy) AMSU-A radiances, AVHRR winds assimilated
Modify thinning/weight in time for AMVs and radiances
Update CRTM updated to v2.2.1 with bug fixes in wind direction, use of FAST microwave emissivity models, reflection correction for cloudy situations
Aircraft data bias corrected
Relocation and storm tracking modified to allow hourly tropical cyclone relocation
GFS Corrections to land surface to reduce summertime warm, dry bias over Great Plains
Post Hourly output fields through 120-hr forecasts
Improved icing probability products and new icing severity product
Five more levels from 10 hPa to 1 hPa in post-processed pgb files

07/19/2017 **GFS** NEMS superstructure upgrade
Land surface Changes to increase ground heat flux under deep snow, unify snow cover fraction and snow albedo in radiation and land surface parameterizations
New high-resolution MODIS-based snow-free albedo, maximum snow albedo, soil type and vegetation type
Surface layer parameterization changed to modify roughness-length formulation and to introduce stability parameter constraint in Monin-Obukov similarity theory to prevent land-surface system from decoupling
GFS NEMS framework
NSST added to couple with GFS to provide more realistic thermal boundary condition

Scale- and Aerosol-Aware deep and shallow convection parameterizations with convective cloudiness enhancement to make code more applicable for higher resolution
Rayleigh damping tuned to improve temperature and wind forecasts in upper stratosphere
Problems in pressure and hybrid coordinates transformation during tropical storm relocation corrected
Post Divergence equation used to compute omega on grid space for GFS NEMS output
GDAS Near Sea Surface Temperature (NSST) Analysis. Fix to cloud water increment bugs. Changes to land surface type specification for CRTM.
Log-Normal wind QC for AMVs.
Meghatropiques SAPHIR radiances, VIIRS AMVs. GOES clear-air water vapor winds, extra GNSS-RO observations, Global Hawk dropsondes. assimilated

01/05/2018 **GDAS** GEOS-16 AMVs assimilated.
02/28/2018 **GDAS** Switch from Meteosat-10 to Meteosat-11 winds
05/30/2018 **GDAS** Assimilation of NOAA 20 advanced microwave and IR sounder data

References

Alpert, J. C., 2004: Sub-grid scale mountain blocking at NCEP. Preprints 20th Conf. on Wea. and Fcst/16th Conf. on NWP, Seattle, WA., Amer. Meteor. Soc. http://www.emc.ncep.noaa.gov/gmb/wd23ja/presentations/mb_16confNWP_P2.4/mbnwp8b.pdf

-----, 2013: Sub-grid scale gravity wave drag and mountain blocking at NCEP. NEMS/GFS Modeling Summer School, NCWCP, College Park, MD., July 20-Aug.1, 2013. https://www.slideserve.com/mervin/sub-grid-scale-gravity-wave-drag-and-mountain-blocking-at-ncep

-----, M. Kanamitsu, P. Caplan, J. Sela, G. White, and E. Kalnay, 1988: Mountain induced gravity wave drag parameterization in the NMC medium range forecast model. Preprints 8th Conf. on NWP, 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 EMC MRF. Preprints 11th Conf. on NWP, Norfolk, VA., Amer. Meteor. Soc., 322-323.

-----, D. L. Carlis, B. A. Ballish, and V. K. Kumar, 2009: Improved forecast skill using pseudo-observations in the NCEP GFS. 13th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans and Land Surface (IOAS-AOLS), Phoenix, AZ, 2009. Amer. Meteor. Soc. http://www.emc.ncep.noaa.gov/gmb/wd23ja/presentations/dropout/Alpert%20et%20al%20AMS%202009%20Final%20a.pdf

Andersson, E., and A, Hollingsworth, 1988: Typhoon bogus observations in the ECMWF data assimilation system. ECMWF Tech. Memo. 148, ECMWF, Reading, U. K.

https://www.ecmwf.int/sites/default/files/elibrary/1988/7754-typhoon-bogus-observations-ecmwf-data-assimilation-system.pdf

Arakawa, A., and Y. Mintz, 1974: The UCLA atmospheric circulation model. Dept. of Meteorology, U. California.

----- and W. H. Schubert, 1974: Interaction of a cumulus ensemble with the large-scale environment. Part I. J. Atmos. Sci., 31, 674-701. *https://doi.org/10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2*

Baker, W. E., H. E. Fleming, M. D. Goldberg, B. B. Katz, and J. C. Derber, 1995: Development and implementation of satellite temperature soundings produced interactively. TPB 422, National Weather Service, U. S. Dept. of Commerce, 8pp. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB422.pdf.

Ballish, B. A., 1981: Initialization, theory and application to the NMC spectral model. Ph. D. Dissertation, University of Maryland, 151 pp.

-----, X. Cao, E. Kalnay and M. Kanamitsu, 1992: Incremental nonlinear normal mode initialization. Mon. Wea. Rev., 120, 1723-1734. *https://doi.org/10.1175/1520-0493(1992)120<1723:INNMI>2.0.CO;2*

Barnston, A. G., and M. K. Tippett, 2013: Predictions of Nino 3.4 SST in CFSv1 and CFSv2: a diagnostic comparison. Climate Dynamics, 41, 1615-1633. https://doi.org/10.1007/s00382-013-1845-2

Bauer, P., A. Thorpe, and G. Brunet. 2015: The quiet revolution of numerical weather prediction. Nature, 525, 47-55. http://dx.doi.org/10.1038/nature14956

Behringer, D. W., 2007: The Global Ocean Data Assimilation System at NCEP. 11th Symposium on IOAS-AOLS, San Antonio, TX, Amer. Meteor. Soc., 14-18.

-----, and Y. Xue, 2004: Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. 8th Symposium on IOAS-AOLS, AMS 84th Annual Meeting, Seattle, Washington, 11-15.

Bergmann, K. H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. Mon. Wea. Rev., 107, 1423-1444. *https://doi.org/10.1175/1520-0493(1979)107<1423:MAOTAW>2.0.CO;2*

Betts, A. K., S-Y. Hong and H.-L. Pan, 1996: Comparison of NCEP-NCAR reanalysis with 1987 FIFE data. Mon. Wea. Rev., 124, 1480-1498. *https://doi.org/10.1175/1520-0493(1996)124<1480:CONNRW>2.0.CO;2*

Bjerknes, V., 1904: Das Problem der Wettervorhesage, Betrachtet von Standpunkte der Mechanik und der Physick (The problem of weather prediction, considered from the viewpoints of mechanics and physics). Meteor. Zeitschrift., 21, 1-7. (translated and edited by Volken, E., and S. Bronniman – Meteorol. Zeitschrift.18 (2009), 663–667). https://www.researchgate.net/publication/240669114_The_problem_of_weather_prediction_considered_from_the_viewpoints_of_mechanics_and_physics

Bonner, W. D., 1989: NMC Overview: Recent Progress and Future Plans. Wea. and Fcst., 4, 275-285. *https://doi.org/10.1175/1520-0434(1989)004<0275:NORPAF>2.0.CO;2*

Bosart, L., 2003: Whither the weather analysis and forecasting process? Wea. and Fcst., 18, 520-529. *https://doi.org/10.1175/1520-0434(2003)18<520:WTWAAF>2.0.CO;2*

Bourke, W., 1974: A multi-level spectral model. I. Formulation and hemispheric integrations. Mon. Wea. Rev., 102, 687-701. *https://doi.org/10.1175/1520-0493(1974)102<0687:AMLSMI>2.0.CO;2*

Bouttier, F., J. Derber, and M. Fisher, 1997: The 1997 revision of the J_b term in 3D/4D-var. ECMWF Research Dept. Tech. Memo 238, available from ECMWF, Shinfield Park, Reading RG29AX, UK https://www.ecmwf.int/sites/default/files/elibrary/1997/8349-1997-revision-jb-term-3d4d-var.pdf

Bristor, C. L. (Ed.), 1975: Central processing and analysis of geostationary satellite data. NOAA Tech. Memo. NESS 64, 155 pp.

Bryan, K., 1969: A numerical method for the study of the World Ocean. J. Comput. Phys., 4, 347-376. https://doi.org/10.1016/0021-9991(69)90004-7

Campana, K., and P. Caplan, eds., 2005: Technical Procedures Bulletin for the T382 Global Forecast System, 9 pp plus 4 figs.

http://www.emc.ncep.noaa.gov/gc_wmb/Documentation/TPBoct05/T382.TPB.FINAL.htm

-----, -----, and J. Sela, 1989: Changes to the Global Forecast Model on November 30, 1988. TPB 383, National Weather Service, U. S. Dept. of Commerce.

http://www.emc.ncep.noaa.gov/users/GFS_history/TPB383.pdf

-----, -----, G. H. White, S. K. Yang and H. M. Juang, 1990: Impact of changes to cloud parameterization on the forecast error of NMC's global model. *7th Conference on Atmospheric Radiation: July 23-27, 1990, San Francisco, Calif*. Boston, Mass: American Meteorological Society, J152-J158.

-----, Y.-T. Hou, K. E. Mitchell, S.-K. Yang, and R. Cullather, 1994: Improved diagnostic cloud parameterization in NMC's global model. Preprints, 10th Conf. on NWP, Portland, OR. Amer. Meteor. Soc., 324-325.

Caplan, P., 1985: Recent Changes in the New Medium-Range Forecast (MRF) Model. NMC Office Note 307, U.S. Dept. of Commerce. 15 pp. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0004/0140848E.pdf

-----, 1995: The 12-14 March 1993 Superstorm: Performance of the NMC Global Medium-Range Model. Bull. Amer. Meteor. Soc., 76, 201-212. *https://doi.org/10.1175/1520-0477(1995)076<0201:TMSPOT>2.0.CO;2*

----- and G. H. White, 1989: Performance of the National Meteorological Center's Medium-Range Model. Wea. and Fcst., 4, 391-400. *https://doi.org/10.1175/1520-0434(1989)004<0391:POTNMC>2.0.CO;2*

----- and S. Saha, 1998: Further changes to the 1998 NCEP operational MRF analysis/forecast system. U.S. Dept. of Commerce, TPB 450. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB450.pdf.

----- and H.-L. Pan, 2000: Changes to the 1999 NCEP operation MRF/AVN global analysis/forecast system. U. S. Dept. of Commerce, TPB 452. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB452.pdf.

-----, G. White, and B. A. Ballish, 1989: Effects of recent changes in NMC models upon global analyses and medium-range forecasts. Proceedings of the 13th Annual Climate Diagnostic Workshop. Atmospheric and Environmental Research Inc., Cambridge, MA, Oct. 32-Nov. 4, 1988. Washington, D.C: U.S. Dept. of Commerce. National Oceanic and Atmospheric Administration, National Weather Service, 416-421.

-----, J. Derber, W. Gemmill, S.-Y. Hong, H.-L. Pan, and D. Parrish, 1997: Changes to the 1995 NCEP Operational Medium-Range Forecast Model Analysis-Forecast System. Wea. and Fcst., 12, 581-594. *https://doi.org/10.1175/1520-0434(1997)012<0581:CTTNOM>2.0.CO;2*

Charney, J.G., 1949: On a physical basis for numerical prediction of large-scale motions in the atmosphere. J. Meteor., 6, 371-385. *https://doi.org/10.1175/1520-0469(1949)006<0372:OAPBFN>2.0.CO;2*

-----, 1954: Numerical prediction of cyclogenesis. Proc. Nat. Acad. Sci. U.S., 40, 99-110. https://doi.org/10.1073/pnas.40.2.99

-----, R. Fjortoft and J. von Neumann, 1950: Numerical integration of the barotropic vorticity equation. Tellus, 6, 390-318. **https://doi.org/10.1111/j.2153-3490.1950.tb00336.x**

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q.-Y.Duan, M. Ek, and A. Betts, 1996: Modeling of land-surface evaporation by four schemes and comparison with FIFE observations. J. Geophys. Res., 101, No. D3, 7251-7268. **https://doi.org/10.1029/95JD02165**

Chou, M.-D., 1992: A solar radiation model for use in climate studies. J. Atmos. Sci., 49, 762-772. *https://doi.org/10.1175/1520-0469(1992)049<0762:ASRMFU>2.0.CO;2*

----- and K.-T. Lee, 1996: Parameterization for the absorption of solar radiation by water vapor and ozone. J. Atmos. Sci., 53, 1203-1208. *https://doi.org/10.1175/1520-0469(1996)053<1203:PFTAOS>2.0.CO;2*

----- and M. J. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Rep. TM-1994-104606 Series on Global Modeling and Data Assimilation, 85 pp. http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=D630AB274A85A6A1D764699AFB3890F2?doi=10.1.1.26.4850&rep=rep1&type=pdf

-----, -----, C. H. Ho, M. M. H. Yan, and K. T. Lee, 1998: Parameterizations for cloud overlapping and shortwave single scattering properties for use in general circulation and cloud ensemble models. J. Clim., 11, 202-214. *https://doi.org/10.1175/1520-0442(1998)011<0202:PFCOAS>2.0.CO;2*

Chun, H.-Y., and J.-J. Baik, 1998: Momentum flux by thermally induced internal gravity waves and its approximation for large-scale models. J. Atmos. Sci., 55, 3299-3310. *https://doi.org/10.1175/1520-0469(1998)055<3299:MFBTII>2.0.CO;2*

-----, M.-D. Song, J.-W. Kim, and J.-J. Baik, 2001: Effects of gravity wave drag induced by cumulus convection on the atmospheric general circulation. J. Atmos. Sci., 58, 302-319. *https://doi.org/10.1175/1520-0469(2001)058<0302:EOGWDI>2.0.CO;2*

-----, I.-S. Song, J.-J. Baik, and Y.-J. Kim, 2004: Impact of a convectively forced gravity wave drag parameterization in NCAR CCM3. J. Climate, 17, 3530-3547. *https://doi.org/10.1175/1520-0442(2004)017<3530:IOACFG>2.0.CO;2*

Clough, S. A., M. J. Iacono, and J.-L. Moncet, 1992: Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor. J. Geophys. Res., 97, 15761-15785. **https://doi.org/10.1029/92JD01419**

-----, 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. https://doi.org/10.1016/j.jqsrt.2004.05.058

Collins, W. G., 1998: Complex quality control of significant level rawinsonde temperatures. J. Atmos. Ocean. Tech., 15, 69-79. *https://doi.org/10.1175/1520-0426(1998)015<0069:CQCOSL>2.0.CO;2*

----- and L. S. Gandin, 1990: Comprehensive hydrostatic quality control at the National Meteorological Center. Mon. Wea. Rev., 118, 2752-2767. *https://doi.org/10.1175/1520-0493(1990)118<2752:CHQCAT>2.0.CO;2*

Cox, M. D., 1984: A primitive, 3-dimensional model of the ocean. GFDL Ocean Group Tech. Rep. No. 1, Geophysical Fluid Dynamics Lab., 143 pp. https://www.gfdl.noaa.gov/wp-content/uploads/files/model_development/ocean/gfdl_tech_report_1_cox_1984.pdf

Cressman, G. P., 1980: A cautionary note on the use of NMC vertical velocity forecasts. NMC Office Note 213, U. S. Dept. of Commerce, 2 pp plus 4 figs. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0003/014081F0.pdf

Cucurull, L., 2010: Improvement in the use of an operational constellation of GFDS radio occultation receivers in weather forecasting. Wea. and Fcst., 25, 749-767. *https://doi.org/10.1175/2009WAF2222302.1*

-----, J. C. Derber, R. Treadon, and R. J. Purser, 2007: Assimilation of global positioning system radio occultation observations into NCEP's Global Data Assimilation System. Mon. Wea. Rev., 135, 3174–3193. *https://doi.org/10.1175/MWR3461.1*

-----, J. C. Derber, and R. J. Purser, 2012: A bending angle forward operator for global positioning system radio occultation measurements. J. Geophys. Res. Atmos., 118, 14-28, **https://doi.org/10.1029/2012JD017782**.

Deaven, D., G. DiMego, B. Ballish, L. Morone, J. Thiebaux, P. Julian, and R. Petersen, 1990: Analysis changes in the Global Data Assimilation System on November 30, 1988. TPB 384, National Weather Service, U. S. Dept. of Commerce, 8 pp. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB384.pdf.

Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. Mon. Wea. Rev., 126, 2287-2299. *https://doi.org/10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2*

----- and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. Tellus A, 51: 195-221. https://doi.org/10.3402/tellusa.v51i2.12316

-----, D. F. Parrish, and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. Wea. and Fcst., 6, 538-547. *https://doi.org/10.1175/1520-0434(1991)006<0538:TNGOAS>2.0.CO;2*

-----, H.-L. Pan, J. Alpert, P. Caplan, G. White, M. Iredell, Y.-T. Hou, K. Campana, and S. Moorthi, 1998: Changes to the 1998 NCEP operational MRF model analysis/forecast system. U.S. Dept. of Commerce TPB 449, 37 pp. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB449.pdf

-----, P. vanDelst, X.J. Su, X. Li, K, Okamoto, and R. Treadon, 2003: Enhanced use of radiance data in NCEP data assimilation systems. Intl. TOVS Study Cong. XIII, Ste. Adele, Canada, 29 Oct.-4 Nov. 2003. http://cimss.ssec.wisc.edu/itwg/itsc/itsc13/proceedings/session1/1_8_derber.pdf

Dey, C., 1989: The evolution of objective analysis methodology at the National Meteorological Center. Wea. And Fcst., 4, 297-312. *https://doi.org/10.1175/1520-0434(1989)004<0297:TEOOAM>2.0.CO;2*

-----, and L.L. Morone, 1985: Evolution of the National Meteorological Center Global Data Assimilation System: January 1982-December 1983. Mon. Wea. Rev., 113, 304-318. *https://doi.org/10.1175/1520-0493(1985)113<0304:EOTNMC>2.0.CO;2*

-----, and B. A. Ballish, 1988: 1987 Summary of Changes to NMC Operational Global Analyses. NMC Office Note 336, U. S. Dept. of Commerce, 5 pp plus 2 tables. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0005/014088BC.pdf

-----, B. A. Ballish, and P. A. Phoebus, 1987: 1985-1986 Summary of Changes to NMC Operational Analyses. NMC Office Note 327, US. Dept. of Commerce, 10 pp. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0004/01408624.pdf

-----, R. A. Petersen, B. A. Ballish, P. M. Caplan, L. L. Morone, H. J. Thiebaux, G. H. White, H. E. Fleming, A. L. Reale, and D. G. Gray, 1989: An evaluation of NESDIS TOVS physical retrievals using data impact studies. NOAA Tech. Memo. NWS NMC 69, NTIS. 25 pp plus 24 figs. https://ntrl.ntis.gov/NTRL/

http://www.emc.ncep.noaa.gov/users/GFS_history/Deyetal1989.pdf

DiMego, G. J., 1988: The National Meteorological Center Regional Analysis System. Mon. Wea. Rev., 116, 977-1000. *https://doi.org/10.1175/1520-0493(1988)116<0977:TNMCRA>2.0.CO;2*

----, 2004: NMC/NCEP supercomputers and the models that consumer them: A sordid tale of codependence. Symp. on the 50th Anniversary of Operational NWP, UMCP, College Park, MD, 15-17 June 2004. http://www.ncep.noaa.gov/nwp50/Presentations/

-----, 2010: 50+ Years of NWP: Supercomputers and the models that consumed them: A sordid tale of codependence. 50th Anniversary Celebration of Atmospheric Science at the University of Albany, Oc.t 4, 2010, Albany, N.Y.

http://www.emc.ncep.noaa.gov/users/GFS_history/DAES50th.computers.pdf

Dorian, T., M. Ek, C. Guastini, J. Han, G. Manikin, J. Meng, V. Tallapragada, H. Wei, G. White, W. Zheng, I. Jirak, A. Dean, S. Weiss. 2017: Reducing systematic errors in GFS sensible weather forecasts. 5th WGNE Workshop on Systematic Errors in Weather and Climate Models, Montreal, CA, June 19-23, 2017.

http://collaboration.cmc.ec.gc.ca/science/rpn/wgne_wse/pdfs/Tue_White.pdf

Dorman, J. L., and P. J. Sellers, 1989: A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple

biosphere model (SiB). J. Appl. Meteor., 28, 813-855. *https://doi.org/10.1175/1520-0450(1989)028<0833:AGCOAR>2.0.CO;2*

Dudhia, A., 2015: Temperature soundings. Encyclopedia of Atmos. Sci., 2nd edition, Northm G. R., J. A. Pyle, and F. Zhang, ed. Academic Press, Vol. 5, 145-156. https://www.sciencedirect.com/topics/earth-and-planetary-sciences/tovs

Edwards, E. N., 2010: A Vast Machine. The MIT Press, Cambridge, MA, 518 pp.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, R. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land-surface model advances in the NCEP operational mesoscale ETA model. J. Geophys. Res., 108, No. D22, 8852. **https://doi.org/10.1029/2002JD003296**

Eliasen, E., B. Machenhauer, and E. Rasmusen, 1970: On a numerical method for integration of the hydrodynamic equations with a spectral representation of the horizontal fields. Institute for Theoretical Meteorology, University of Copenhagen, Rep. No. 2, 35 pp.

English, S. J., R. J. Renshaw, P. C. Dibben, A. J. Smith, P. J. Rayer, C. Poulsen, F. W. Saunders, and J. R. Eyre, 2000: A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts. Quart. J. Roy. Met. Soc., 126, 2911-2931. https://doi.org/10.1002/qj.49712656915

Environmental Modeling Center, 2003: The GFS Atmospheric Model. NCEP Office Note 442, U. S. Dept. of Commerce, 14 pp. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on442.pdf

Fels, S. B., and M. D. Schwarzkopf, 1975: The simplified exchange approximation: A new method for radiative transfer calculations. J. Atmos. Sci., 32, 1475-1488. *https://doi.org/10.1175/1520-0469(1975)032<1475:TSEAAN>2.0.CO;2*

Flattery, T., 1971: Spectral models for global analysis and forecasting. Proc. Sixth AWS Technical Exchange Conf., U.S. Naval Academy, Air Weather Service Tech. Rep. 242, 42-54.

Fritsch, J. M., and C. F. Chappell, 1980: Numerical prediction of convectively-driven mesoscale pressure systems. Part I: Convective parameterization. J. Atmos. Sci., 37, 1722-1733. *https://doi.org/10.1175/1520-0469(1980)037<1722:NPOCDM>2.0.CO;2*

Gandin, L., 1963: Objective analysis of meteorological fields. Gidrometeorologicheskoe Isdatel'stvo, Leningrad. Translated from Russian, Israel Program for Scientific Translation, Jerusalem, 1965, 242 pp.

-----, 1988: Complex quality control of meteorological observations. Mon. Wea. Rev., 116, 1137-1156. *https://doi.org/10.1175/1520-0493(1988)116<1137:CQCOMO>2.0.CO;2*

-----, L. L. Morone and W. G. Collins, 1993: Two years of operational comprehensive hydrostatic quality control at the National Meteorological Center. Wea. and Fcst., 8, 57-72. *https://doi.org/10.1175/1520-0434(1993)008<0057:TYOOCH>2.0.CO;2*

Gerrity, J. P., Jr, Ed., 1986: The MRF/GDAS Pre-Implementation Test. NMC Office Note 322, U.S. Dept. of Commerce, 17 pp. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0004/01408605.pdf

-----, J. H. Ward and G. H. White, 1985: The new NMC Medium-range Forecast Model—An Introductory Note. NMC Office Note 315, US Dept. of Commerce.10 pp plus 5 figs. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0004/014085DC.pdf

Glahn, B., 2012a: The United States National Weather Service: the first 100 years. Part 1. Pilot Imaging, Rockville, MD, 65 pp. https://commons.wikimedia.org/wiki/File:The_United_States_Weather_Service,_First_100_years_by_Bob_Glahn_part_1.pdf

-----, 2012b: The United States National Weather Service: the first 100 years. Part 2. Pilot Imaging, Rockville, MD, 55 pp. https://commons.wikimedia.org/wiki/File:The_United_States_Weather_Service,_First_100_years_by_Bob_Glahn_part_2.pdf

Global Modeling Branch, 1994: Upgrade of the Global Forecast System-1994. Changes in cloud parameterization, surface physics, analysis system, and water vapor data. TPB 417, National Weather Service, U. S. Dept. of Commerce, 22 pp. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB417.pdf

Global Climate and Weather Modeling Branch, 2004: SSI analysis system 2004. NCEP Office Note 443, U. S. Dept. of Commerce, 11 pp. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on443.pdf

-----, 2016: The Global Forecast System (GFS) – Global Spectral Model (GSM). (GSM version 13.0.2). http://www.emc.ncep.noaa.gov/GFS/doc.php

Goodberlet, M. A., C. T. Swift, and J. Wilkerson, 1989: Remote sensing of ocean surface winds with the Special Sensor Microwave/Imager. J. Geophys. Res., 94, 547-555. **https://doi.org/10.1029/JC094iC10p14547**

Grell, G. A., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations. Mon. Wea. Rev., 121, 764-787. *https://doi.org/10.1175/1520-0493(1993)121<0764:PEOAUB>2.0.CO;2*

Griffies, S. M., M. J. Harrison, R. C. Pacanowski and A. Rosati, 2004: Technical guide to MOM4. GFDL Ocean Group Technical Report No. 5, 337 pp. http://data1.gfdl.noaa.gov/~arl/pubrel/o/mom4p1/src/mom4p1/doc/mom4p0_guide.pdf

Halperin, D. J., H. E. Fuelberg, R. E. Hart, and J. H. Cossuth, 2016: Verification of tropical cyclone genesis forecasts from global numerical models: Comparisons between the North Atlantic and Eastern North Pacific basins. Wea. and Fcst., 32, 947-955. *https://doi.org/10.1175/WAF-D-15-0157.1*

Han, J., and H.-L. Pan, 2011: Revision of convection and vertical diffusion schemes in the NCEP Global Forecast System. Wea. and. Fcst., 26, 520-533. *https://doi.org/10.1175/WAF-D-10-05038.1*

-----, M. L. Witek, J. Tiexeira, R. Sun, H.-L. Pan, J. K. Fletcher, and C. S. Bretherton, 2016: Implementation in the NCEP GFS of a hybrid eddy-diffusivity mass-flux (EDMF) boundary layer parameterization with dissipative heating and modified stable boundary layer mixing. Wea. and Fcst., 31, 341-352. *https://doi.org/10.1175/WAF-D-15-0053.1*

-----, W. Wang, Y. C. Kwon, S.-Y. Hong, V. Tallapragada, and F. Yang, 2017: Updates in the NCEP GFS Cumulus Convection Schemes with Scale and Aerosol Awareness. Wea. and Fcst., 32, 2005-2017. *https://doi.org/10.1175/WAF-D-17-0046.1*

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. https://repository.library.noaa.gov/view/noaa/1157

Harper, K., L. W. Uccellini, E. Kalnay, K. Carey, and L. Morone, 2007: 50th anniversary of operational numerical weather prediction. Bull. Amer. Metero. Soc., 88, 639-650. *https://doi.org/10.1175/BAMS-88-5-639*

Harshvardhan, D. A. Randall, T. G. Corsetti, and D. A. Dazlich, 1989: Earth radiation budget and cloudiness simulations with a general circulation model. J. Atmos. Sci., 46, 1922-1942. *https://doi.org/10.1175/1520-0469(1989)046<1922:ERBACS>2.0.CO;2*

Helfand, H.M., J. C. Jusem, J. Pfaentner, J. Tennenbaum and E. Kalnay, 1987: The effect of a gravity wave drag parameterization scheme on GLA fourth-order GCM forecasts. Collection of papers presented at the WMO/IUGG NWP Symposium, Tokyo, 4-8 August 1986, 729-742. https://doi.org/10.2151/jmsj1965.64A.0_729

Hirano, R. Y., 1992: The National Meteorological Center's Historical 36- (30-) Hour S1 Score Record. NMC Office Note 389, National Weather Service, U. S. Dept. of Commerce, 10pp. www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0005/014089CB.pdf

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. *https://doi.org/10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2*

Hou, D., Z. Toth, and Y. Zhu, 2006: A stochastic parameterization scheme within NCEP global ensemble forecast system. *18th Conf. on Probability and Statistics,* Atlanta, GA, Amer. Meteor. Soc., 4.5. https://ams.confex.com/ams/Annual2006/techprogram/paper_101401.htm.

-----, -----, -----, and W. Yang, 2008: Impact of a stochastic perturbation scheme on global ensemble forecast. *19th Conf. on Probability and Statistics,* New Orleans, LA, Amer. Meteor. Soc., 1.1. https://ams.confex.com/ams/88Annual/techprogram/paper_134165.htm.

-----, -----, -----, -----, and R. Wobus, 2012: A stochastic total tendency perturbation scheme representing model-related uncertainties in the NCEP Global Ensemble System. Environmental Modeling Center, NCWCP, College Park, MD. 51 pp. http://www.emc.ncep.noaa.gov/gmb/yzhu/gif/pub/Manuscript_STTP_Tellus_A_HOU-1.pdf

Hou, Y.-T., S, Moorthi and K. Campana, 2002: Parameterization of solar radiation transfer in the NCEP models. NCEP Office Note 441, U. S. Dept. of Commerce, 34 pp plus 1 fig. and 11 tables. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on441.pdf

Iacano, M. J., R. Pincus, and C. Hannay, 2006: Representing Sub-Grid Scale Cloud Variability with MonteCarlo Independent Column Approximation and RRTMG in the National Center for Atmospheric Research Community Atmosphere Model, CAM3. Sixteenth ARM Science Team Meeting Proceedings, Albuquerque, NM, March 27 - 31, 2006. https://www.arm.gov/publications/proceedings/conf16/extended_abs/iacono_mj.pdf

Iredell, M., 2013: NEMS/GFS Overview. NEMS/GFS Modeling Summer School, NCWCP, College Park, MD., July 29-August 3, 2013. https://earthsystemcog.org/projects/gfsmodelingschool/talks

-----, 2015: EMC FYQ1 Upgrade Review: GFS Upgrade. http://www.emc.ncep.noaa.gov/impdoc/GFS/GFS-CCB.pdf.

----- and P. Caplan, 1997: Four-times daily runs of the AVN model. TPB 442, National Weather Service, U. S. Dept. of Commerce, 2 pp plus 1 fig. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB442.pdf

-----, H.-L. Pan, and P. Caplan, 2002: Changes to the 2002 NCEP operational MRF/AVN Global Analysis/Forecast System. http://www.emc.ncep.noaa.gov/gmb/STATS/tpb97/TPB02/html/v3.html

Ji, M., A. Kumar and A. Leetma, 1994a: A multiseason climate forecast system at the National Meteorological Center. Bull. Amer. Meteor. Soc., 75, 569-577. *https://doi.org/10.1175/1520-0477(1994)075<0569:AMCFSA>2.0.CO;2*

-----, -----, -----, 1994b: An experimental coupled forecast system at the National Meteorological Center: Some early results. Tellus, 46A, 398-418. **https://doi.org/10.1034/j.1600-0870.1994.t01-3-00006.x**

-----, A. Leetma and V. E. Kousky, 1996: Coupled model predictions of ENSO during the 1980s and the 1990s at the National Centers for Environmental Prediction. J. Clim., 9, 3105-3120. *https://doi.org/10.1175/1520-0442(1996)009<3105:CMPOED>2.0.CO;2*

Johansson, A., 2008: Convectively forced gravity wave drag in the CFS and GFS. EMC Seminar, Jan. 20, 2008. http://www.emc.ncep.noaa.gov/seminars/index.html

Juang, H.-M. H., 2004: A reduced spectral transform for the NCEP seasonal forecast global spectral atmospheric model. Mon. Wea. Rev., 132, 1019-1035. *https://doi.org/10.1175/1520-0493(2004)132<1019:ARSTFT>2.0.CO;2*

-----, 2005: Discrete generalized hybrid vertical coordinates by a mass, energy and angular momentum conserving vertical finite-differencing scheme. NCEP Office Note 445, National Weather Service, U. S. Dept. of Commerce, 33 pp. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on445.pdf

-----, 2016: Improving spherical harmonic coefficients for the fine resolution NCEP global atmospheric spectral modeling. Working Group on Numerical Experimentation Blue Book 2016, 6-05. http://bluebook.meteoinfo.ru/uploads/2016/individual-articles/06_Juang_Hann_Ming_Henry_Numerics.pdf

Kalnay, E., 2003: Atmospheric modeling, data assimilation and predictability. Cambridge University Press, New York, NY, 368 pp.

-----, M. Kanamitsu and W. E. Baker, 1990: Global Numerical Weather Prediction at the National Meteorological Center. Bull. Amer. Met. Soc., 71, 1410-1428. *https://doi.org/10.1175/1520-0477(1990)071<1410:GNWPAT>2.0.CO;2*

-----, S. J. Lord and R. D. McPherson, 1998: Maturity of Operational Numerical Weather Prediction: Medium Range. Bull. Amer. Met. Soc., 79, 2753-2759. *https://doi.org/10.1175/1520-0477(1998)079<2753:MOONWP>2.0.CO;2*

-----, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A, Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project. Bull. Amer. Meteor. Soc., 77, 437-471. *https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2*

Kanamitsu, M., 1989: Description of the NMC Global Data Assimilation and Forecast System. Wea. and Fcst., 4, 335-342. *https://doi.org/10.1175/1520-0434(1989)004<0335:DOTNGD>2.0.CO;2*

-----, J. C. Alpert, K. A. Campana, P. M. Caplan, D. G. Deaven, M. Iredell, B. Katz, H.-L. Pan, J. Sela, and G. H. White, 1991: Recent changes implemented into the global forecast system at NMC. Wea. and Fcst., 6, 425-435. *https://doi.org/10.1175/1520-0434(1991)006<0425:RCIITG>2.0.CO;2*

-----, A. Kumar, H.-M. H. Juang, J.-K. Schemm, W. Wang, F.Yang, S.-Y. Hong, P.Peng, W. Chen, S. Moorthi and M. Ji, 2002a: NCEP dynamical seasonal forecast system 2000. Bull. Amer. Meterol. Soc., 83, 1019-1037. *https://doi.org/10.1175/1520-0477(2002)083<1019:NDSFS>2.3.CO;2*

-----, W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002b: NCEP-DOE AMIP-II reanalysis (R-2). Bull. Amer. Meteorol. Soc., 83, 1631-1643. *https://doi.org/10.1175/BAMS-83-11-1631*

Kar, S. K., 2013: Semi-Lagrangian dynamics in GFS. NEMS/GFS Modeling Summer School, NCWCP, College Park, MD., July 29-August 3, 2013. https://earthsystemcog.org/projects/gfsmodelingschool/talks

Kessler, E., 1969: *On the Distribution and Continuity of Water Substance in Atmospheric Circulation*. *Meteor. Monogr.*, No. 32, Amer. Meteor. Soc., 84 pp. https://doi.org/10.1007/978-1-935704-36-2_1

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, D. L. Williamson, and P. J. Rasch, 1998: The National Center for Atmospheric Research community climate model CCM3. J. Climate, 11, 1131-1149. *https://doi.org/10.1175/1520-0442(1998)011<1131:TNCFAR>2.0.CO;2*

Kim, Y.-J., and A. Arakawa, 1995: Improvement of orographic gravity-wave parameterization using a mesoscale gravity wave drag model. J. Atmos. Sci., 52, 1875-1902. *https://doi.org/10.1175/1520-0469(1995)052<1875:IOOGWP>2.0.CO;2*.

Kistler, R. E., and D. F. Parish, 1982: Evolution of the NMC data assimilation system: September 1978-January 1982. Mon. Wea. Rev., 110, 1335-1346. *https://doi.org/10.1175/1520-0493(1982)110<1335:EOTNDA>2.0.CO;2*

-----, E, Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne and M. Fiorino, 2001: The NCEP-NCAR 50 year reanalysis Monthly mean CD-ROM and Documentation. Bull. Amer. Meteor. Soc., 82, 247-267. *https://doi.org/10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2*

Kleist, D. T., 2011: Assimilation of tropical cyclone advisory minimum sea level pressure in the NCEP Global Data Assimilation System. Wea. and Fcst., 26, 1085-1091. *https://doi.org/10.1175/WAF-D-11-00045.1*

----- and K. Ide, 2015: An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. Mon. Wea. Rev., 143, 452-470. *https://doi.org/10.1175/MWR-D-13-00350.1*

-----, D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009a: Improving incremental balance in the GSI 3DVAR analysis system. Mon. Wea. Rev., 117, 1046-1060. *https://doi.org/10.1175/2008MWR2623.1*

-----, -----, -----, -----. W.-S. Wu, and S. Lord, 2009b: Introduction of the GSI into the NCEP Global Data Assimilation System. Wea. and Fcst., 24, 1691-1705. *https://doi.org/10.1175/2009WAF2222201.1*

Koren, V., J. Schaake, K. Mitchell, Q.-Y. Duan, F. Chen, and J. Baker, 1999: A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. J.Geophys. Res., 104, No. D16, 19569-19585. https://doi.org/10.1029/1999JD900232

Krashnopolsky, V., W. Gemmill, and L. Breaker, 1995: Improved SSM/I wind speed retrievals at high wind speed. NCEP Tech. Note, OMB Contribution 111, 46 pp. http://polar.ncep.noaa.gov/mmab/papers/tn111/omb111.pdf

Kuligowski, R. J., R. Treadon, and W. Chen, 2002: Impacts of improved error analysis on the assimilation of polar satellite passive microwave precipitation estimates into the NCEP Global Data Assimilation System. 1st International Precipitation Working Group Workshop, Madrid, Spain, 23-27 Sept. 2002. http://www.isac.cnr.it/~ipwg/meetings/madrid-2002/pdf/kuligowski.pdf

Kuo, H. L., 1965: On formation and intensification of tropical cyclones through latent heat release by cumulus convection. J. Atmos. Sci., 22, 40-63. *https://doi.org/10.1175/1520-0469(1965)022<0040:OFAIOT>2.0.CO;2*

-----, 1974: Further studies of the parameterization of the influence of cumulus convection on large-scale flow. J. Atmos. Sci., 31, 1231-1240. *https://doi.org/10.1175/1520-0469(1974)031<1232:FSOTPO>2.0.CO;2*

Kurihara, Y., M. A. Bender, and R. J. Ross, 1993: An initialization scheme of hurricane models by vortex specification. Mon. Wea. Rev., 121, 2030-2045. *https://doi.org/10.1175/1520-0493(1993)121<2030:AISOHM>2.0.CO;2*

-----, -----, R. T. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL Hurricane Prediction System. Mon. Wea. Rev., 123, 2791-2801. *https://doi.org/10.1175/1520-0493(1995)123<2791:IITGHP>2.0.CO;2*

Lacis, A., and J. E. Hansen, 1974: A parameterization for the absorption of solar radiation in the earth's atmosphere. J. Atmos. Sci., 31, 118-133. *https://doi.org/10.1175/1520-0469(1974)031<0118:APFTAO>2.0.CO;2*

Lander, J., and B. J. Hoskins, 1997: Believable scales and parameterizations in a spectral transform model. Mon. Wea. Rev., 125, 292-303. *https://doi.org/10.1175/1520-0493(1997)125<0292:BSAPIA>2.0.CO;2*

Lazo, J. K., R. E. Morss, and J. L. Demuth, 2009: 300 billion served: Sources, perceptions, uses, and values of weather forecasts. Bull. Amer. Meteor. Soc., 90, 785-798. *https://doi.org/10.1175/2008BAMS2604.1*

LeMarshall, J., J. Jung, J. Derber, M. Chahine, R. Treadon, S. J. Lord, M. Goldberg, W. Wolf, H. C. Liu, J. Joiner, J. Woolen, R. Todling, P. van Delst, and Y. Tahara, 2006: Improving global analysis and forecasting with AIRS. Bull. Amer. Met. Soc., 87, 891-894. *https://doi.org/10.1175/BAMS-87-7-891*

Leith, C. E., 1971: Atmospheric predictability and two-dimensional turbulence. J. Atmos. Sci., 28, 145-161. *https://doi.org/10.1175/1520-0469(1971)028<0145:APATDT>2.0.CO;2*

Leonardo, N. M., and B. A. Colle, 2017: Verification of multi-model ensemble forecasts of North Atlantic tropical cyclones. Wea. and Fcst., 32, 2083-2101. *https://doi.org/10.1175/WAF-D-17-0058.1*

Lin, S.-J., 2004: A "vertically Lagrangian" finite volume dynamic core for global models. Mon. Wea. Rev., 132, 2293-2307. *https://doi.org/10.1175/1520-0493(2004)132<2293:AVLFDC>2.0.CO;2*

Liu, Q., T. Marchok, H.-L. Pan, M. Bender and S. Lord, 2000: Improvements in hurricane initialization and forecasting at NCEP with global and regional (GFDL) models. TPB 472, National Weather Service, U. S. Dept. of Commerce, 7 pp.
http://www.emc.ncep.noaa.gov/users/GFS_history/TPB472.pdf

-----, S. Lord, N. Surgi, Y. Zhu, R. Wobus, Z. Toth, and T. Marchok, 2006: Hurricane relocation in Global Ensemble Forecast System. Preprints, 27th Conf. on Hurricanes and Tropical Meterology, Monterey, CA., Amer, Meteor. Soc., P5.13.
http://www.emc.ncep.noaa.gov/users/GFS_history/Liuetal2006.pdf

Lord, S. J., 1978: Development and observational verification of cumulus cloud parameterization. Ph. D. dissertation, Univ. Calif., Los Angeles, 359 pp.

-----, 1991: A bogussing system for vortex circulations in the National Meteorological Center global forecast model. Preprints, 19th Conf. on Hurricanes and Tropical Meteorology, Miami, Fl., May 6-10, 1991, 328-330.

Lott, F and M. J. Miller: 1997: A new subgridscale orographic drag parameterization: Its formulation and testing. QJRMS, 123, 101-127. **https://doi.org/10.1002/qj.49712353704**

Lynch, P., 2008: The ENIAC Forecasts: A Re-creation. Bull. Am. Meteor. Soc., 89, 45-55. *https://doi.org/10.1175/BAMS-89-1-45*

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model, with applications to non-linear normal mode initialization. Beitr. Phys. Atmos., 50, 253-271. https://www.researchgate.net/publication/266516443_On_the_dynamics_of_gravity_oscillations_in_a_shallow_water_model_with_application_to_normal-mode_initialization

Magnuson, L., J.-R. Bidlot, S. T. K. Lang, A. Thorpe, N. Wedi, and M Yamaguchi, 2014: Evaluation of medium-range forecasts for Hurricane Sandy. Mon. Wea. Rev., 142, 1962-1981. *https://doi.org/10.1175/MWR-D-13-00228.1*

Mahrt, L., and H.-L. Pan, 1984: A two-layer model of soil hydrology. Bound.-Layer Meteor., 29, 1-20. https://doi.org/10.1007/BF00119116

Mass, C., 2014: Is numerical weather prediction one of mankind's greatest achievements? http://cliffmass.blogspot.com/2014/12/is-numerical-weather-prediction-one-of.html

Masuda, K., T. Takshima and Y. Takayama, 1988: Emissivity of pure and sea waters for the model sea surface in the infrared window regions. Remote Sens. Environ., 24, 313-329. https://doi.org/10.1016/0034-4257(88)90032-6

McCormack, J. P., S. D. Eckermann, D. E. Siskind, and T. McGee, 2006: CHEM2D-OPP: A new gas phase photochemistry parameterization for high altitude NWP and climate models. Atmos. Chem. Phys., 6, 4943-4972. https://doi.org/10.5194/acp-6-4943-2006

McMillin, L. M., L. J. Crone, and T. J. Kleepies, 1995: Atmospheric transmittance of an absorbing gas. 5. Improvements to the OPTRAN approach. Applied Optics, 34, 8396-8399. **https://doi.org/10.1364/AO.34.008396**

McNally, A. P., J. C. Derber, W.-S. Wu, and B. B. Katz, 1998: Further changes to the 1997 NCEP operational MRF Model Analysis/Forecast System: The use of TOVS level 1-b radiances and increased vertical diffusion. U.S. Dept. of Commerce, TPB 445, 32 pp. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB445.pdf

-----, -----, -----, and -----, 2000: The use of T level-1B radiances in the NCEP SSI analysis system. Q.J.R.M.S., 126, 689-724. **https://doi.org/10.1002/qj.49712656315**

McPherson, R.D., 1980: Evolution and present status of objective analysis/assimilation at NMC. NMC Office Note 216, US Dept. of Commerce. 22pp. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0003/014081FA.pdf

-----, ed., 1988: Review of the NMC Numerical Guidance Suite in 1987 and a Preview of Changes in 1988: Part II. NWS Western Region Headquarters Technical Attachment No. 88-15, 3pp plus 9 fig. https://www.weather.gov/media/wrh/online_publications/TAs/ta8815.pdf

-----., 1994: The National Centers for Environmental Prediction: Operational climate, ocean and weather prediction for the 21st century. Bull. Amer. Meteor. Soc., 75, 363-373. *https://doi.org/10.1175/1520-0477(1994)075<0363:TNCFEP>2.0.CO;2*

-----, R.D., K.H. Bergman, R.E. Kistler, G.E. Rasch, and D.S. Gordon, 1979: The NMC operational global data assimilation system. Mon. Wea. Rev., 107, 1445-1461. *https://doi.org/10.1175/1520-0493(1979)107<1445:TNOGDA>2.0.CO;2*

Menzel, W. P., 2001: Cloud tracking with satellite imagery: From the pioneering work of Tec Fujita to the present. Bull. Amer. Meteorol. Soc., 82, 33-47. https://doi.org/10.1175/1520-0477(2001)082<0033:CTWSIF>2.3.CO;2

Mesinger, F., 1996: Improvement in quantitative precipitation forecasts with the Eta regional model at the National Centers for Environmental Prediction. Bull. Amer. Meteor. Soc., 77, 2637-2650. *https://doi.org/10.1175/1520-0477(1996)077<2637:IIQPFW>2.0.CO;2*

-----, Z. I. Janjic, S. Nickovic, D. Gavrilov, and D. G. Deaven, 1988: The step-mountain coordinate: Model description and performance for cases of Alpine lee cyclogenesis and for a case of an Appalachian redevelopment. Mon. Wea. Rev., 116, 1493-1518. *https://doi.org/10.1175/1520-0493(1988)116<1493:TSMCMD>2.0.CO;2*

-----, M. Rancic, and R. J. Purser, 2018: Numerical methods in atmospheric model. Oxford Research Encyclopedia of Climate Science, in press. http://climatescience.oxfordre.com/page/forthcoming/

Michalakes, J., and J. Whitaker, 2016: Testing performance and scaling for NOAA's gnext generation global modeling system. 17th ECMWF Workshop on High performance Computing in Meteorology, Reading, U. K., Oct. 26. 2016. https://www.ecmwf.int/sites/default/files/elibrary/2016/16813-testing-performance-and-scaling-noaas-next-generation-global-modeling-system.pdf

Mitchell, K. E., amd D. C. Hahn, 1989: Development of a cloud forecast system from the GL baseline global spectral model. Geophysics Lab, Hansom AFB, Ma. GL-TR-89-0343. http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA231595

Mittenmaier, M., N. Roberts, and S. A. Thompson, 2011: A long-term assessment of precipitation forecast skill using the Fractional Skill Score. Meteor. Appl., 20, 176-186. **https://doi.org/10.1002/met.296**

Miyakoda, K., and J. Sirutis, 1977: Comparative integrations of global models with various parameterized processes of subgrid-scale vertical transports: Description of the parameterizations. Contrib. Atmos. Phys., 50, 445-487.

----- and -----, 1986: Manual of the E-physics. Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 97 pp. https://www.gfdl.noaa.gov/bibliography/related_files/Manual_of_the_E-Physics.pdf

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono. And S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res., 102, 16663-16682. **https://doi.org/10.1029/97JD00237**

Moore, P., 2015: *The Weather Experiment: The Pioneers Who Sought to See the Future*. Farrar, Straus and Girouz, New York, N. Y., 395 pp.

Moorthi, S., and M. J. Suarez, 1992: Relaxed Arakawa-Schubert: A parameterization of moist convection for general circulation models. Mon. Wea. Rev., 120, 978-1002. *https://doi.org/10.1175/1520-0493(1992)120<0978:RASAPO>2.0.CO;2*

-----, H.-L. Pan and P. Caplan, 2001: Changes to the 2001 NCEP operational AVN/MRF Global Analysis/Forecast System. TPB 484, National Weather Service, U. S. Dept. of Commerce, 14 pp. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB484.pdf

Orzag, S. A., 1970: Transform methods for calculation of vector coupled sums: Application to the spectral form of the vorticity equations. J. Atmos. Sci., 22, 40-63.

Pacanowski, R. C., and S. M. Griffies, 1998: MOM 3.0 Manual. NOAA/GFDL, 682 pp. https://www.gfdl.noaa.gov/wp-content/uploads/files/model_development/ocean/mom3_manual.pdf

Palmer, T. N., G. J. Shutts and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave

drag parameterization. Quart. J. Roy. Meteor. Soc., 112, 1001-1039.

https://doi.org/10.1002/qj.49711247406

Pan, H.-L., 1990: A simple parameterization scheme of evapotranspiration over land for the NMC Medium-range Forecast Model. Mon. Wea. Rev., 118, 2500-2512. *https://doi.org/10.1175/1520-0493(1990)118<2500:ASPSOE>2.0.CO;2*

-----and L. Mahrt, 1987: Interaction between soil hydrology and boundary-layer development. Bound.-Layer Meteor., 38, 185-220. https://doi.org/10.1007/BF00121563

----- and W.-S. Wu, 1995: Implementing a mass-flux convective parameterization package for the NMC Medium Range Forecast Model. NMC Office Note 409, 21 pp plus 19 figs. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0005/01408A42.pdf

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. Mon. Wea. Rev., 120, 1747-1763. *https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2*

-----, -----, R. J. Purser, W.-S. Wu and Z.-X. Pu, 1997: The NCEP Global Analysis System: Recent improvements and future plans. J. Meteor. Soc. Japan, 75, 359-365. https://doi.org/10.2151/jmsj1965.75.1B_359

Pegion, P., J. Whitaker, T. Hamill, G. Bates, M. Gehne, and W. Kolcynski, 2016: Stochastic parameterization development in the NOAA/NCEP Global Forecast System. ECMWF, Reading, U. K. https://www.ecmwf.int/sites/default/files/elibrary/2016/16741-stochastic-parameterization-development-noaancep-global-forecast-system.pdf

Persson, A., 2004: The 1954 start of operational numerical prediction in Sweden. Symp. on the 50th Anniversary of Operational NWP, UMCP, College Park, MD, 15-17 June 2004. http://www.ncep.noaa.gov/nwp50/Presentations/

Peters, C., V. Gerald, P. Woiceshyn, and W. Gemmill, 1994: Operational reprocessed ERS-1 scatterometer winds: A documentation. OPC Contribution 96, U. S. Dept. of Commerce, 12 pp. http://polar.ncep.noaa.gov/mmab/papers/tn96/opc96.pdf

Philander, S. G. H., W. J. Hurlin, and A. D. Seigel, 1987: A model of the seasonal cycle in the tropical Pacific Ocean. J. Phys. Oceanogr.,17,1986-2002. *https://doi.org/10.1175/1520-0485(1987)017<1986:SOTSCO>2.0.CO;2*

Phillips, N. A., 1959: Numerical integration of the primitive equations on the hemisphere. Mon. Wea. Rev., 87, 333-345. *https://doi.org/10.1175/1520-0493(1959)087<0333:NIOTPE>2.0.CO;2*

-----, 1979: The nested grid model. NOAA Tech. Rep. NWS-22, 79 pp. http://mi.nws.noaa.gov/oh/hdsc/Technical_reports/TR22.pdf

Pierrehumbert, R. T., 1986: An essay on the parameterization of orographic gravity wave drag. Seminar/Workshop 1986 Observation, Theory and Modelling of Orographic Effects. ECMWF, Shinfield Park, Reading UK, 251-282. https://www.ecmwf.int/sites/default/files/elibrary/1986/11673-essay-parameterization-orographic-gravity-wave-drag.pdf

Putnam, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. J. Computat. Phys., 227, 55-78. https://doi.org/10.1016/j.jcp.2007.07.022

Rabier, F., 2005: Overview of global data assimilation developments in numerical weather centres. Q. J. R. Meteorol. Soc., 131, 3215-3233. **https://doi.org/10.1256/qj.05.129**

Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. J. Clim., 1, 75-86. *https://doi.org/10.1175/1520-0442(1988)001<0075:ARTGSS>2.0.CO;2*

-----, 1991: A new global sea surface temperature analysis. TPB 393, National Weather Service, U. S. Dept. of Commerce, 2pp +2 figs. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB393.pdf.

Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, 236 pp. (Reprint, with a new introduction by Sidney Chapman, Dover Publications,1965, 236 pp.)

Rood, R., A. R. Douglas, J. A. Kaye, M. A. Geller, C. Y. Chen, D. J. Allen, E. M.Larsen, E. R. Nash, and J. E. Nielsen, 1991: Three-dimensional simulations of wintertime ozone variability in the lower stratosphere. J. Geophys. Res., 96, # D3, 5055-5071. **https://doi.org/10.1029/90JD02537**

Rossby, C.-G., 1940: Planetary flow patterns in the atmosphere. Quart. J. Roy. Meteor. Soc., 66 (Suppl.), 68-87. http://empslocal.ex.ac.uk/people/staff/gv219/classics.d/Rossby-planflowQJ40.pdf

-----, and Coauthors, 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. J. Mar. Res., 2, 38-55. http://peabody.yale.edu/sites/default/files/documents/scientific-publications/jmr02-01-06-CG_ROSSBYetal.pdf

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Pena, S. Lord, G. White, W. Ebisuzaki, P. Peng, and P. Xie, 2006: The NCEP Climate Forecast System. J. Climate, 19, 3483-3517. *https://doi.org/10.1175/JCLI3812.1*

-----, S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y-T. Hou, H.-Y. Chuang, H.-M. H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. van Delst, D. Keyser, J. Derber, M. Ek, J.

Meng, H. Wei, R. Yang, S. Lord, H. van den Dool, A. Kumar, W. Wang, C. Long, M. Chelliah, Y. Xue, B. Huang, J-K. Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C.-Z. Zou, Q. Liu, Y. Chen, Y. Han, L. Cucurull, R. W. Reynolds, G. Rutledge, and M. Goldberg, 2010: The NCEP Climate Forecast System Reanalysis. Bull. Amer. Meteor. Soc, 91, 1015-1057. *https://doi.org/10.1175/2010BAMS3001.1*

-----, -----, X. Wu, J. Wang, S. Nadiga, P. Tripp, D, Behringer, Y-T, Hou, H.-Y. Chuang, M. Iredell, M. Ek, J. Meng, R. Yang, M. P. Mendez, H. van den Dool, Q. Zhang, W. Wang, M. Chen, and E. Becker, 2014: The NCEP Climate Forecast System Version 2. J. Clim., 27, 2185-2208. *https://doi.org/10.1175/JCLI-D-12-00823.1*

Schwarzkopf, M. D., and S. B. Fels, 1985: Improvements to the algorithm for computing CO2 transmissivities and cooling rates. J. Geophys. Res., 90, D6, 10541-10550. **https://doi.org/10.1029/JD090iD06p10541**

----- and -----, 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. **https://doi.org/10.1029/89JD01598**

Sela, J., 1980: Spectral modeling at the National Meteorological Center. Mon. Wea. Rev., 108, 1279-1292. *https://doi.org/10.1175/1520-0493(1980)108<1279:SMATNM>2.0.CO;2*

-----, 2009: The implementation of the sigma pressure hybrid coordinate in the GFS. NCEP Office Note 461, U. S. Dept. of Commerce, 25 pp. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on461.pdf

-----, 2010: The derivation of the sigma pressure hybrid coordinate semi-Lagrangian model equations of the GFS. NCEP Office Note 462, U. S. Dept. of Commerce, 31 pp. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on462.pdf

Shuman, F.G., 1989: History of numerical weather prediction at the National Meteorological Center. Wea. And Fcst., 4, 286-296. *https://doi.org/10.1175/1520-0434(1989)004<0286:HONWPA>2.0.CO;2*

----- and J. Hovermale, 1968: An operational six-layer primitive equation model. J. Appl. Meteor., 7, 525-547. *https://doi.org/10.1175/1520-0450(1968)007<0525:AOSLPE>2.0.CO;2*

Slingo, A., 1989: A GCM parameterization for the shortwave radiative properties of water clouds. J. Atmos. Sci., 46, 1419-1427. *https://doi.org/10.1175/1520-0469(1989)046<1419:AGPFTS>2.0.CO;2*

Slingo, J., 1987: The development and verification of a cloud prediction scheme for the ECMWF model. Quart. J. Roy. Met. Soc., 113, 899-927.

https://doi.org/10.1002/qj.49711347710

-----, 2017: Weather and climate: in the eye of the storm. FT Magazine, https://www.ft.com/content/7ce2c3cc-1f03-11e7-a454-ab04428977f9.

Stackpole, J.D., 1973: The NMC 8-layer Global and Hemispheric Primitive Equation Models (8L-GLOPEP & 8L-HEMPEP) on a Longitude-Latitude Grid. NMC Office Note 90, U.S. Dept. of Commerce, 17 pp+1 fig. http://www.ncep.noaa.gov/officenotes/NOAA-NPM-NCEPON-0001/013FD8F4.pdf

-----, 1978: The NMC 9-layer Global Primitive Equation Model on a Latitude-Longitude Grid. NMC Office Note 178, U. S. Dept. of Commerce, 31 pp. http://www.lib.ncep.noaa.gov/ncepofficenotes/files/013BA1FA.pdf

Stephens, G. L., 1984: The parameterization of radiation for numerical weather prediction and climate models. Mon. Wea. Rev.,112, 826-867. *https://doi.org/10.1175/1520-0493(1984)112<0826:TPORFN>2.0.CO;2*

Stoffelen, A., B. Becker, J. Eyre, and H. Roquet, 1994: Theoretical studies of the impact of Doppler wind lidar data: preparation of a data base. European Space Agency Contract Report, ESA Rep. CR(P)-3943, 187 pp.

Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and cloud studies with mesoscale numerical weather prediction models. Mon. Wea. Rev., 112, 826-867. *https://doi.org/10.1175/1520-0493(1989)117<1641:CACPSW>2.0.CO;2*

Surgi, N., H.-L. Pan, and S. J. Lord, 1998: Improvement of the NCEP global model over the tropics: An evaluation of model performance during the 1995 hurricane season. Mon. Wea. Rev., 126, 1287-1305. *https://doi.org/10.1175/1520-0493(1998)126<1287:IOTNGM>2.0.CO;2*

Swapna, P., M. K. Roxy, K. Aparna, K. Kalkarni, A. G. Parjeesh, K. Ashok, R. Krishnan, S. Moorthi, A. Kumar, and B. N. Goswami, 2015: Bull. Amer. Meteorol. Soc., 96, 1351-1367. *https://doi.org/10.1175/BAMS-D-13-00276.1*

Teweles, S., and H. Wobus, 1954: Verification of prognostic charts. Bull. Amer. Meteor. Soc., 35, 455-461.

Thompson, P. D., and W. L. Gates, 1956: A test of numerical prediction methods based on the barotropic and two-parameter baroclinic models. J. Meteor., 13, 127-141. *https://doi.org/10.1175/1520-0469(1956)013<0127:ATONPM>2.0.CO;2*

Tiedtke, M., 1983: The sensitivity of the time-mean large-scale flow to cumulus convection in the ECMWF model. Workshop on Cumulus Convection in Large-Scale Numerical Models. ECMWF, 297-316. https://www.ecmwf.int/sites/default/files/elibrary/1983/12733-sensitivity-time-mean-large-scale-flow-cumulus-convection-ecmwf-model.pdf

Toepfer, F., and T. Schneider, 2016: Next Generation Global Prediction System (NGGPS) Phase 2 Atmospheric Dynamic Core Evaluation. UMAC Presentation, June 22, 2016.

https://www.weather.gov/media/sti/nggps/Phase%202%20Dycore%20Evaluation%20Briefing%2022%20June%202016%20UMAC%20Final%20for%20posting%2006272016.pdf

Torn, R. D., J. S. Whitaker, P. Pegion, T. M. Hamill, and G. J. Hakim, 2015: Diagnosis of the source of GFS medium-range errors in Hurricane Sandy (2012). Mon. Wea. Rev., 143, 132-152. *https://doi.org/10.1175/MWR-D-14-00086.1*

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. Bull. Amer. Meteor. Soc., 74, 2317-2330. *https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2*

Tracton, S., 2010: A snow surprise, 10 years ago today. Capital Weather Gang, Washington Post, Jan. 25, 2010. http://voices.washingtonpost.com/capitalweathergang/2010/01/major_snowstorm_ambushes_washi.html

Tracton, M. S., K. Mo, W. Chen, E. Kalnay, R, Kistler and G. White, 1989: Dynamical Extended Range Forecasting (DERF) at the National Meteorological Center. Mon. Wea. Rev., 117, 1604-1635. *https://doi.org/10.1175/1520-0493(1989)117<1604:DERFAT>2.0.CO;2*

----- and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. Wea. and Fcst., 8, 379-398. *https://doi.org/10.1175/1520-0434(1993)008<0379:OEPATN>2.0.CO;2*

Treadon, R., H.-L. Pan, W.-S. Wu, Y. Lin, W. S. Olson, and R. J. Kuligowski, 2003: Global and regional moisture analyses at NCEP. ECMWF/GEWEX Workshop on Humidity Analysis, 8-11 July 2002, ECMWF, Shinfield Park, UK, 33-47. https://www.ecmwf.int/sites/default/files/elibrary/2003/12809-global-and-reginal-moisture-analyses-ncep.pdf

Trenberth, K. E., and J. G. Olson, 1988: An evaluation and intercomparison of global analyses from the National Meteorological Center and the European Centre for Medium Range Weather Forecasts. Bull. Amer. Meteor. Soc., 69, 1047-1057. *https://doi.org/10.1175/1520-0477(1988)069<1047:AEAIOG>2.0.CO;2*

Tribbia, J.J., and R. A. Anthes, 1987: Scientific basis of modern weather prediction. Science, 237, 493-499. DOI:10.1126/science.237.4814.493

Troen, I., and L. Mahrt, 1986: A simple model of the atmospheric boundary layer: Sensitivity to surface evaporation. Bound.-Lay. Meteor., 37, 129-148. https://doi.org/10.1007/BF00122760

Wallace, J. M., S. Tibaldi, and A. J. Simmons, 1983: Reduction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. Quart. J. Roy. Met. Soc., 109, 683-717. **https://doi.org/10.1002/qj.49710946202**

Wang, X., D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based Ensemble-Variational Hybrid Data Assimilation for NCEP Global Forecast System: Single-resolution experiments. Mon. Wea. Rev., 141, 4098-4117. *https://doi.org/10.1175/MWR-D-12-00141.1*

----- and T. Lei, 2014: GSI-based four-dimensional ensemble-variational (4DEnVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP Global Forecast System. Mon. Wea. Rev., 142, 3303-3324. *https://doi.org/10.1175/MWR-D-13-00303.1*

Wang., Z., W. Li, M. S. Peng, X. Jiang, R. McTaggart-Cowan, and C. A. Davis, 2018: Predictive skill and predictability of North Atlantic tropical cyclogenesis in different synoptic flow regimes. J. Atmos. Sci., 75, 361-378. *https://doi.org/10.1175/JAS-D-17-0094.1*

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. Tellus, 60A, 62-79. https://doi.org/10.1111/j.1600-0870.2007.00273.x

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. Mon. Wea. Rev., 136, 463-482. *https://doi.org/10.1175/2007MWR2018.1*

White, G. H., 1983: Estimates of the seasonal mean vertical velocity fields of the extratropical Northern Hemisphere. Mon. Wea. Rev., 111, 1418-1433. *https://doi.org/10.1175/1520-0493(1983)111<1418:EOTSMV>2.0.CO;2*

-----, G. Manikin, C. Guastini, and T. Dorian, 2017: The Environmental Modeling Center's Model Evaluation Group. 28th Conf. Wea. and Fcst./24th Conference on NWP. Amer. Met. Soc., Seattle, WA, Jan. 22-26, 2017.

Winton, M., 2000: A reformulated three-layer sea ice model. J. Atmos. and Ocean. Tech., 17, 525-531. *https://doi.org/10.1175/1520-0426(2000)017<0525:ARTLSI>2.0.CO;2*

Wu, X., I. Simmonds, and W. Budd, 1997: Modeling of Antarctic sea ice in a general circulation model. J. Clim., 10, 593-609. *https://doi.org/10.1175/1520-0442(1997)010<0593:MOASII>2.0.CO;2*

----- and R. Grumbine, 2009: Sea-ice in the CFS Reanalysis. 34th Ann. Clim. Diag. and Predic. Wksp., Oct. 26-30, 2009, Monterey, CA. http://met.nps.edu/climate_CDPW09/documents/POSTERS/P1.18_Wu_34th_CDPW_Oct09.pdf

Wu, W.-S., M. Iredell, S. Saha, and P. Caplan, 1997: Changes to the 1997 NCEP operational MRF model analysis/forecast system. U. S. Dept. of Commerce, TPB 443. http://www.emc.ncep.noaa.gov/users/GFS_history/TPB443.pdf

-----, R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. Mon. Wea. Rev., 130, 2905-2916. *https://doi.org/10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2*

Xu, K. M., and D. A. Randall, 1996: A semi-empirical cloudiness parameterization for use in climate models. J. Atmos. Sci., 53, 3084-3102. *https://doi.org/10.1175/1520-0469(1996)053<3084:ASCPFU>2.0.CO;2*

Yamaguchi, M., J. Ishida, H. Sato, and M. Nakagawa, 2017: WGNE intercomparison of tropical Cyclone forecasts by operational NWP models: A quarte century and beyond. Bull. Amer. Meteor. Soc., 98, 2337-2349. *https://doi.org/10.1175/BAMS-D-16-0133.1*

Yang, F., 2009: On the Negative Water Vapor in the NCEP GFS: Sources and Solution. 23rd Conference on Weather Analysis and Forecasting/19th Conference on NWP, 1-5 June 2009, Omaha, NE.

-----, 2012: Review of NCEP GFS forecast skills and major upgrades. ESRL/GSD weekly meeting, May 4, 2012. https://www.slideserve.com/villette-leclerc/review-of-ncep-gfs-forecast-skills-and-major-upgrades.

-----, 2013: Review of GFS Forecast Skills in 2012. http://www.emc.ncep.noaa.gov/gmb/STATS_vsdb/longterm/

-----, K. Mitchell, Y.-T. Hou, Y. Dai, X. Zeng, Z. Wang, and X.-Z. Liang, 2008: Dependence of land surface albedo on solar zenith angle: observations and model parameterizations. *Journal of Applied Meteorology and Climatology*. No.11, Vol **47**, 2963-2982. *https://doi.org/10.1175/2008JAMC1843.1*

-----, A. Collard, R. Treadon, and J. Derber, 2013: On the GFS low forecast skills in the Southern Hemisphere: Errors from Analysis and Model. 4th WGNE Workshop on systematic

errors, Met Office, Exeter, UK, 15-19 April 2013. https://www.dropbox.com/sh/sz9heq3b5emkd9h/ZZxsZoTL5J?n=145108712&oref=e&preview=Holistic_YangFL.pdf

Yu, T.-W., M. Iredell, and D. Keyser, 1997: Global data assimilation and forecast experiments using SSM/I wind speed data derived from a neural network application. Wea. and Fcst., 12, 859-865. *https://doi.org/10.1175/1520-0434(1997)012<0859:GDAAFE>2.0.CO;2*

Zeng, X., M. Zhao, and R. E. Dickinson, 1998: Intercomparison of bulk aerodynamic algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. J. Climate, 11, 2628-2644. *https://doi.org/10.1175/1520-0442(1998)011<2628:IOBAAF>2.0.CO;2*

Zhao, Q. Y., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP models. Mon. Wea. Rev., 125, 1931-1953. *https://doi.org/10.1175/1520-0493(1997)125<1931:APCSFO>2.0.CO;2*

Zheng, W., M. Ek, K. Mitchell, H. Wei and J. Meng, 2017: Improving the stable surface layer in the NCEP Global Forecast System. Mon. Wea. Rev., 145, 3969-3987. *https://doi.org/10.1175/MWR-D-16-0438.1*

Zhou, X., Y. Zhu, D. Hou, and D. Kleist, 2016: A comparison of perturbations from an ensemble transform and an ensemble Kalman filter for the NCEP Global Ensemble Forecast System. Wea. and Fcst., 31, 2057-2074. *https://doi.org/10.1175/WAF-D-16-0109.1*

-----, -----, -----, Y. Luo, J. Peng, and R. Wobus, 2017: Performance of the new NCEP global ensemble forecast system in a parallel experiment. Wea. and Fcst., 32, 1989-2004. *https://doi.org/10.1175/WAF-D-17-0023.1*

Zhu, Y., and Z. Toth, 2006: Global Ensemble and NAEFS. RFC Short-Term Ensemble Workshop. Nov. 30, 2006. http://www.nws.noaa.gov/oh/hrl/hsmb/docs/hep/events_announce/Zhu_GEFS_NAEFS_Final.pdf

Zhu, Y., J. C. Derber, R. J. Purser, B. A. Ballish, and J. Whiting, 2015: Variational correction of aircraft temperature bias in the NCEP's GSI analysis system. Mon. Wea. Rev., 143, 3774-3803. *https://doi.org/10.1175/MWR-D-14-00235.1*

-----, E. Liu, R. Mahajan, C. Thomas, D. Groff, R. van Delst, A. Collard, D. Kleist, R. Treadon, and J. C. Derber, 2016: All-sky microwave radiance assimilation in NCEP's GSI analysis system. Mon. Wea. Rev., 144, 4709-4735. *https://doi.org/10.1175/MWR-D-15-0445.1*

1952	IBM 701	1 kiloflop
1957	IBM 704	8 kiloflops
1960	IBM 7090	67 kiloflops
1963	IBM 7094	100 kiloflops
1966	CDC 6600	3 megaflops
1971	IBM 360-195	6 megaflops
1972	3 IBM 360-195s	18 megaflops
1981	CDC Cyber 205	400 megaflops
1987	2 CDC Cyber 205s	800 megaflops
1990	Cray Y-MP	2.6 gigaflops
1994	Cray C9016	15.3 gigaflops
1999	IBM SP2	700 gigaflops
2003	IBM SP2 upgrade	2.5 teraflops
2005	IBM P5	7 teraflops
2007	IBM P6	17 teraflops
2012	IBM X86	1.48 petaflops
2015	IBM Cray	4.09 (5.57) petaflops
2018	IBM Dell	2.84 (8.41) petaflops

Table 1. NMC/NCEP computers. Flop totals in brackets since 2012 are cumulative, reflecting the total capacity of more than one machine. Data provided by Geoff DiMego (2004) and Mike Kane (personal communication)

Change History of GFS Configurations

Mon/Year	Levels	Truncations	Z-cor/dyncore	Major components upgrade
Aug 1980	12	R30 (375km)	Sigma Eulerian	first global spectral model, rhomboidal
Oct 1983	12	R40 (300km)	Sigma Eulerian	
Apr 1985	18	R40 (300km)	Sigma Eulerian	GFDL Physics, silhouette mountains
Aug 1987	18	T80 (150km)	Sigma Eulerian	First triangular truncation, diurnal cycle
Mar 1991	18	T126 (105km)	Sigma Eulerian	Mean orography, SSI (June 1991)
Aug 1993	28	T126 (105km)	Sigma Eulerian	Arakawa-Schubert convection
Jun 1998	42	T170 (80km)	Sigma Eulerian	Prognostic ozone; SW from GFDL to NASA
Oct 1998	28	T126 (105km)	Sigma Eulerian	The restoration
Jan 2000	42	T170 (80km)	Sigma Eulerian	First on IBM
Oct 2002	64	T254 (55km)	Sigma Eulerian	RRTM LW
May2005	64	T382 (35km)	Sigma Eulerian	2L OSU to 4L NOAH LSM; high-res to 180 h
May2007	64	T382 (35km)	Hybrid Eulerian	SSI to GSI
Jul 2010	64	T574 (23km)	Hybrid Eulerian	RRTM SW; New shallow cnvtion; TVD tracer
Jan 2015	64	T1534 (13km)	Hybrid Semi-Lag	SLG; Hybrid EDMF; McICA
May 2016	64	T1534 (13km)	Hybrid Semi-Lag	4-D hybrid EN-Var DA
June 2017	64	T1534 (13km)	Hybrid Semi-Lag	NEMS GSM, improved physics
Dec 2018	64	FV3 (13km)	Finite-Volume	NGGPS FV3 dycore

Table 2: Changes in GFS resolution, dynamic core, physics and data assimilation
Provided by Fanglin Yang.

Version	Implementation	Initial uncertainty	TS relocation	Model uncertainty	Resolution	Forecast length (d)	Ensemble members	Daily frequency
V1.0	1992.12	BV	None	None	T62L18	12	2	00UTC
V2.0	1994.3				T62L18	16	10(00UTC) 4(12UTC)	00,12UTC
V3.0	2000.6				T126L28(0-2.5) T62L28(2.5-16)		10	
V4.0	2001.1				T126(0-3.5) T62L28(3.5-16)			
V5.0	2004.3				T126L28(0-7.5) T62L28(7.5-16)			00,06,12, 18UTC
V6.0	2005.8		TSR		T126L28			
V7.0	2006.5	BV- ETR					14	
V8.0	2007.3						20	
V9.0	2010.2			STTP	T190L28			
V10.0	2012.2				T254L42 (0-8) T190L42 (8-16)			
V11.0	2015.12	EnKF (f06)			**T_L574L64 (0-8) T_L382L64 (8-16)**			

Table 3: GEFS configuration 1992-2015. Provided by Yuejian Zhu.

Comparison of Vertical Structures			
Layers	64	42	28
Top (hPa)	.266	2	3
Layers above 100 hPa	24	10	7
Layers below 850 hPa	13	10	6
Thickness of lowest layer(m)	20	30	40

Table 4 Comparison of vertical structure in different versions of the GFS

	NH day 5	NH day 8	SH day 5	SH day 8
CDAS	.735	.436	.651	.359
CMC	.881	.609	.871	.577
GFS	.887	.626	.875	.599
UKMO	.898		.891	
ECMWF	.918	.679	.916	.669

Table 5 500 hPa Anomaly correlations averaged over Mar. 2017-Feb. 2018 for 5 and 8 day forecasts for NH and SH. CDAS is the T62 28 level version of the GFS of the mid-1990s used in the NCEP-NCAR reanalysis, CMC is the Canadian Meteorological Centre global model, UKMO is the Meteorological Office's global model whose forecasts extend only to 6 days, and ECMWF is the European Centre for Medium-range Weather Forecasts. Data provided by Fanglin Yang

Deterministic hi-res	**0.8 anomaly correlation**	**0.6 anomaly correlation**
CMC	5.8	7.8
GFS	6.1	8.1
ECMWF	6.6	8.5
Ensemble		
CMC	6.6	9.6
GEFS	6.7	9.5
NAEFS	6.9	9.9
EPS	7.4	10.4

Table 6 The 500 hPa forecast length in days averaged over Feb. 2017-Feb. 2018 for the NH when the anomaly correlation reaches .8 (considered the limit of good forecasts) and .6 (the limit of useful forecasts) for deterministic high-resolution models and for ensemble means. NAEFS is the combined GEFS and Canadian ensemble; EPS is the European Centre Ensemble Prediction System. Data provided by Fanglin Yang

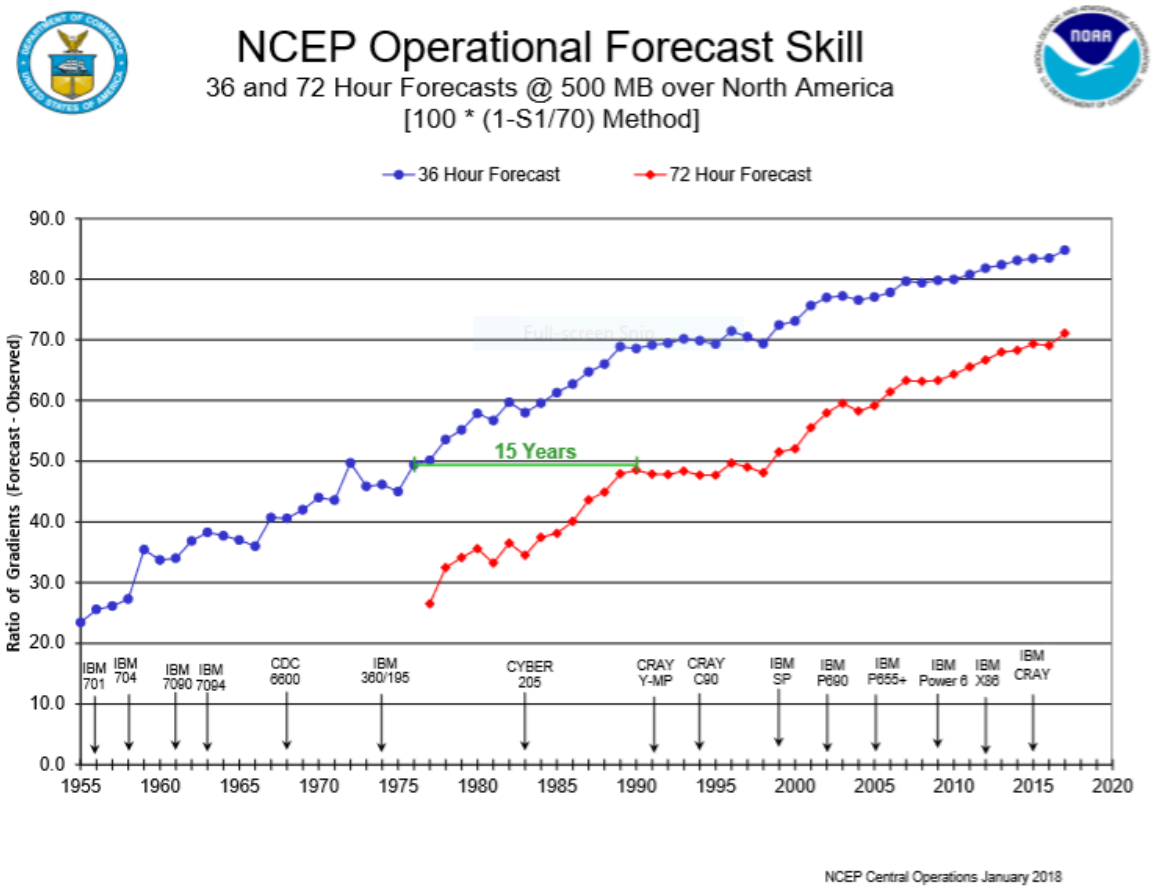


Figure 1 The annual mean S1 skill score (Twerles and Wobus, 1954; Hirano, 1992) from 1955-2017, measuring the relative error in the pressure gradient of 500 hPa height over North America. The skill of spectral GFS forecasts is shown since 1980; the skill of manual and earlier models' forecasts is shown before. An S1 score of 70 (0 on the scale of Figure1) was considered useless; an S1 score of 20 (approximately 71 in Figure 1) was deemed equivalent to the difference in hand analyses over North America drawn by different skilled synoptic meteorologists and considered a perfect forecast. The figure was produced by Steve Lilly.

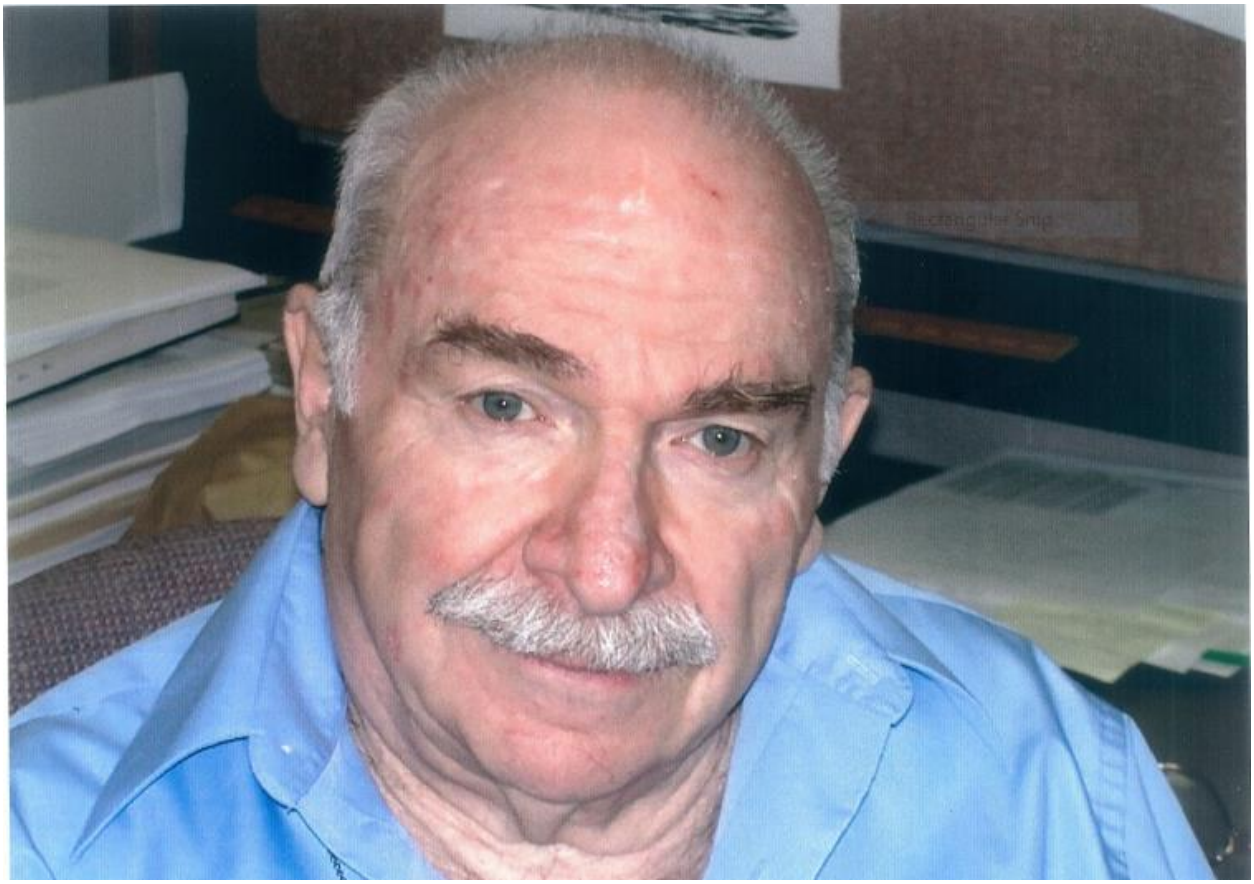


Figure 2: Dr. Joseph Sela created and developed the spectral GFS and worked on improving it from 1975 until his death in 2010 as the GFS was successfully and consistently forecasting the Feb 5-6 2010 Snowmageddon east coast snowstorm from 6 days in advance. This paper is dedicated to him.

Figure 3: Five of the eight directors of EMC. Not shown are the first two directors of the Development Divison: Dr. Fred Shuman and Dr. John Brown, and the current EMC director, Dr. Brian Gross. Shown are Drs. Henrik Tolman, Bill Lapenta, Eugenia Kalnay, Steve Lord, and Mike Farrar.

Figure 4 (top) The Development Division in the early 1990's at the World Weather Building.

(bottom) EMC in the NCWCP Auditorium in 2017. The bottom picture is by Michiko Masutani.

Annual Mean 500-hPa Height day-5 Anomaly Correlation

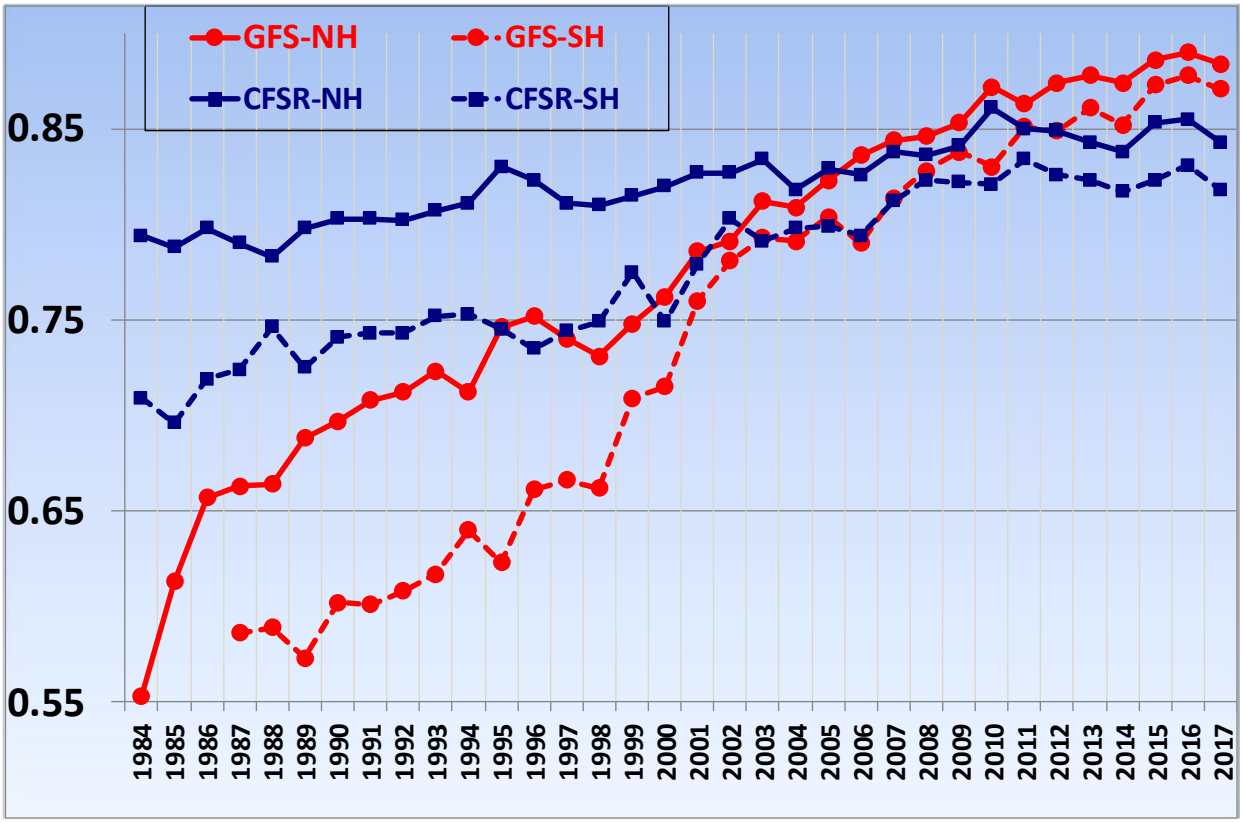


Figure 5a: Annual mean 500 hPa height anomaly correlations for 5 day forecasts in the NH and SH by the operational GFS and by the CFSR, a frozen coupled T384L64 data assimilation/T126L64 version of the operational GFS of 2010 (Saha *et al*., 2010). The CFSR shows the effects of improvement in observations; the GFS shows the effect of all improvements to the GFS since 1984. Provided by Fanglin Yang

Annual mean 500-hPa HGT Day 5 anomaly correlation GFS minus CFSR

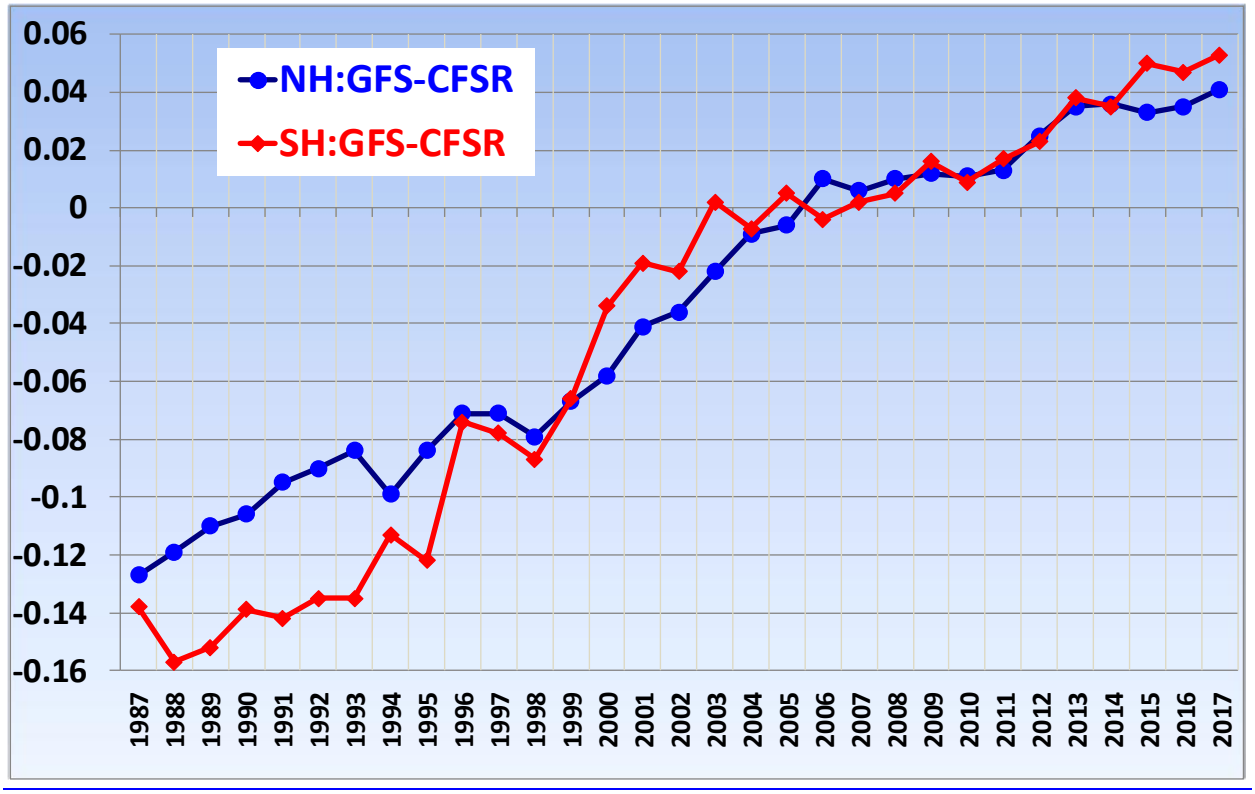


Figure 5b Annual mean difference in 500 hPa height anomaly correlation between CFSR and GFS. Provided by Fanglin Yang.

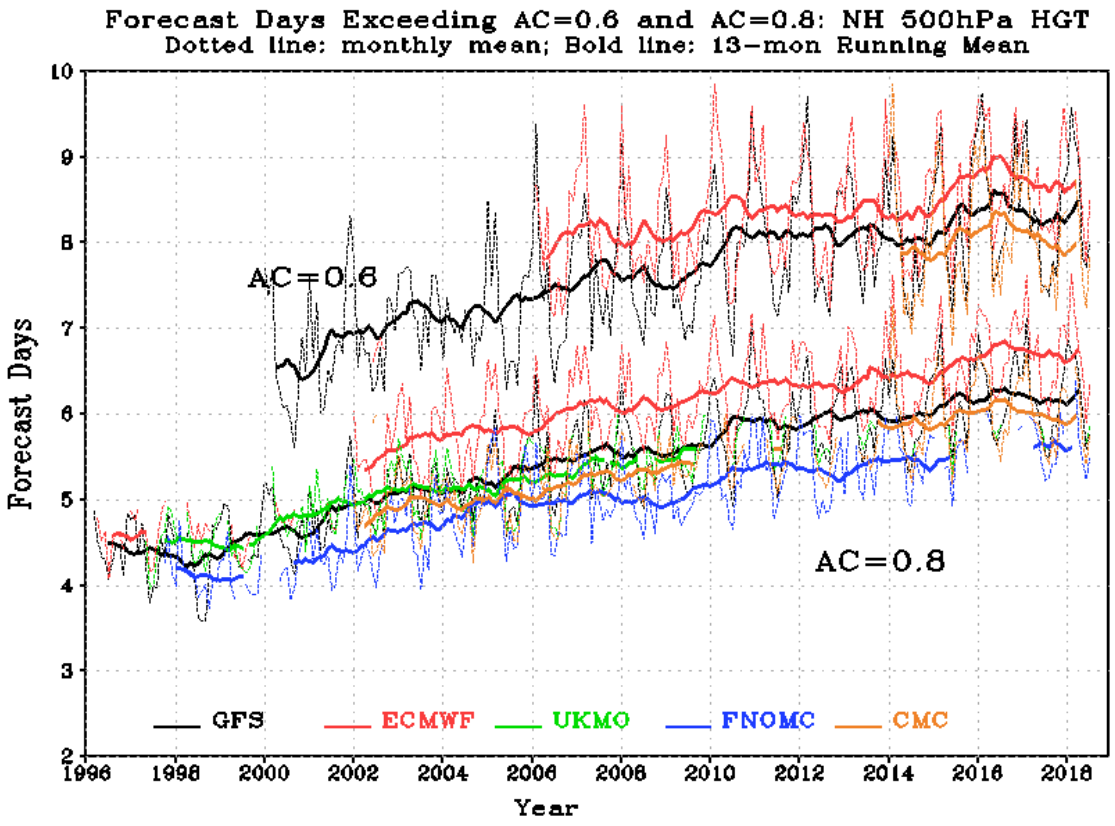


Figure 6a. The 13-month running mean length of forecast in days it takes for 500 hPa height forecasts anomaly correlations to reach .8 (the limit of good forecasts) and .6 (the limit of useful forecasts) for global forecasts by the GFS, ECMWF, United Kingdom Meteorological Office (UKMO), CMC and the U. S. Navy's Fleet Numerical Meteorology and Oceanography Center (FNOMC)'s global deterministic forecasts for NH. Thinner lines represent monthly mean values through July 2018. UKMO forecasts extend only to 6 days; currently they are somewhat better than the GFS out to 6 days. Provided by Fanglin Yang.

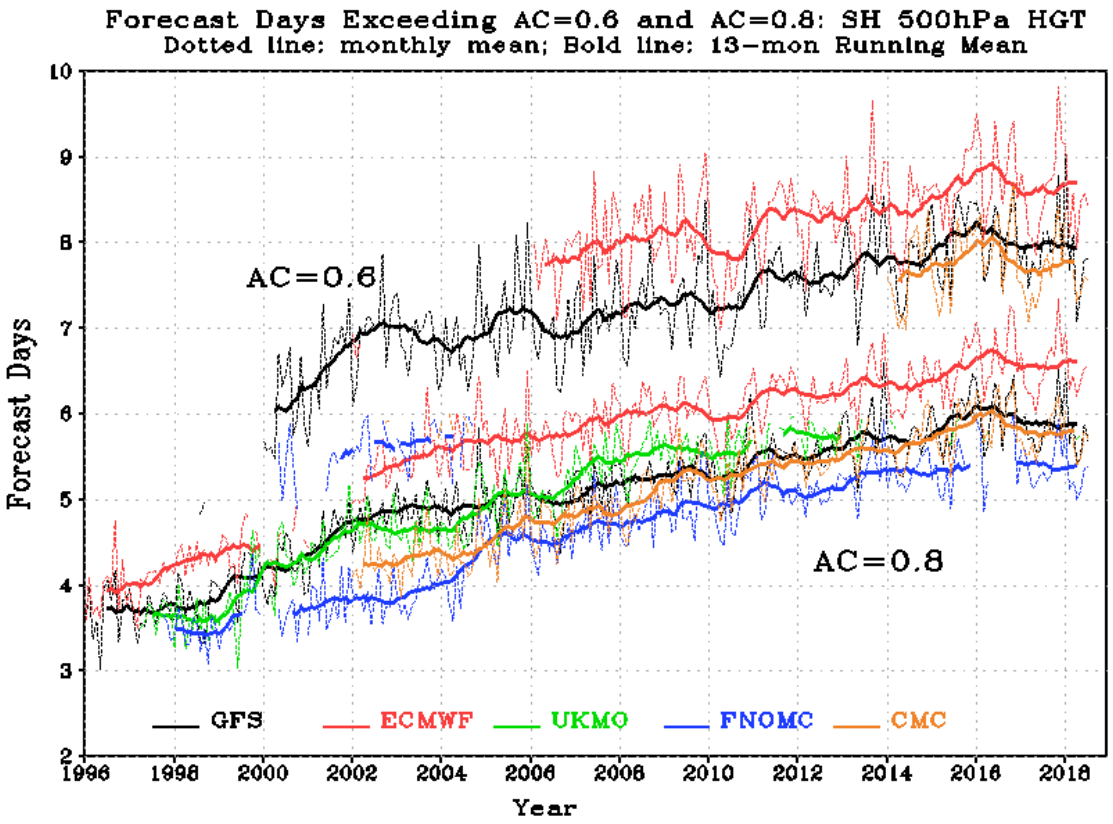


Figure 6b As in Figure 6a, except for the SH. Provided by Fanglin Yang

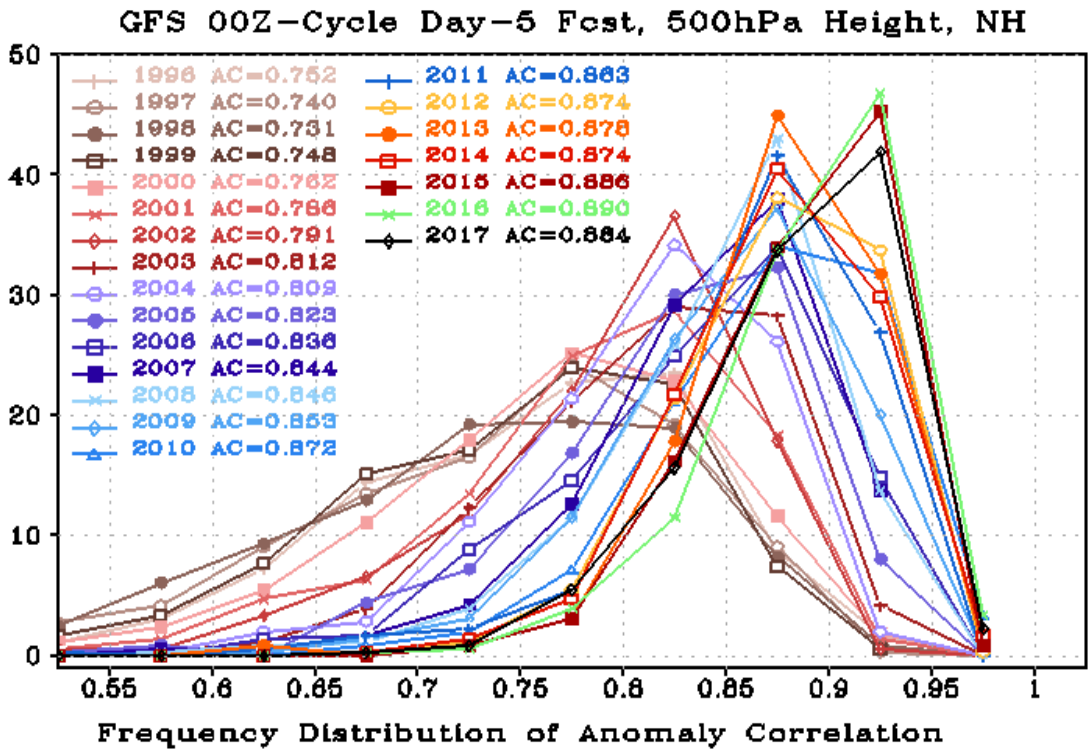


Figure 7a: Frequency distribution of daily anomaly correlations of GFS 000 UTC NH day 5 forecast of 500 hPa height by year 1996-2017. Provided by Fanglin Yang.

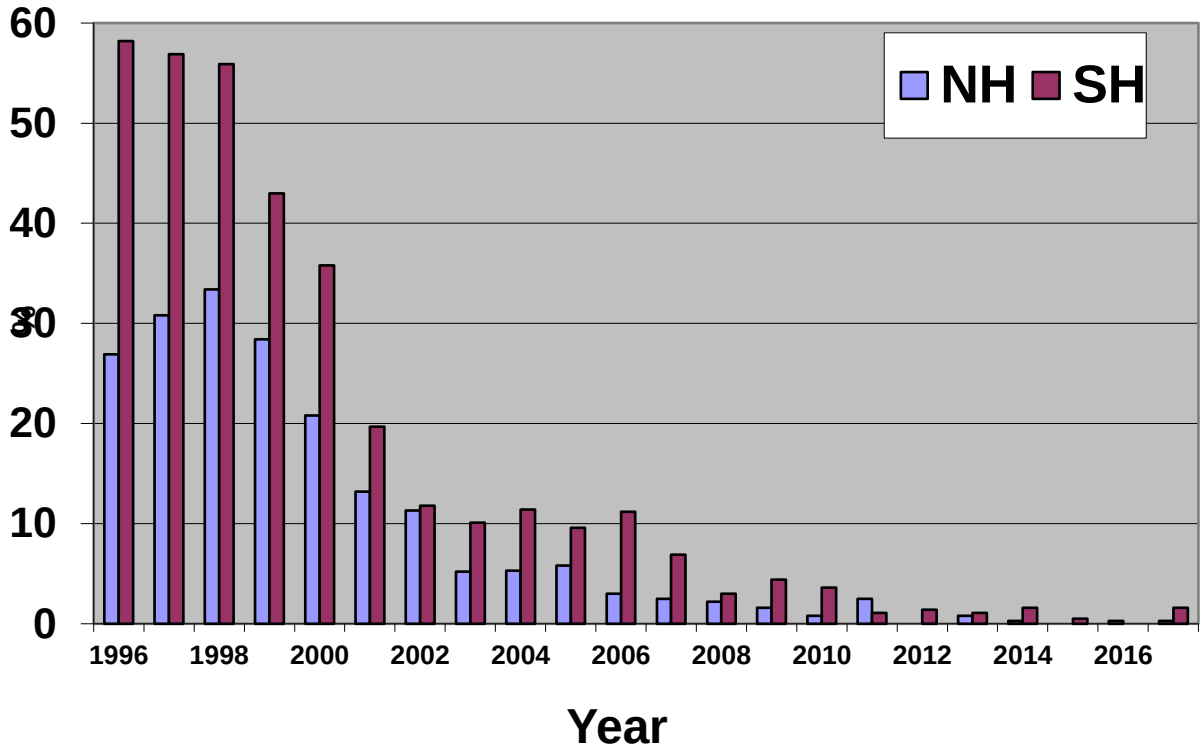


Figure 7b: Annual percent of GFS 500 hPa geopotential height forecasts with anomaly correlations below .7 from 1996 to 2017. Provided by Fanglin Yang

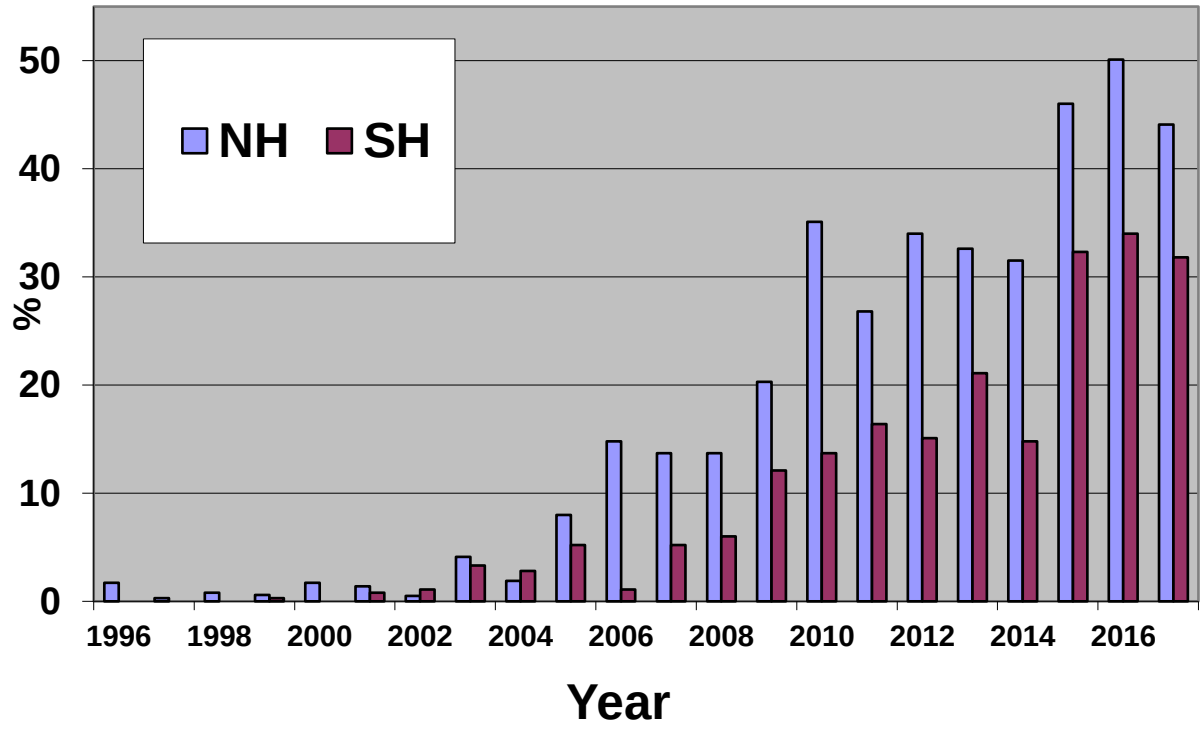


Figure 7c: Annual percent of GFS 500 hPa geopotential height forecasts with anomaly correlations above .9 from 1996 to 2017. Provided by Fanglin Yang

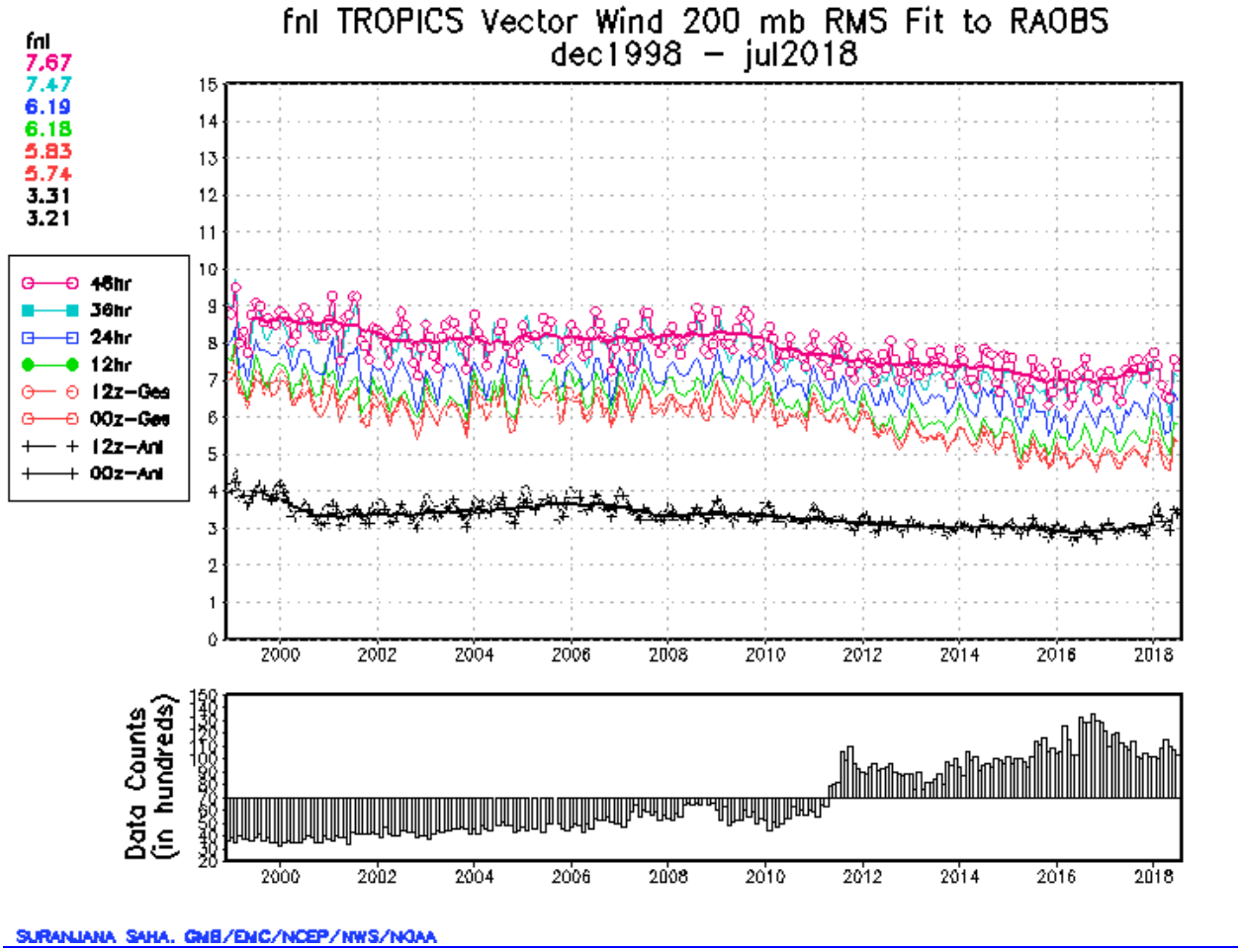


Figure 8a. RMS fit of GFS analyses and forecasts to 200 hPa winds measured by rawinsondes over the tropics. Provided by Suranjana Saha and Jack Woollen.

(http://www.emc.ncep.noaa.gov/gmb/ssaha/)

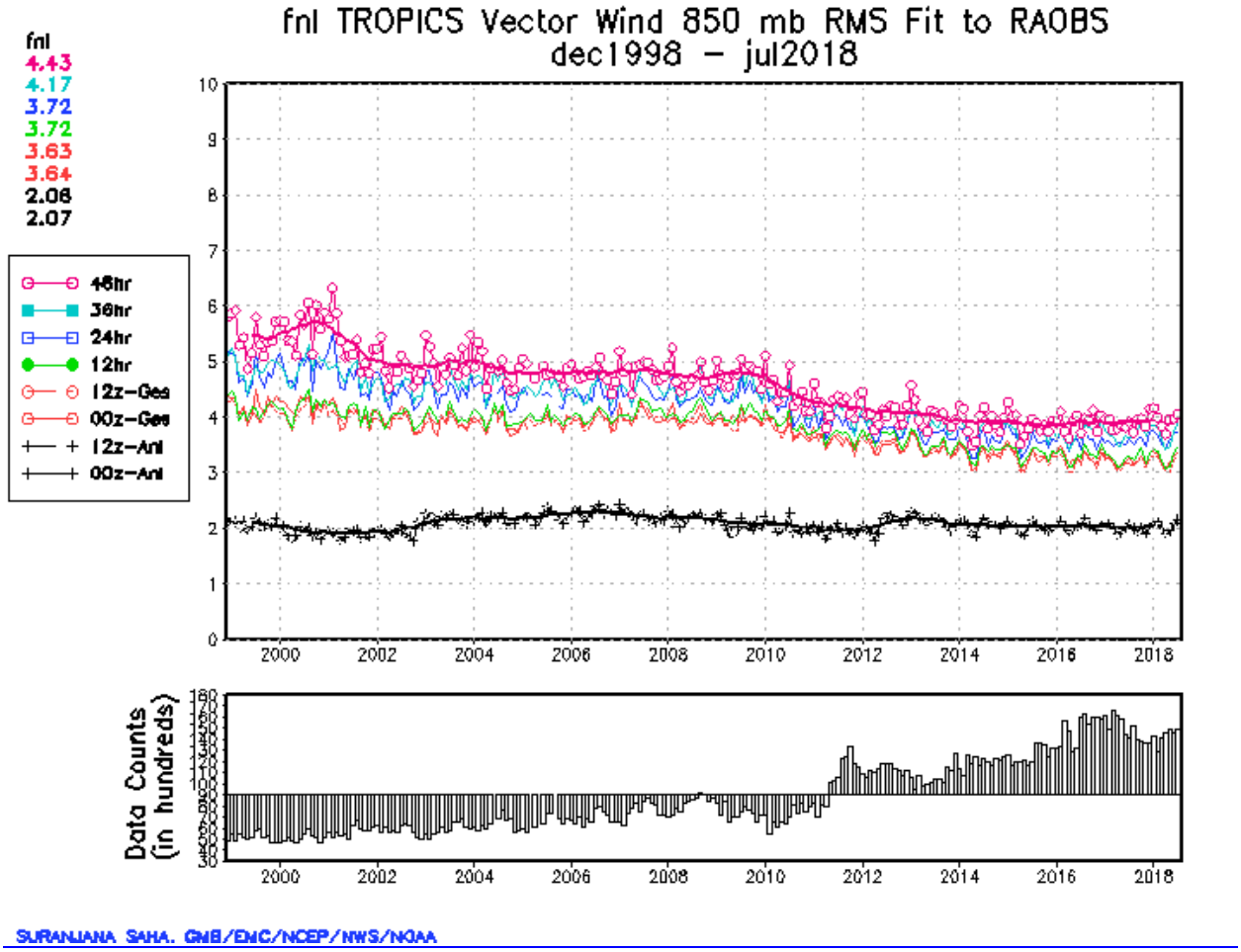
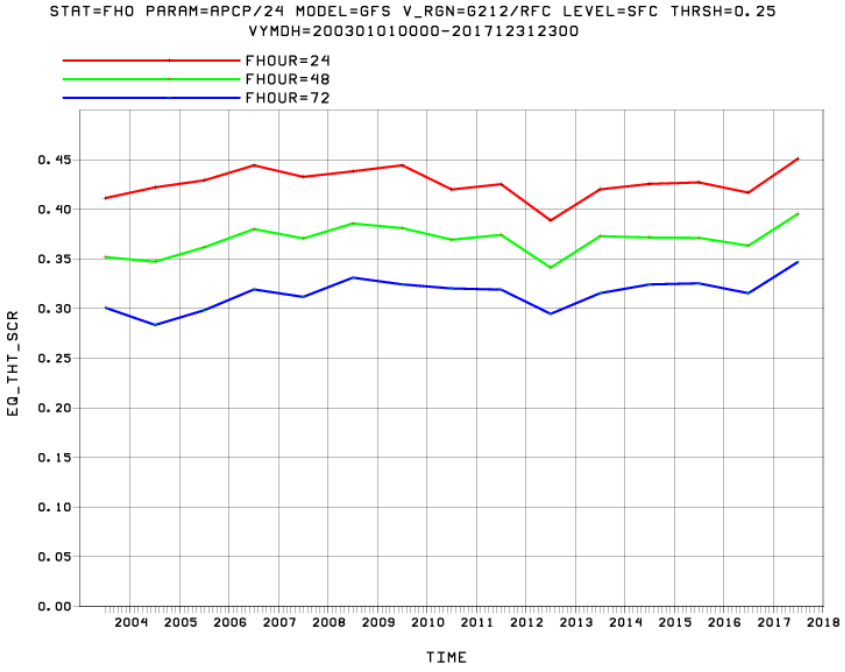
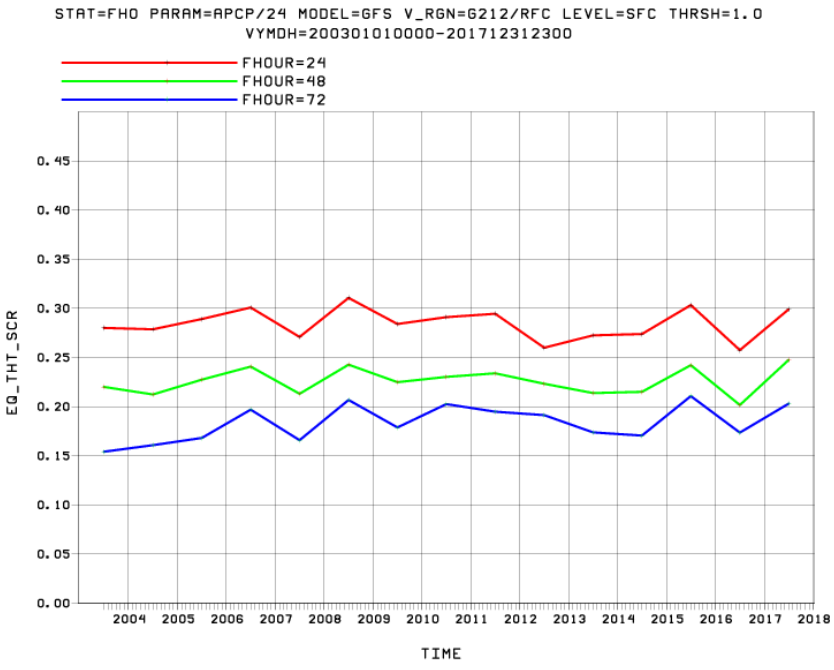


Figure 8b RMS fit of GFS analyses and forecasts to 200 hPa winds measured by rawinsondes over the tropics. Provided by Suranjana Saha and Jack Woollen.

(http://www.emc.ncep.noaa.gov/gmb/ssaha/)

GFS 24/48/72h annual ETS, 2003-2017, 40 km





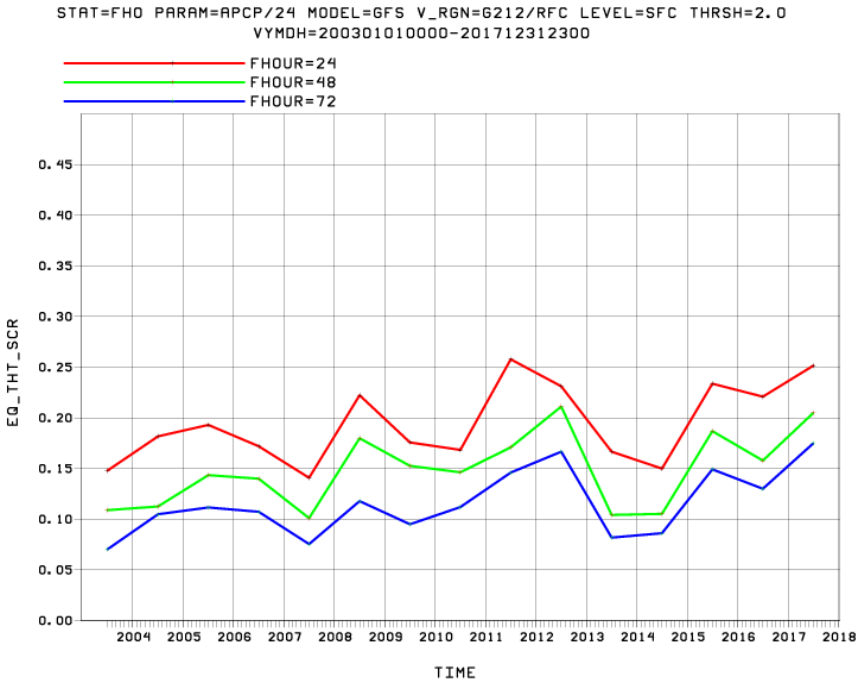


Figure 9a: Annual mean equitable threat score (Mesinger, 1966) for the continental US for GFS 24, 48 and 72 h forecasts on a 40 km grid of precipitation over the continental US for amounts of (top) 0.25, (middle) 1 and (bottom) 2 inches/day from 2003 to 2017. Provided by Ying Lin.

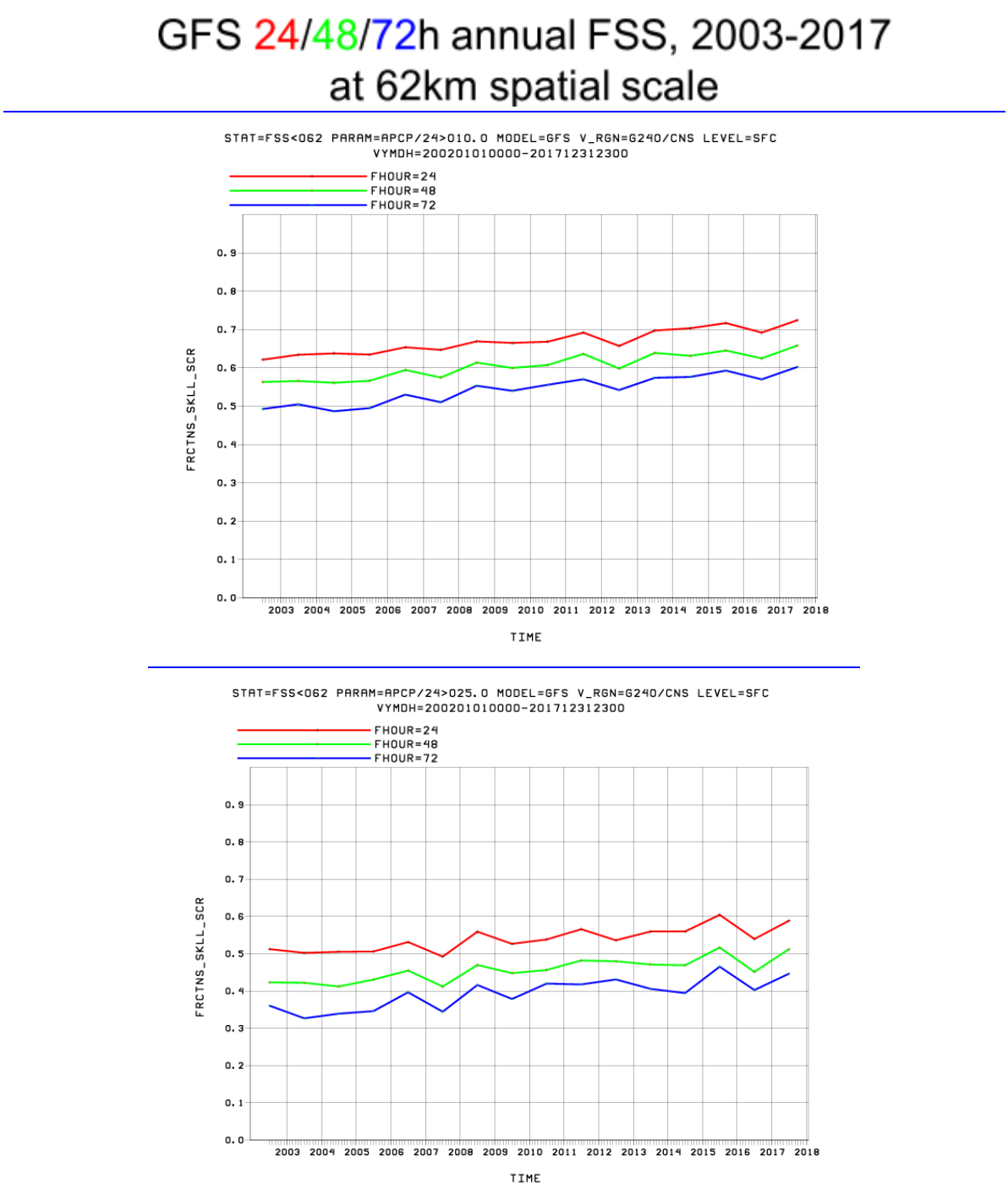


Figure 9b: Annual average fractional skill score (Mittenmaier *et al*., 2011) for GFS 1, 2 and 3 day forecasts of precipitation over the continental US for rates of (top) 10 and (bottom) 25 mm/day at a 62 km spatial scale from 2003 to 2017. Provided by Ying Lin.

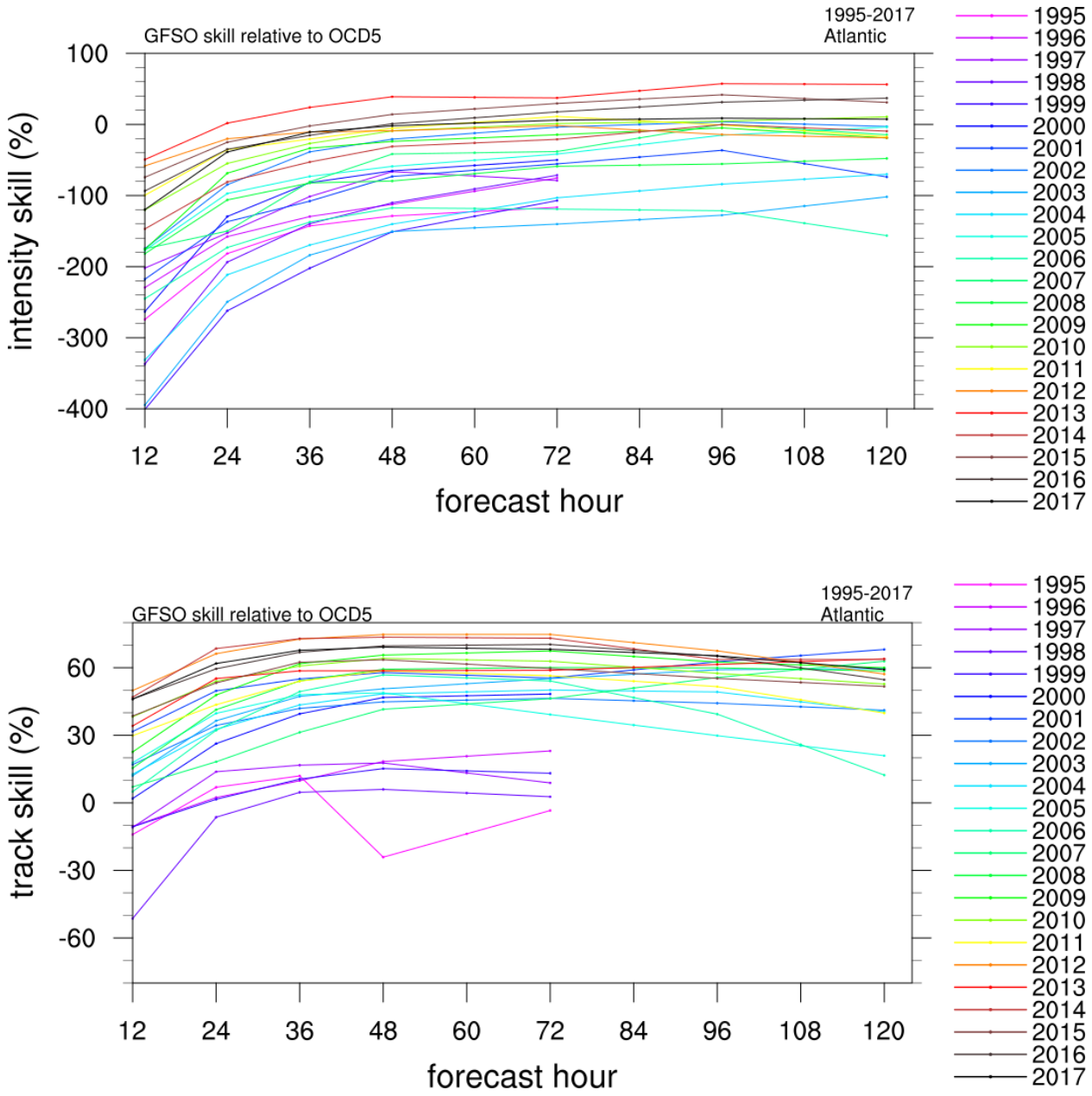


Figure 10a. (top) Intensity and (bottom) Track skill of tropical cyclone forecasts by GFS for Atlantic basin for 1995-2017 relative to skill of the NHC statistical baseline OCD5 based on persistence and climatology (https://www.nhc.noaa.gov/modelsummary.shtml). Provided by Andrew Penny.

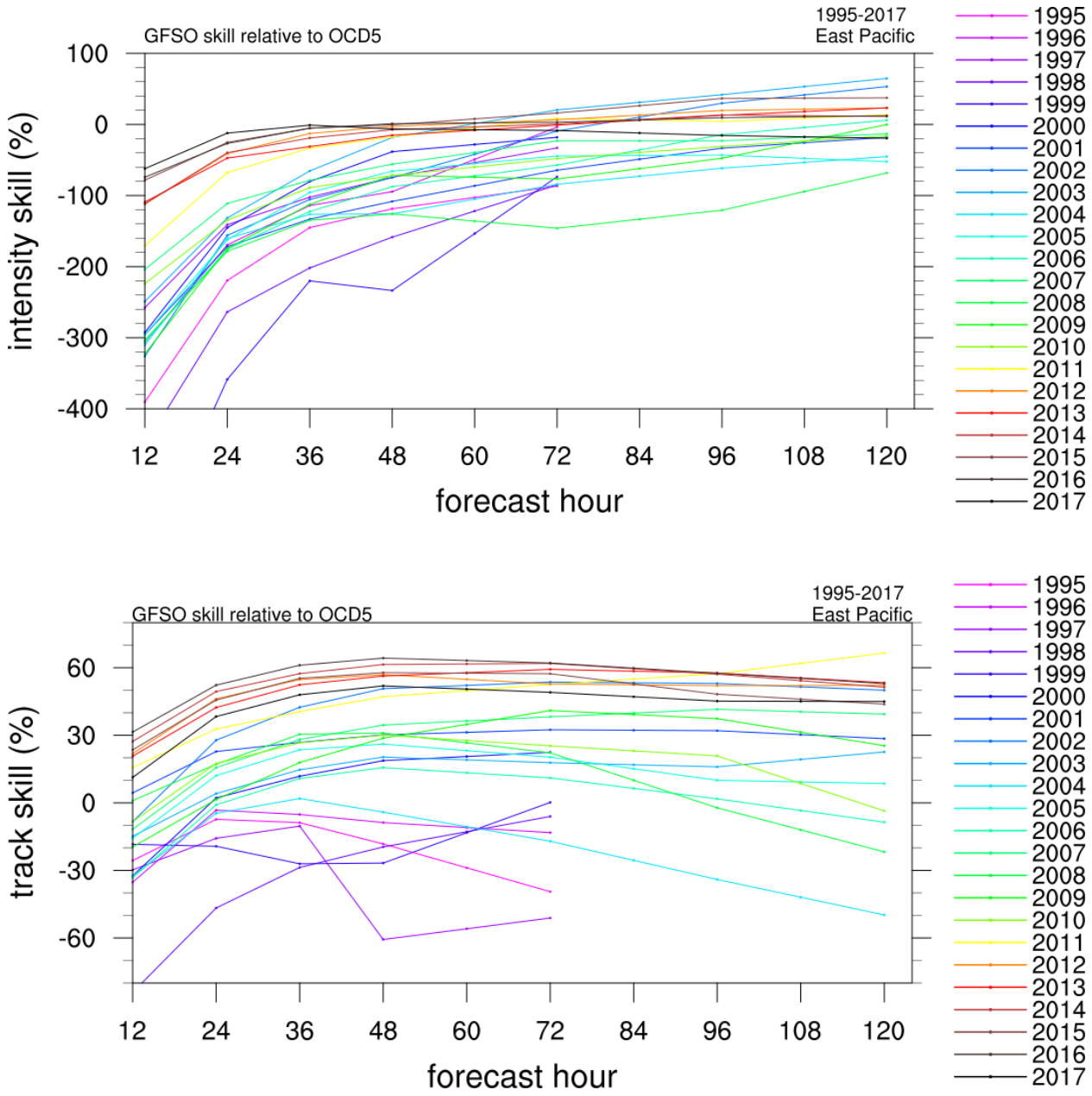


Figure 10b. (top) Intensity and (bottom) Track skill of tropical cyclone forecasts by GFS for the east Pacific basin for 1995-2017 relative to skill of the NHC statistical baseline OCD5 based on persistence and climatology (https://www.nhc.noaa.gov/modelsummary.shtml). Provided by Andrew Penny.

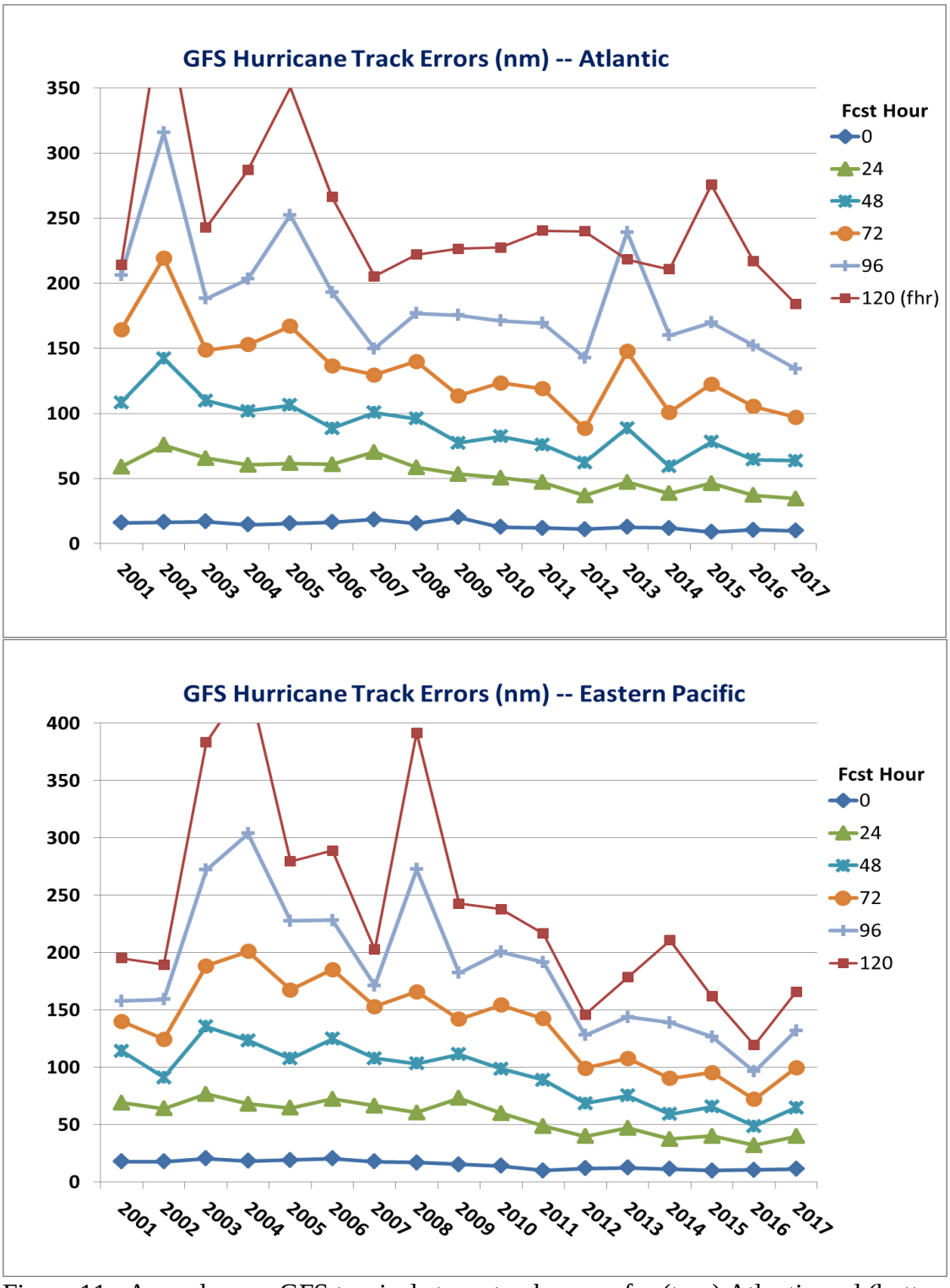


Figure 11: Annual mean GFS tropical storm track errors for (top) Atlantic and (bottom) east Pacific for 2001-2017. Furnished by Fanglin Yang.

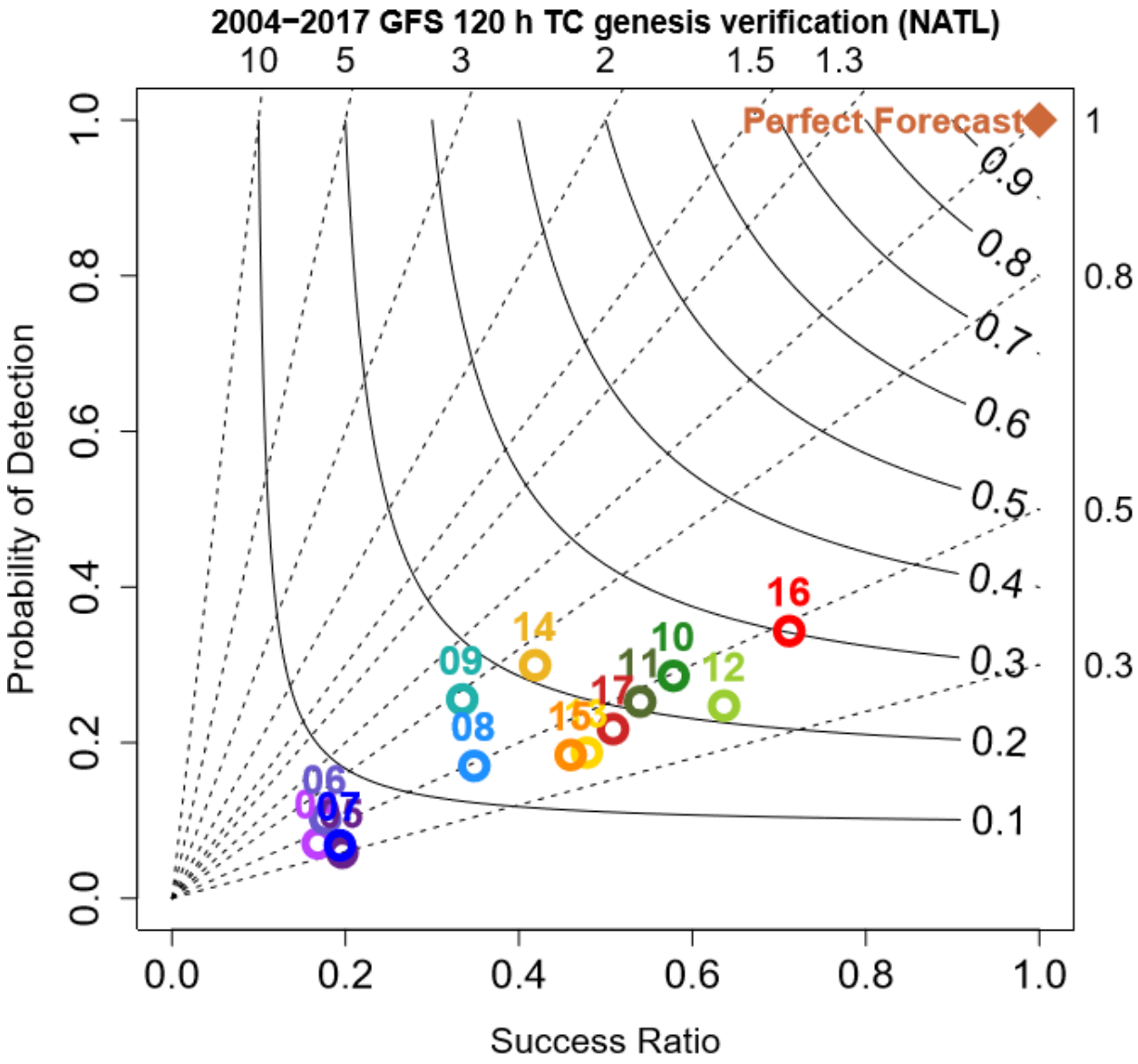


Figure 12a. Verification of GFS forecasts of tropical storm genesis in the Atlantic. Genesis events in forecast hours 6-120 verified against NHC best track analyses to determine whether genesis actually occurred within 120 h of model initial time. Probability of detection is the number of hits divided by the number of hits and misses; success ratio is 1 minus the number of false alarms divided by the number of hits and false alarms (Halperin *et al*., 2016). Furnished by Dan Halperin.

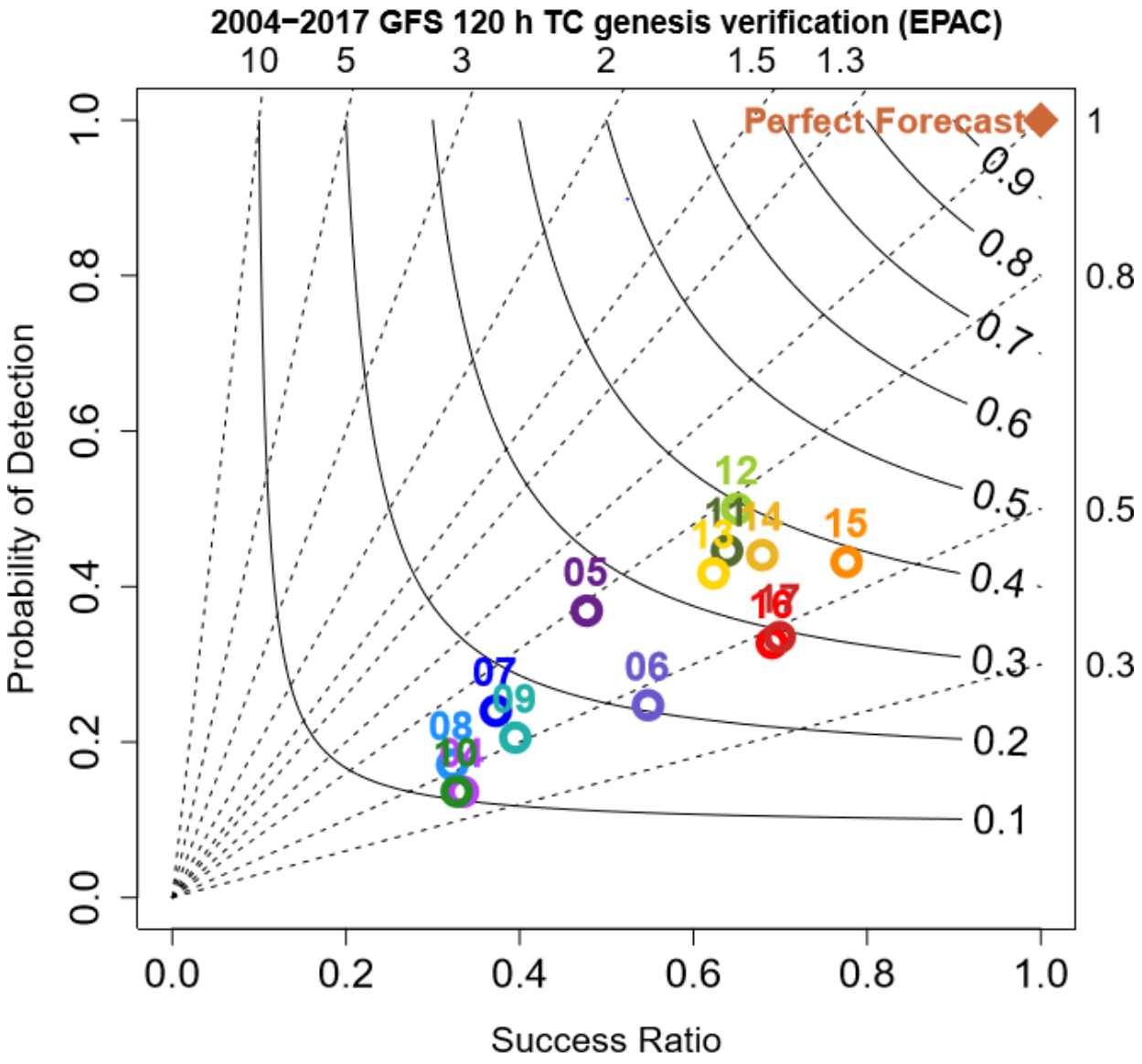


Figure 12b. As in Figure 12a, except for the East Pacific.

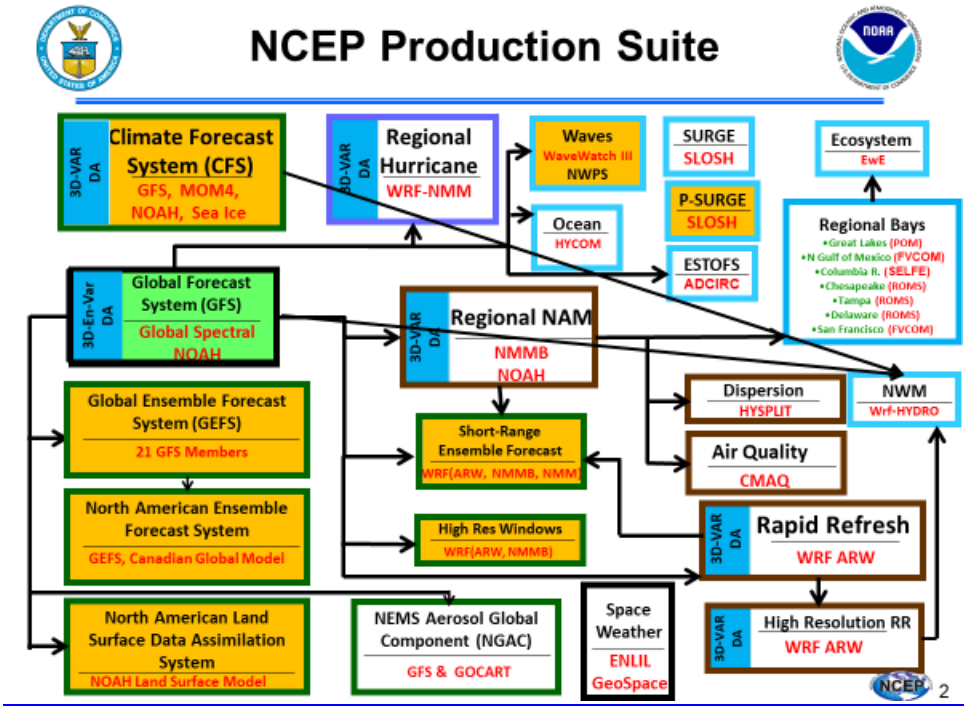


Figure 13. The NCEP Production Suite as of March 2018. Provided by Bill Lapenta.