Briefing to EMC CCB
September 30, 2020

GFS V16.0.0 -- Upgrades for Q2FY2021

Presented by:

Fanglin Yang
On Science, Infrastructure and Product Changes

Geoffrey Manikin
on Performance Evaluation

Based on work done by EMC MDAB, VPPP, and EIB branches, in collaboration with OAR/GFDL, OAR/PSL, OAR/GSL and NCAR, and various GFS downstream code managers and external collaborators

GFSv16 Acknowledgements

Project Manager: Vijay Tallapragada
Project Leads: Fanglin Yang, Russ Treadon
Model: Jongil Han, Jack Kain, Weizhong Zheng, George Gayno, Helin Wei, Ruiyu Sun, Moorthi Shrinivas, Anning Chen, Rongqian Yang, Xingren Wu, Xiaqiong Zhou, Valery Yudin, Youlong Xia, Jesse Meng, Henrique Alves, Roberto Padilla, Jessica Meixner, Ali Abdolali, Jian-Wen Bao, Henry Juang, Sajal Kar
Data Assimilation: Jeff Whitaker, Catherine Thomas, Cory Martin, Wanshu Wu, Phil Pegion, Haixia Liu, Kristen Bathmann, Andrew Collard, Xu Li, Yanqiu Zhu, Xiujuan Su, Shelley Melchior, Sudhir Nadiga, Steve Stegall
Post & VV: Hui-ya Chuang, Wen Meng, Boi Vuong, Mallory Row, Guang-Ping Lou, Yali Mao, Jiayi Peng, Deanna Spindler, Todd Spindler, Roberto Padilla, Geoff Manikin, Alicia Bentley, Logan Dawson, Shannon Shields, Chris MacIntosh, Philippe Papin
Infrastructure: Jun Wang, Kate Friedman, Mark Iredell, Hang Lei, Eric Rogers, George Vandenberghe, James Abeles, Gerhard Theurich, Edward Hartnett, Farida Adimi
Management: Ivanka Stajner, Vijay Tallapragada, Arun Chawla, Jason Levit, Avichal Mehra, Daryl Kleist, Fanglin Yang
NCO: Carissa Klemmer, Steven Earle, Anne Myckow, Dataflow team
External Collaborators: CPC, GFDL, GSL, PSL and NCAR/CGD
STI SOO Team: Robert Ballard, Warren Blier, Mike Fowle, Chris Karstens, Mark Klein, David Levin, Bill Martin, Emily Niebuhr, Jack Settelmaier, Steverino Silberberg, Ben Trabing, Brian Zachry

GDAS/GFS Version 16
Status as of September 28, 2020

G

Project Information & Highlights

Project Manager: Vijay Tallapragada

Leads: Fanglin Yang and Russ Treadon (EMC), Steven Earle (NCO)

Scope: Develop and incorporate new capabilities into the NCEP GFS with 13 km resolution and 127 levels, including advanced physics and DA system, including GLDAS in DA cycle, and coupling to a wave model (one-way). Additional capabilities from the NGGPS community were also incorporated (project plan & charter)

Expected benefits: higher model vertical resolution, extended model domain up to the mesopause, improved model physics, advanced data assimilation, improved model forecast skills.

Dependencies: gravity-wave drag parameterization; wave coupling, and DA upgrade; Satisfactory evaluation by stakeholders and downstream products

Schedule

G

Milestones & Deliverables	Date	Status
Freeze model code and data assimilation system	5/19/20	Complete
EMC/NCO EE2 kick off meeting	6/11/20	Complete
PNS due to HQ	6/18/20	Complete
Complete full retrospective/real time runs and evaluation	8/31/20	Complete
Complete Field evaluation	9/25/20	Complete
OD Brief	10/5/20	**planned**
Deliver final system code and SCN to NCO	10/09/20	**planned**
Start 30-day evaluation and IT testing	12/21/20	**planned**
Operational Implementation	2/3/21	**planned**

Issues/Risks

G

Risk: None

Resources

G

Staff: 3 Fed FTEs + 10 contractor FTEs; including Dev (FV3, physics, DA, post processing, V&V, and infrastructure)

Funding Source: STI/NGGPS

Compute: EMC Dev: (+100%)**; Parallels:** (+100%); **Ops**: 800 nodes HWM

Archive: Parallels: 2 PB HPSS for 1-year retros; **Ops**: 7 TB online and 1 TB HPSS per cycle

R Management Attention Required	Y Potential Management Attention Needed	G On Target

Topics

- Science changes
- System configuration and resource requirement
- Product changes
- Downstream model evaluation

- Performance evaluation
- Downstream user evaluation
- Benefits and concerns

Change History of GFS Configuration

Mon/Year	Lev	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
Aug 1987	18	T80 (150km)	Sigma Eulerian	First triangular truncation; diurnal cycle
Mar 1991	18	T126 (105km)	Sigma Eulerian	
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	T170 (80km)	Sigma Eulerian	the restoration
Jan 2000	42	T170 (80km)	Sigma Eulerian	first on IBM
Oct 2002	64	T254 (55km)	Sigma Eulerian	RRTM LW;
May 2005	64	T382 (35km)	Sigma Eulerian	2L OSU to 4L NOAH LSM; high-res to 180hr
May 2007	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 etc
May2016	64	T1534 (13km)	Hybrid Semi-Lag	4-D Hybrid En-Var DA
Jun2017	64	T1534 (13km)	Hybrid Semi-Lag	NEMS GSM, advanced physics
Jun 2019	64	FV3 (13km)	Finite-Volume	NGGPS FV3 dycore, GFDL MP
Feb 2021	127	FV3 (13km)	Finite-Volume	IAU, LETKF, TKE-EDMF, uGWD

18 years !

GFS.v16: Major Upgrades to Forecast Model

Model resolution:

Increased vertical layers from 64 to 127 & raised model top from 54 km to 80 km

Physics updates

PBL/turbulence: Replaced K-EDMF with sa-TKE-EDMF
Revised background diffusivity as a stability dependent function

GWD: Added a parameterization for subgrid scale nonstationary gravity-wave drag

Radiation: Updated calculation of solar radiation absorption by water clouds; Updated cloud overlap assumptions.

Microphysics: Updated GFDL microphysics scheme for computing ice cloud effective radius

Noah LSM: Revised ground heat flux calculation over snow covered surface; Introduced vegetation impact on surface energy budget over urban area

Coupling to Wave

One-way coupling of atmospheric model with Global Wave Model (WaveWatch III)

Coupling to GLDAS

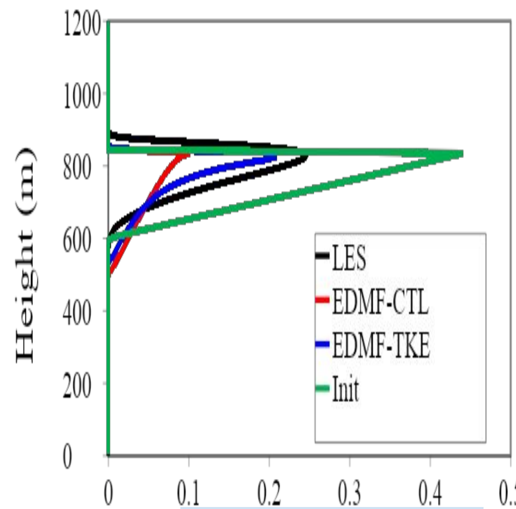
Spin up land states using CPC Gauge precipitation in the GDAS 00Z cycle

Advanced Features of TKE-EDMF
Vertical Turbulent Mixing Scheme over the K-EDMF

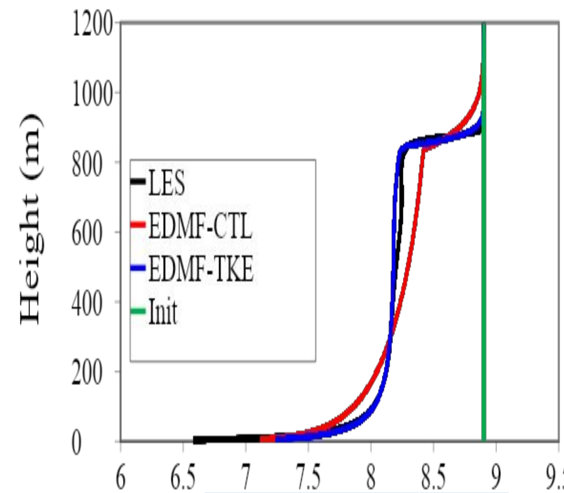
- **Higher-order accuracy in turbulence representation**, less diffusive than K-EDMF
- Advection of turbulence by the grid-mean flows
- Inclusion of **moist processes**
- Mass-flux representation for the **nonlocal momentum mixing**
- EDMF parameterization for the **stratocumulus-top-driven turbulence mixing**
- **Scale awareness**
- **Interaction of TKE with cumulus convection**

SCM simulation vs LES

For the marine stratocumulus-topped boundary layer, the TKE-EDMF better simulates the liquid water and wind speed profiles than the K-EDMF (*EDMF-CTL*) compared to the LES. The simulated liquid water profile from TKE-EDMF is correctly less diffusive. Also, the TKE-EDMF displays a well-mixed feature of momentum similar to the LES, whereas the K-EDMF fails to simulate the well-mixed momentum due to the lack of nonlocal momentum mixing.



cloud water

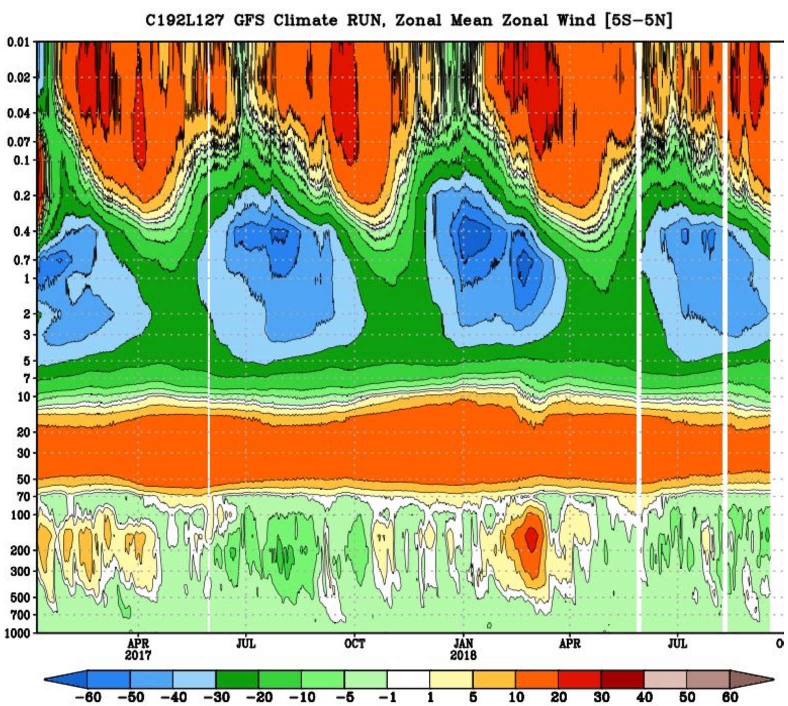


wind speed

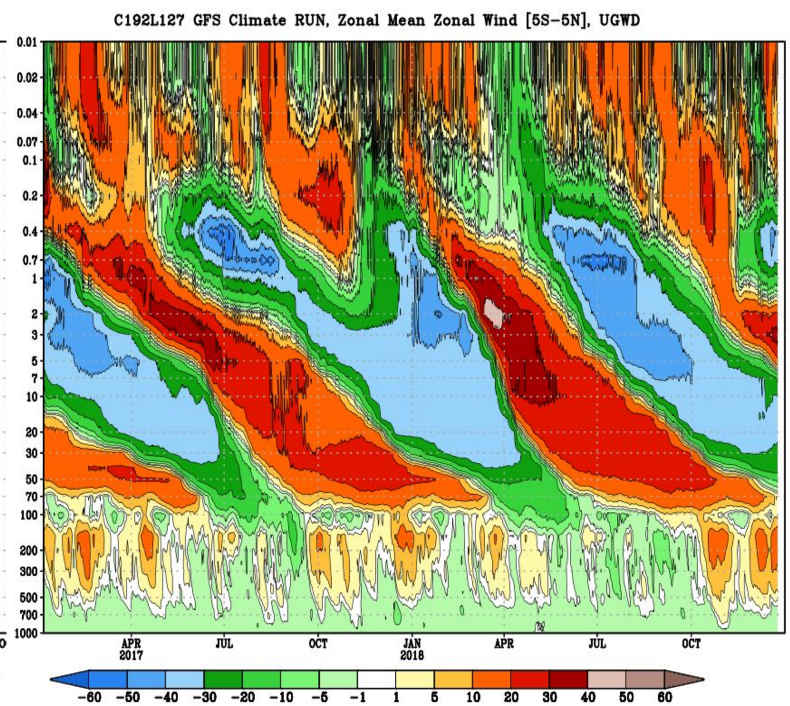
Non-Stationary GWD: Impact on QBO/SAO

In collaboration with CIRES, UCB

CONTROL



W/ non-stationary GWD



Prescribed Forcing (MERRA2 AGCM)

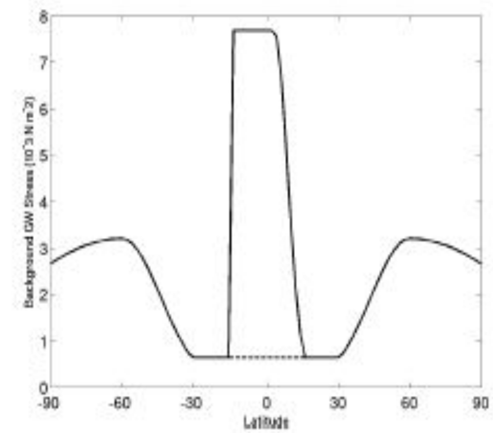


Figure 3. Background non-orographic drag from the MERRA (dashed line) and MERRA2 AGCM (solid line) simulations. The dashed line underlies the solid line outside of the tropics.

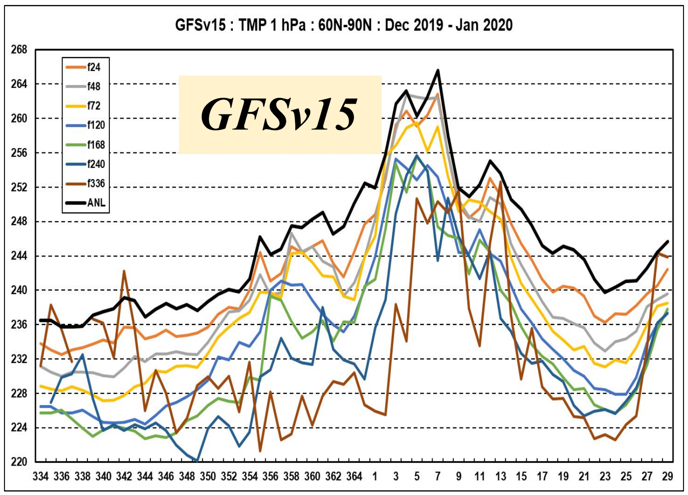
- **Current operational model cannot simulate the QBO**
- **A QBO-like feature is captured in GFS.v16 "climate" run with the non-stationary GWD physics included; However, the periodicity is too short, appears to be a downward propagating SAO.**

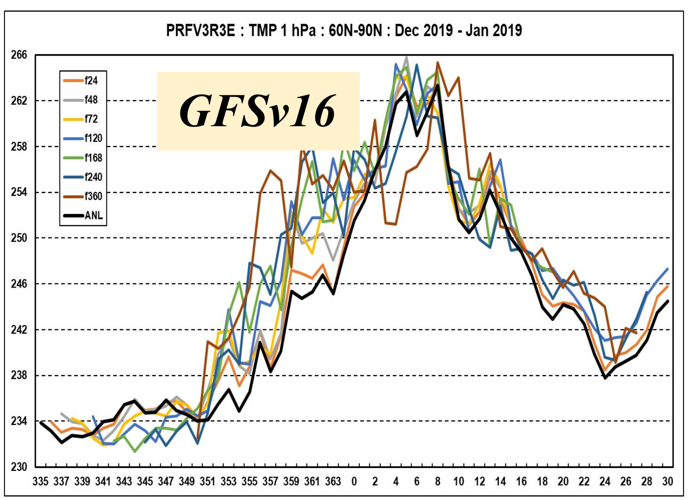
Forecast Improvements in the Stratosphere

Courtesy of Craig Long, NCEP/CPC

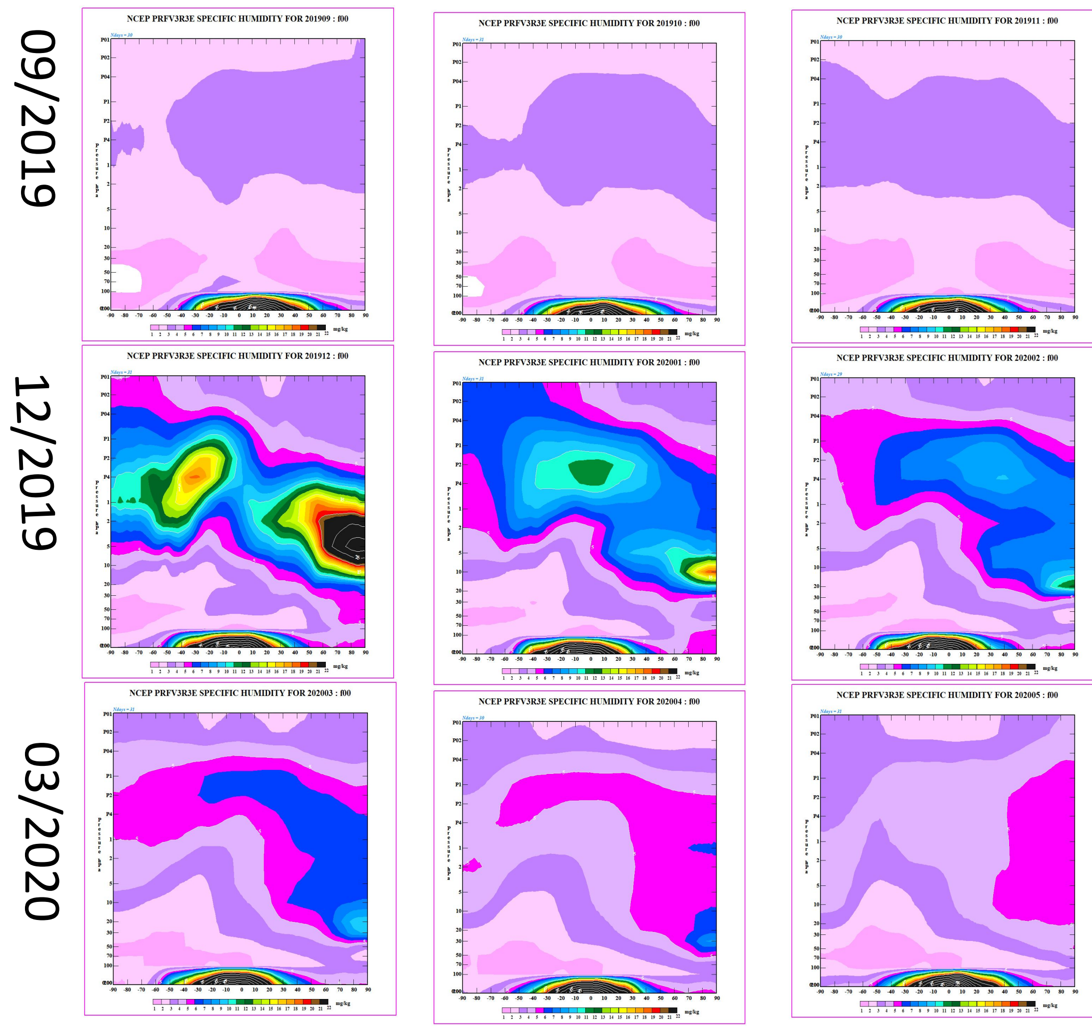
Improved 1-hPa Temperatures :

60N-90N Dec 2019 – Jan 2020





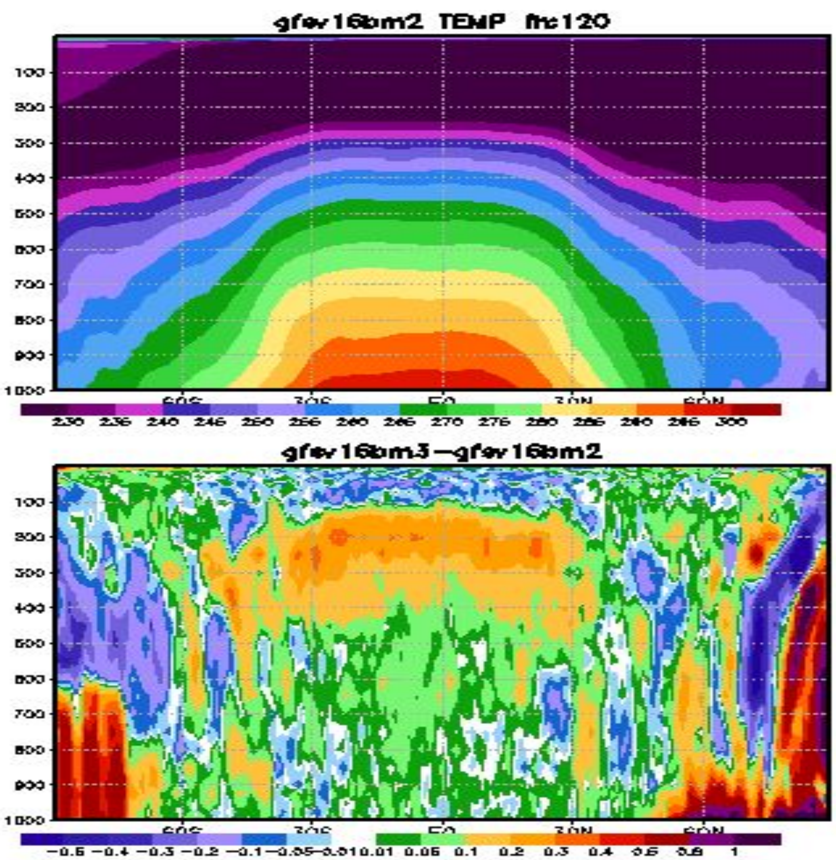
Captured water vapor seasonal cycle in the stratosphere, compares well well with UARS HALOE observations (no shown)



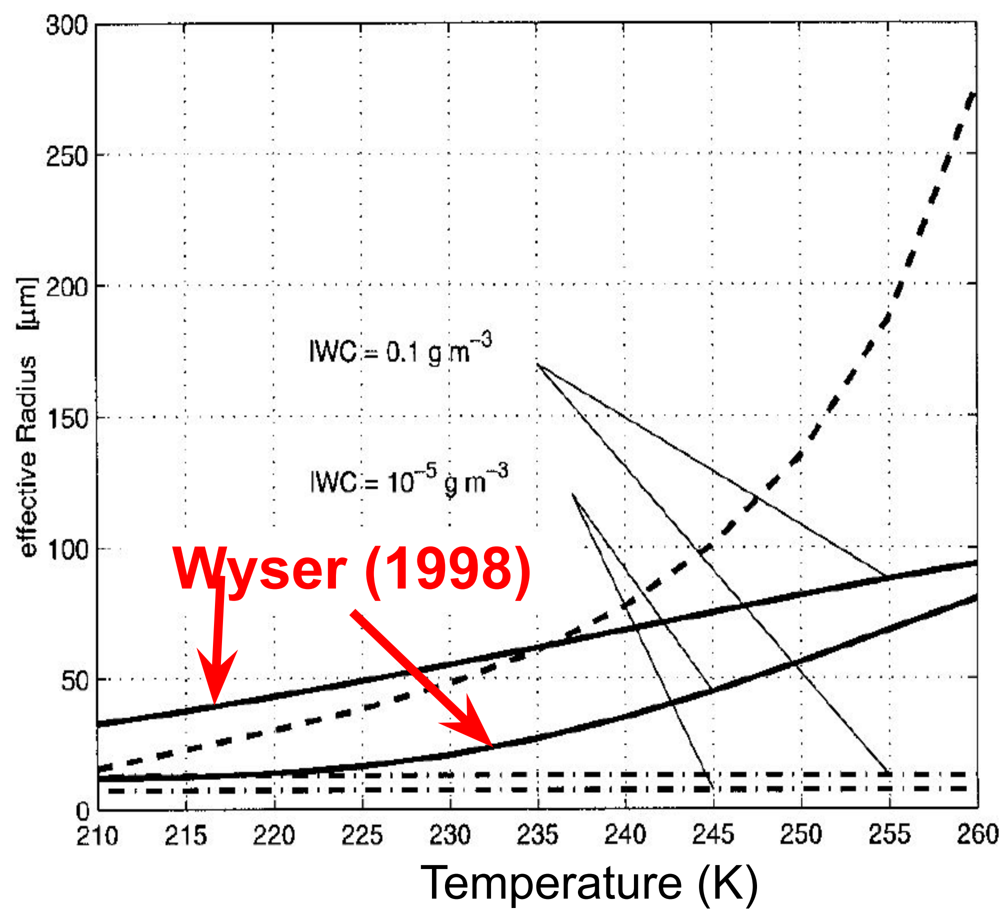
Improved Interaction between Ice Clouds and Radiation

In collaboration with GFDL

Use Wyser (1998) formula to calculate $r_{eff\text{-}ice}$ as a function of q_i and T for q_i > qmin instead of using a constant $r_{eff\text{-}ice}$



reduced tropospheric cold bias



GFS V16: One-Way Coupling to WAVE

Multi_1

- Arctic Polar Stereographic
  - 18 km resolution
  - 50°N to 90°N
- Global grid: 30 arc min
- Regional grids: 10 arc min
  - ak_10m
  - wc_10m
  - at_10m
  - ep_10m
- Coastal grids: 4 arc min
  - ak_4m
  - wc_4m
  - at_4m

GFSv16 wave

- Arctic Polar Stereographic
  - 9 km resolution
  - 50°N to 90°N
- Global core
  - 16 km (10 arcmin)
  - 15°S to 52.5°N
- Southern Ocean
  - 25 km (15 arcmin)
  - 10.5°S to 79.5°S

- New RTOFS ocean surface current forcing up to 192h,
- Forecasts will be extended from 180 hr to 384 hr.

Major Upgrades to Data Assimilation

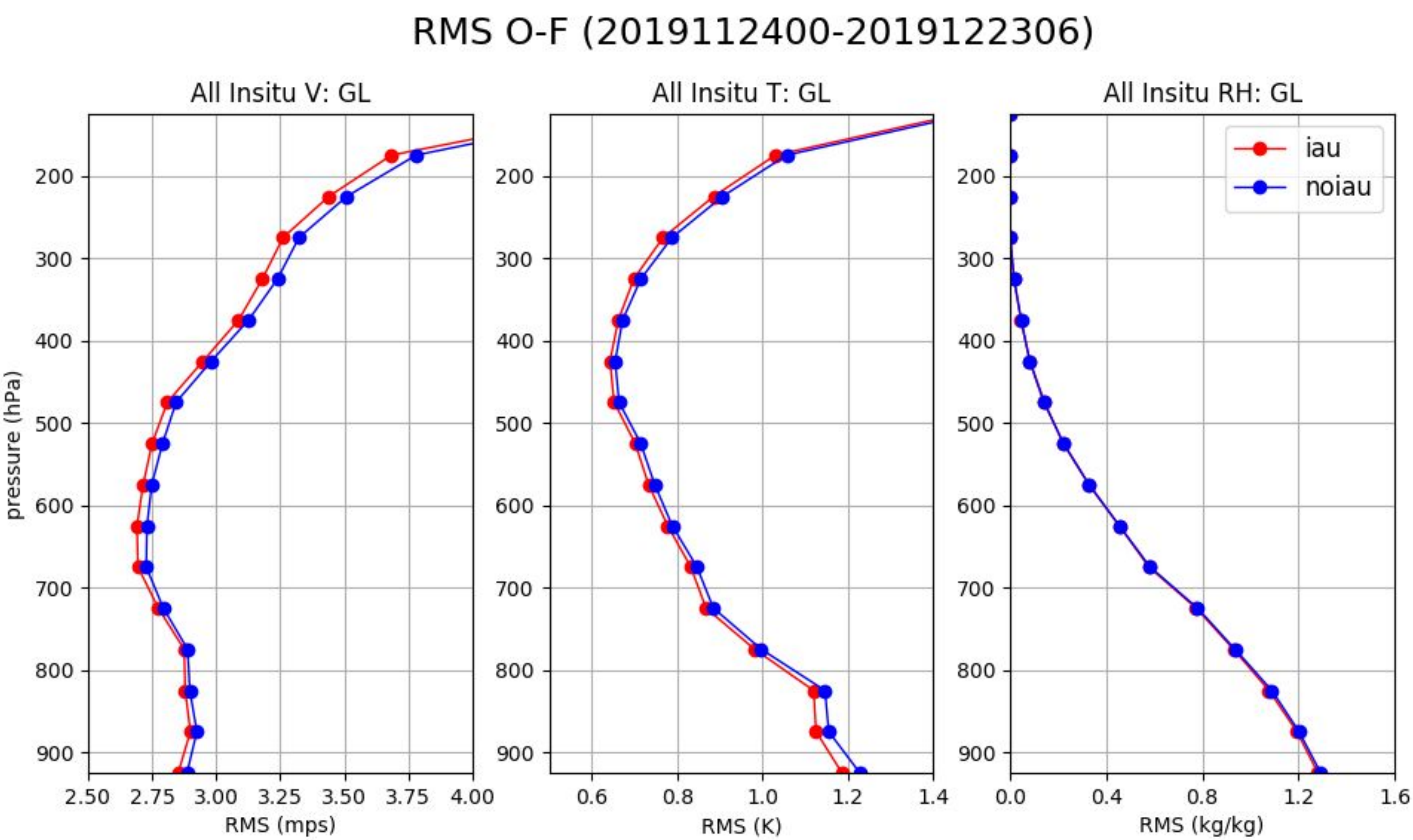
- **Local Ensemble Kalman Filter (LETKF)**
  with model space localization and linearized observation operator
  to replace the Ensemble Square Root Filter (EnSRF)

- **4-Dimensional Incremental Analysis Update (4D-IAU)**

- T**urn on SKEB** in EnKF forecasts
- Update variational QC
- Apply Hilbert curve to aircraft data
- **Correlated observation error** for CrIS over sea surfaces and IASI over sea and land
- Update temperature aircraft bias correction with safeguard
- Assimilate AMSU-A channel 14 and ATMS channel 15 w/o bias correction
- **Assimilate CSR data from ABI_G16, AHI_Himawari8, and SEVIRI_M08**
- **Assimilate AVHRR** from NOAA-19 and Metop-B for NSST
- **Assimilate additional GPSRO** (add Metop-C GRAS, More Cosmic-2)
- **Assimilate** high-density flight-level wind, temperature, and moisture observations **(HDOB) in tropical storm environment**
- Reduce the distance threshold for inner core dropsonde data to 55km (from 111km or 3*RMW) and add a wind threshold of 32 m/s to a**llow more dropsonde data being assimilated**
- **Use CRTM v2.3.0**

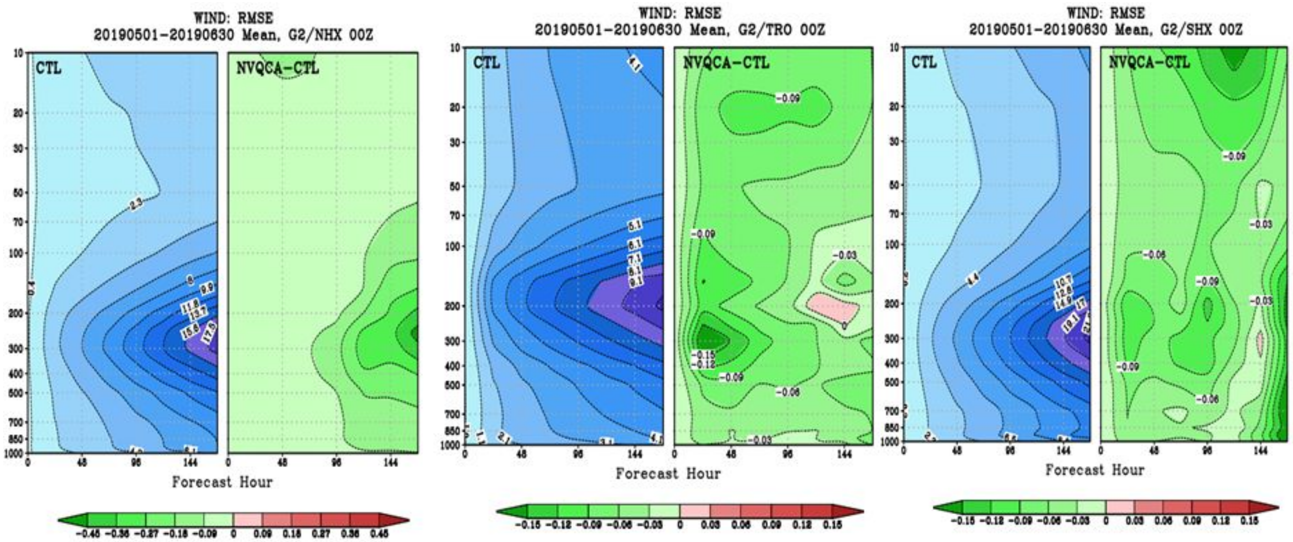
Incremental Analysis Update (IAU)

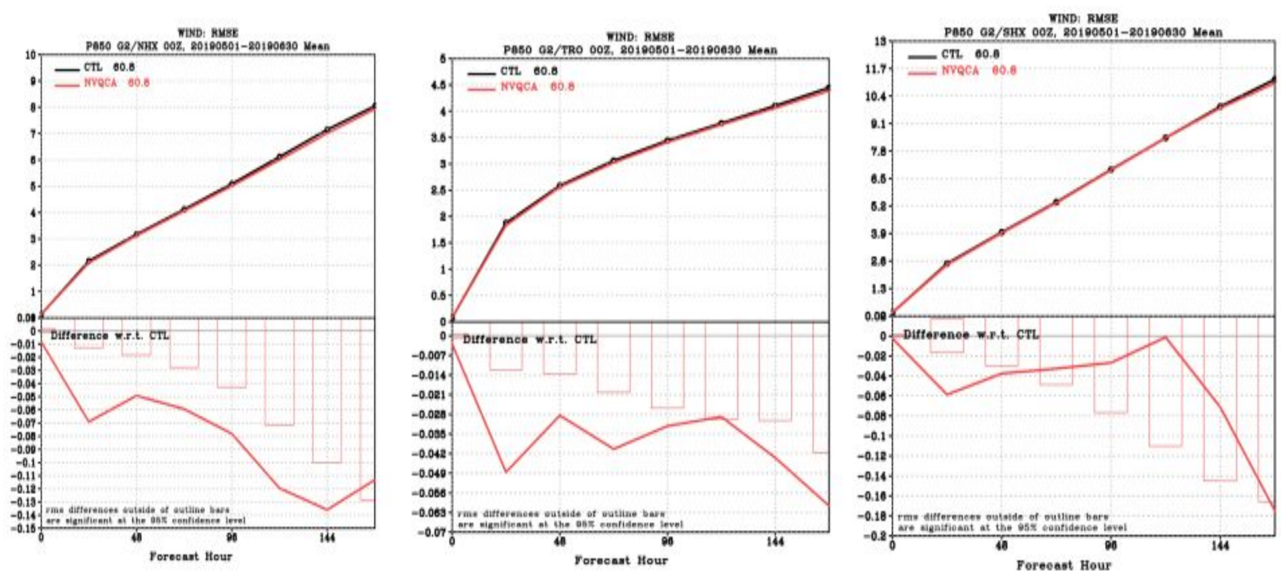
In collaboration with OAR/PSL



New Variational Quality Control

- A **new variational quality control** is applied to **conventional observations**.
- Previous variational quality control could not be applied in the first iterations of minimization due to the possibility of multiple minima in the cost function.
- New probability density function formulation greatly reduces the possibility of multiple minima.
- **Greatest impact in wind RMSE** and in the northern hemisphere.

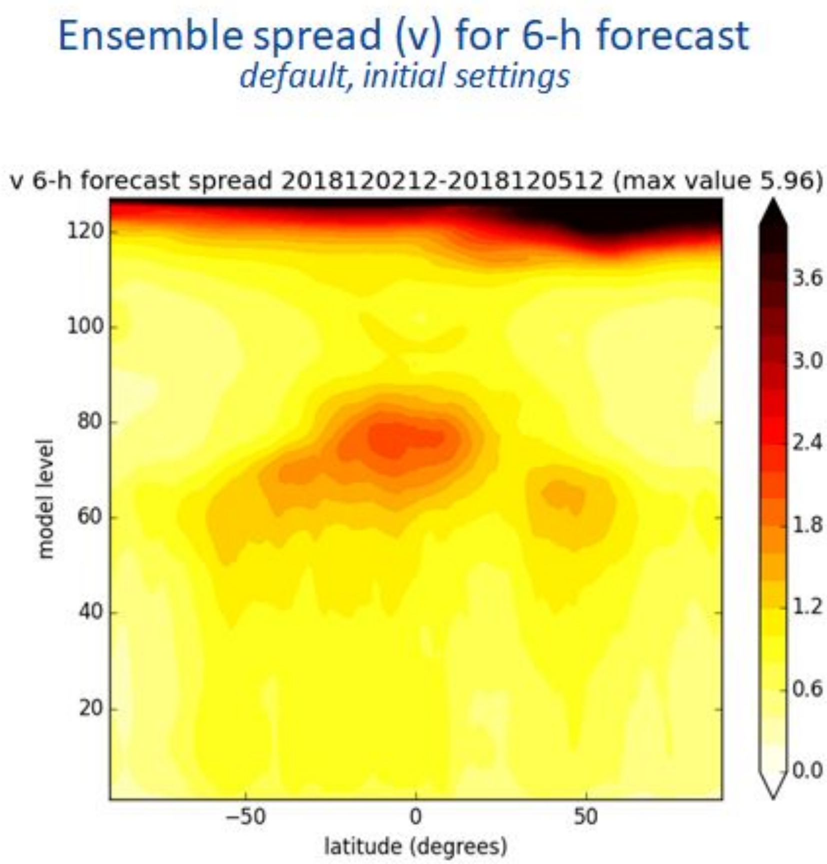


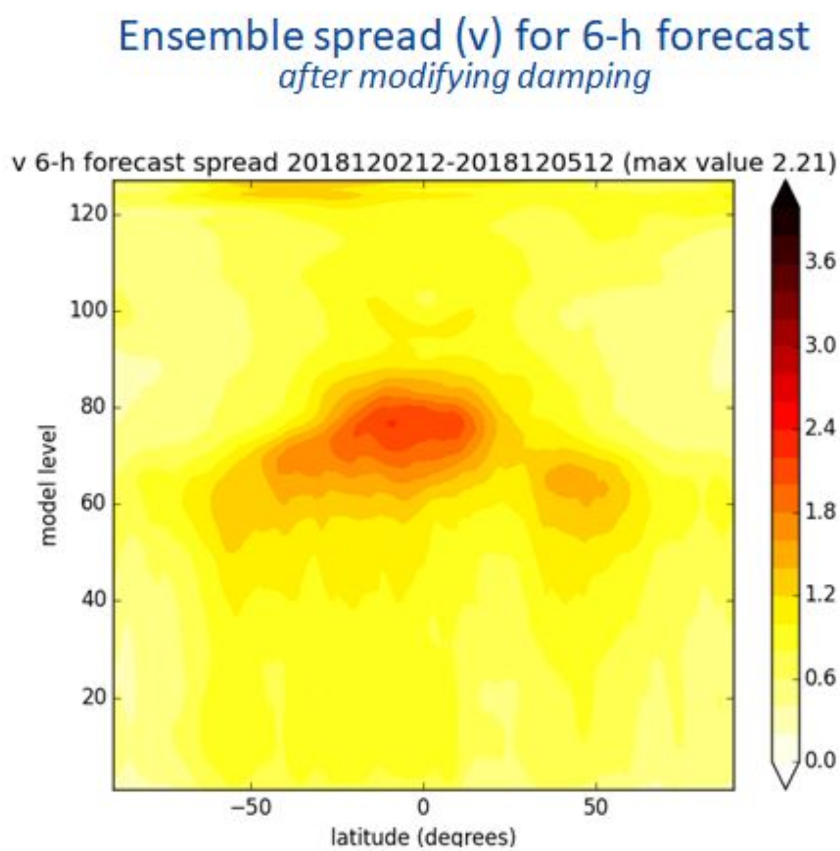


Reduce Ensemble Spread at Model Top

In collaboration with OAR/PSL

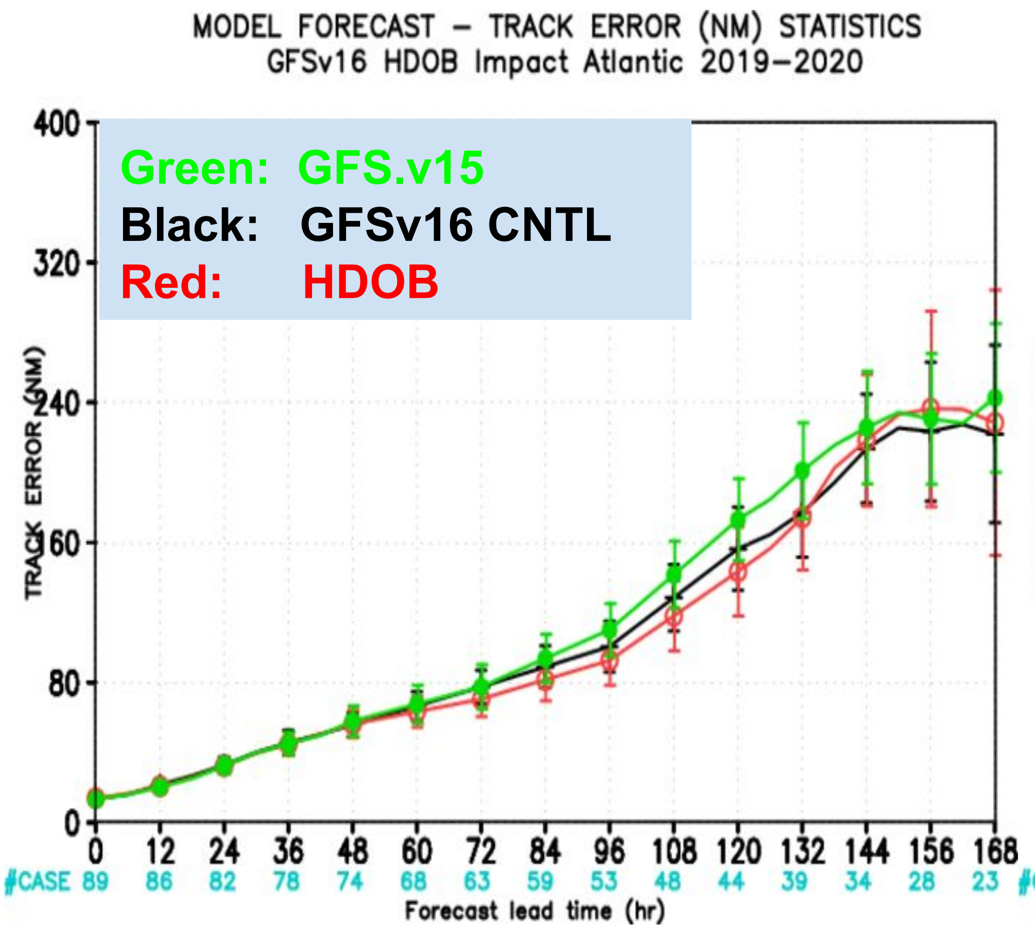
Initial attempts at cycling (with ensemble) were poor due to large spread in upper layers. New damping was added to reduce the spread



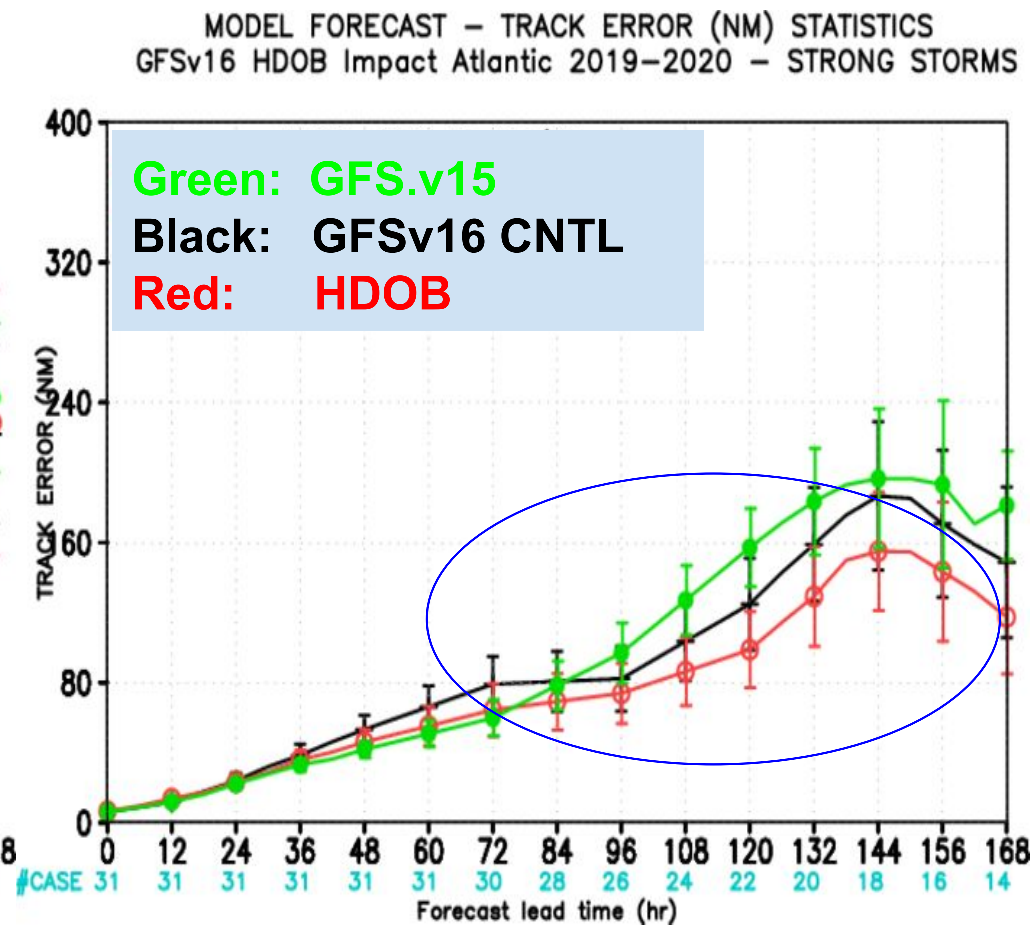


Impact of Aircraft High Density Observations and Dropsondes on Tropical Cyclone Forecasts

Track Error -- Atlantic Basin
2019 ~ 2020 Hurricane Seasons



All Storms



Strong Storms vmax>50kts

I/O Change – Use NetCDF to replace nemsio

GFS.v15 (C768L64) history files in nemsio format:

atmf 16.8 GB
sfc 2.8 GB

GFS.v16 (C768L127), in nemsio format

atmf 33.6 GB
sfc 5.6 GB

A decision was made to write out GFS.v16 forecast history files (atmf and sfcf) in netCDF format with compression. Parallel I/O was developed with updated netCDF and HDF libraries.

compression ratio:

Atmf 3d	5x	(33.6 GB to 6.7 GB),	lossy compression
sfc 2d	2.5x	(2.8 GB to 1.1 GB),	lossless compression

Pre-Processing Changes

obsproc_global and **obsproc_prep** was updated to process new satellite observations, high density aircraft observations, and to work with model history files in netCDF format.

exglobal_dump.sh.ecf modified to generate BUFR dump files for the following data types:

- GOES-16, -17 Clear Sky Radiance data (gsrcsr)
- GOES-16 All Sky Radiance data (gsrasr)
- OMPS Limb Profiler (ompslp)
- Himawari-8 Clear Sky Radiance data (ahicsr)
- VIIRS SST (Clear Sky w/o Land Radiance data) from NPP & NOAA-20 (sstvcw)
- VIIRS SST (Probably Clear Sky w/o Land Radiance data) from NPP & NOAA-20 (sstvpw)
- LEO-GEO Satellite AMVs from UWisc (leogeo)
- High Density obs from reconnaissance aircraft (hdob)

exglobal_dump.sh.ecf modified to remove legacy/obsolete bufr dump file processing:

- GOES-15 data
- Legacy VIIRS AMV data
- Obsolete EUMETSAT CrIS data (escris, escrsf)

`JGLOBAL_PREP` **and** **`exglobal_makeprepbufr.sh`**

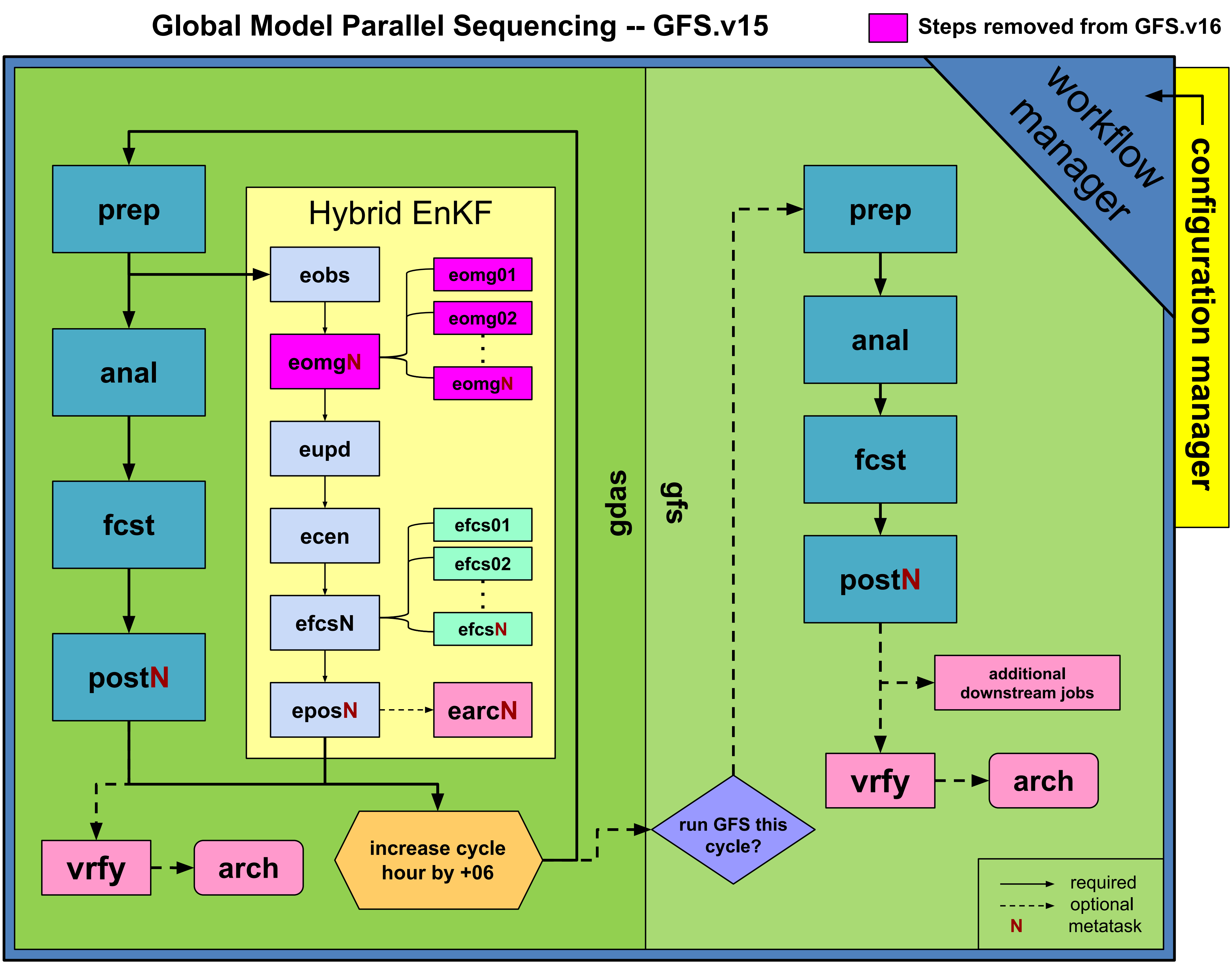
- Updated to handle netcdf history filename patterns

Inline Post and Offline Post

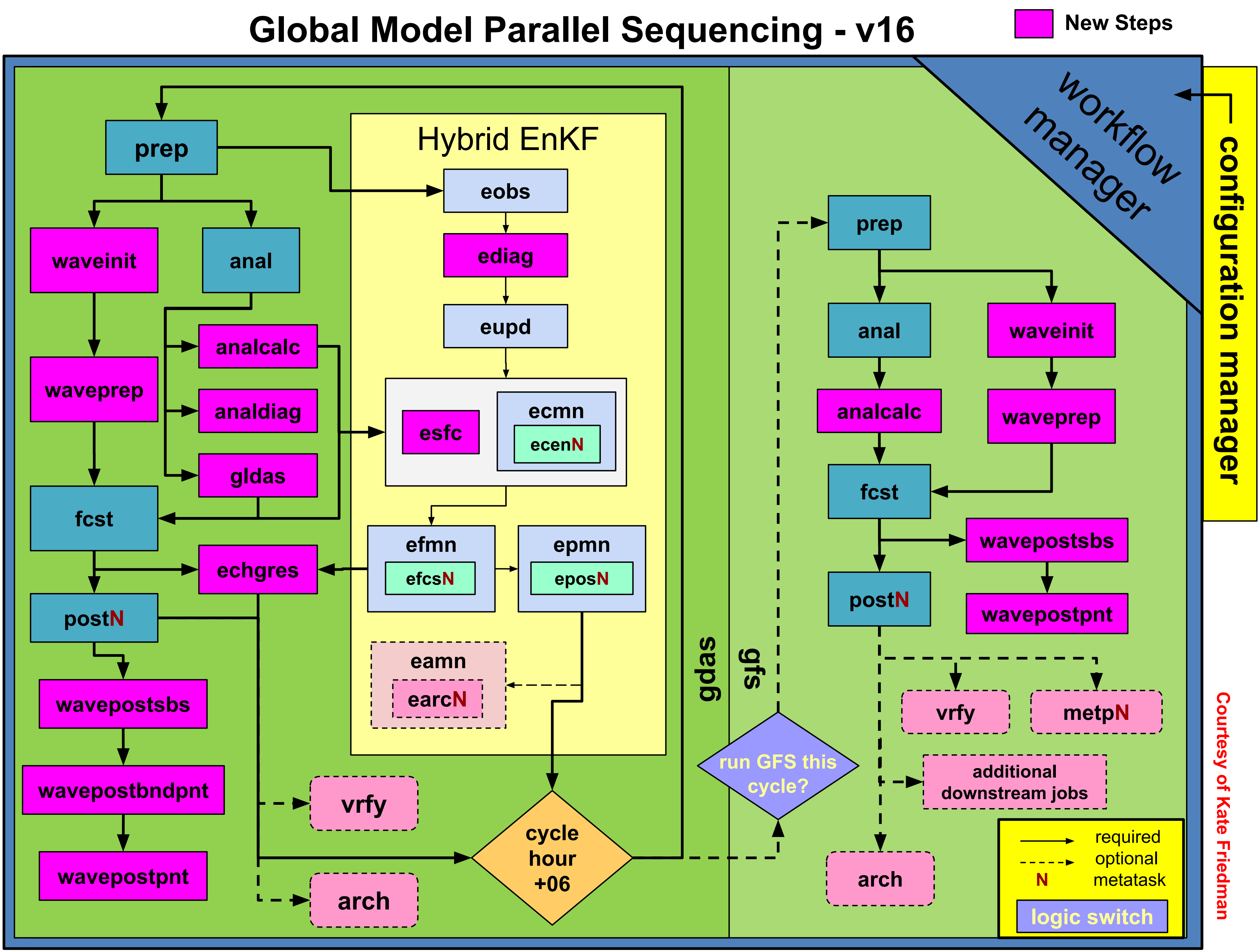
- **Inline post** was introduced to GFS.v16
  - Inline post makes use of forecast data saved in memory for post processing, *reduces I/O activity, and speeds up the entire forecast system*.
  - A **Post library** was created using the offline post Fortran programs. It can be called by the Write Grid Component within the forecast model.
  - Since lossy compression is applied for writing out forecast history files, *inline post generates more accurate products* than the standalone offline post.

- **GFS.v15**
  - ALL master, flux, simulated satellite radiance, and GTG files are made by the offline post.

- **GFS.v16**
  - Master and flux files are produced by the inline post.
  - Simulated satellite radiance and WAFS files are still made by the offline post.


Global Model Parallel Sequencing -- GFS.v15
Steps removed from GFS.v16
workflow manager
configuration manager
prep
anal
fcst
postN
vrfy
arch
Hybrid EnKF
eobs
eomgN
eomg01
eomg02
eomgN
eupd
ecen
efcsN
efcs01
efcs02
efcsN
eposN
earcN
increase cycle hour by +06
gdas
gfs
run GFS this cycle?
prep
anal
fcst
postN
additional downstream jobs
vrfy
arch
required
optional
N
metatask

Global Model Parallel Sequencing - v16



Product Changes -- Atmosphere

- Unification: All isobaric state fields in pgrb2 files will have the same 41 levels at all forecast hours and analysis time
- New products:
  - **Add more pressure levels** (at 0.01, 0.02, 0.04, 0.07, 0.1, 0.2, 0.7 hPas) in the upper stratosphere and the mesosphere in pgrb2 files
  - Other **new products** include *cloud ceiling, total column and low/mid/high cloud fractions, and radar reflectivity at 1 km/4 km and 1st/2nd model level above ground*.
  - FSU storm genesis verification stats
- Replaced products:
  - *Replace filtered Shuell SLP with unfiltered one using same ID PRMSL*
  - *Replace legacy synthetic nadir GOES 12/13 with synthetic nadir ABI GOES-R products*
- Products moved to different files:
  - Isobaric SPFH moved from pgrb2b to pgrb2 files
  - GTG and Icing severity moved to new file gfs.tHHz.wafs_0p25.fFF.grib2
- Changes in Grib2 IDs:
  - low/middle/high cloud from TCDC to LCDC, MCDC, HCDC
  - Icing Severity parameter from 234 to 37, mnemonics from ICSEV to ICESEV

Product Changes -- Atmosphere (Cont'd)

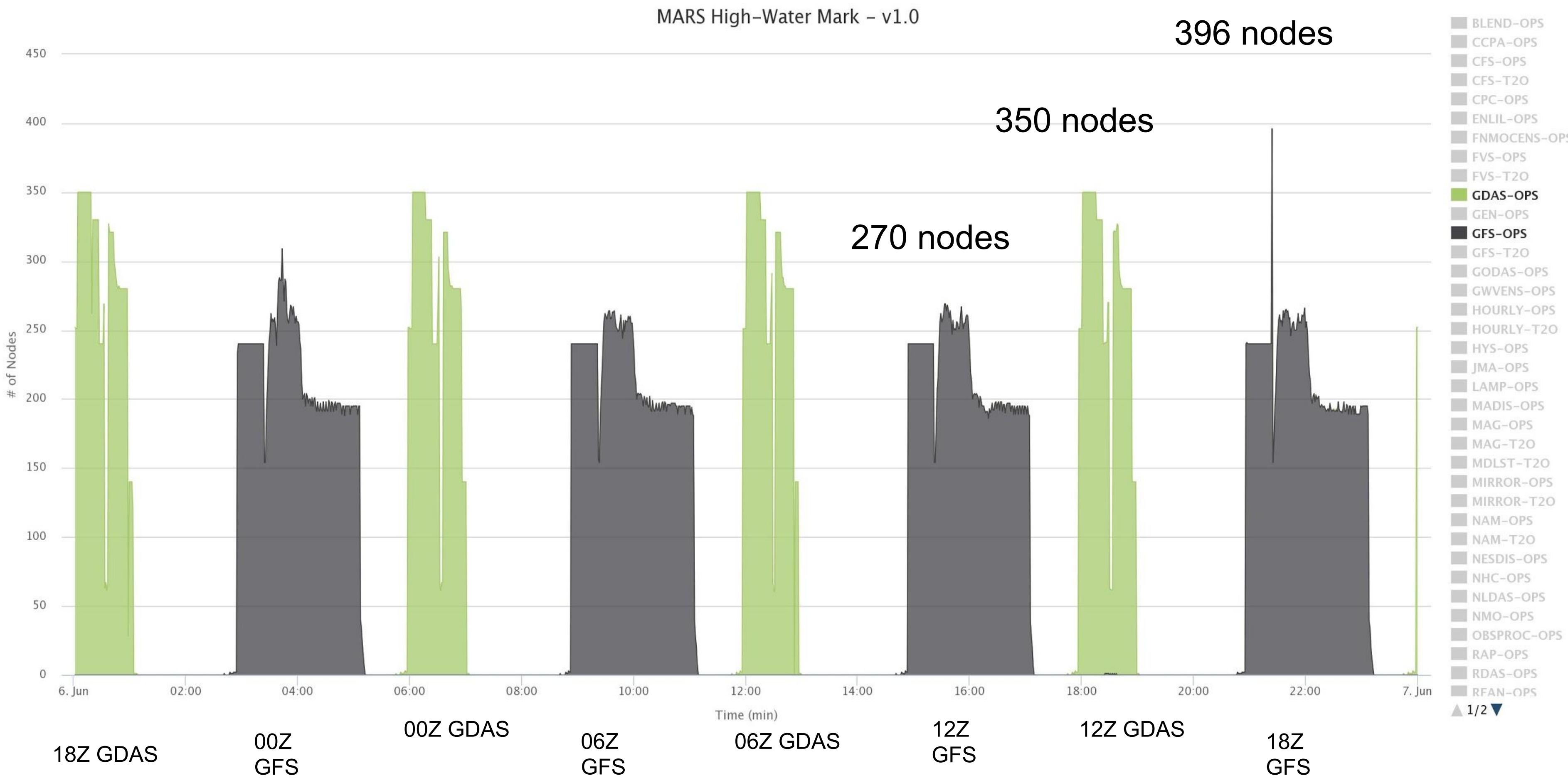
- **Delay in product delivery:** synthetic GOES products

- **Removed products:**
  - Three legacy bulletins (navy bull, wintemv bull, gdas bull)
  - 5-wave height (5WAVH) in all GFS pgb files, AWIPS 20km grids (CONUS, Alaska, Puerto Rico, Pacific region) and AWIPS LAT/LON 1.0 degree grid
  - Lifted Index in GFS Flux files
  - SPFH at 16 levels from pgrb2b
  - A PNS on GFS V16 product removal has been sent out in April. No objection within 30 day period.

- **Update to GFS Bufr sounding output**
  - Increase vertical levels from 64 to 127
  - Remove terrain adjustments of temperature and SPFH profiles from station elevation
  - Sea-surface pressure is reduced from model surface height
  - Changed surface evaporation value and unit from watts/m^2 (surface latent heat net flux) to kg/m^2 (evaporation)

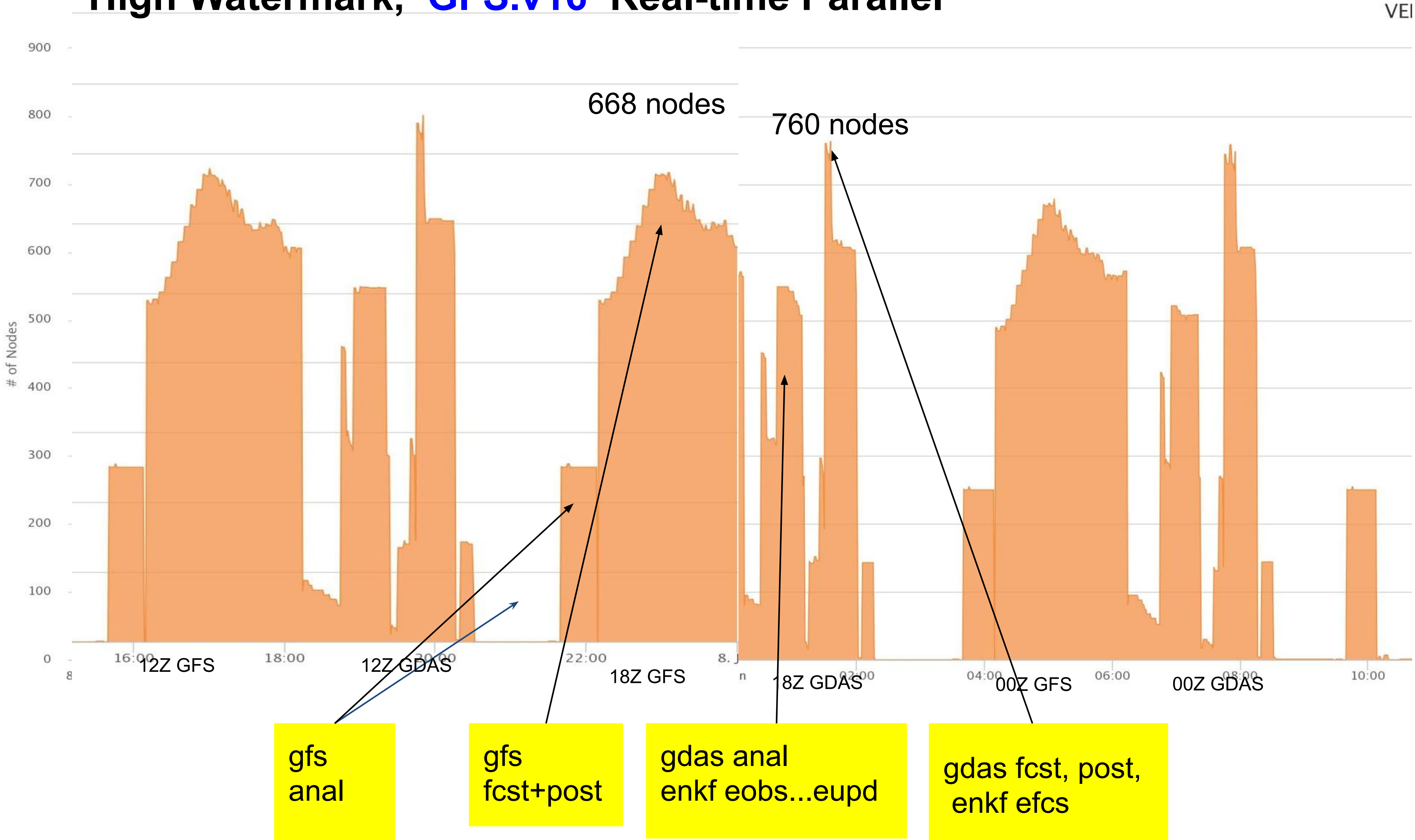
Changes in Computing Cost

High Watermark, GFS.v15 current operation



Changes in Computing Cost

High Watermark, GFS.v16 Real-time Parallel



Computational Cost
Timing and Node Usage

	GFS v15 time (min)	GFS v15 nodes	GFS v16 time (min)	GFS v16 nodes
gfs_analysis	28.0 - 28.7	240	28.1 - 29.4	250
gdas_analysis_high	32.2 - 33.0	240	38.2 - 39.3	250
gfs_forecast_high	100.8 - 103.4 (6.38 min/day)	148	122.8 - 124.2 (7.72 m/day)	484
wave_fcst	53.8 - 54	18	122.8 - 124.2	60
gdas_forecast_high	11.5 - 11.7	28	21.10- 21.5	119
enkf_update	6.5 - 6.8	90	25.6 - 26.7	240
enkf_fcst_XX	19.7 - 19.8	14 x 20 = 280	28.5 - 31.5	15 x 40 = 600

Delays in v16 with respect to v15

- gfs prep + anal + fcst
  - v15: 134.4 minutes
  - v16: 158.2 minutes (23.8 minutes longer, but still within 8 min/day)
- gdas 06, 12, 18Z
  - v15: prep + anal + fcst = 47.5
  - v16: prep + anal + fcst = 67.0 (19.5 minutes longer)
- gdas 00Z
  - v15: prep + anal + fcst = 48.3
  - v16: prep +anal + gldas + fcst = 72.3 (24 minutes longer)
- enkf
  - v15: eobs + eomg + eupd + ecen + efcs + epos = 53.2
  - v16: eobs + ediag + eupd + analdiag + ecen + efcs = 80.7 (27.5 minutes longer

Data Volume in COMROT (in TB, 4 cycles)

Daily totals (Tb)	**GFS v15**	**GFS v16**
gdas.$PDY/$cyc	1.25	0.92
enkfgdas.$PDY/$cyc	23.11	15.65
gfs.$PDY/$cyc	18.63	12.28
gdaswave.$PDY/$cyc	-na-	0.02
gfswave.$PDY/$cyc	0.02 (Multi-1)	0.04
rtofs.$PDY	-na-	0.02
TOTAL Daily Tb	**42.99**	**29.11**

HPSS Archive for operation (GB/cycle)
see details here

Daily totals (Tb)	**GFS v15**	**GFS v16**	**notes**
gdas	150	106	
enkfgdas & restart	840	1736	Increase due to IAU
enkfdgas history	1320	512	Reduction due to netCDF
gfs	813	512	Reduction due to netCDF
Multi_1 WAVE	16	60	Resolution, +GDAS, 7 → 16 day
TOTAL	**3139**	**2937**	

Changes in GFS Data Volume Disseminated to NOMADS (per cycle)

gdas	+/- GB
Model sfc output	- 4
pgrb2 at 0.25:	+ 2
gfs	
Model sfc output	- 18
Model atm output	- 251
pgrb2 at 0.25:	+ 42
pgrb2 at 0.50:	+ 9
pgrb2 at 1.0:	+ 2
Flux	+ 16
pgrbfull at 0.5:	+ 9
enkfgdas	
sfc output	- 96

- size of fcst files decreased by 369 GB
- size of pgrb products increased by 80 GB
- In total, reduced by 289 GB.

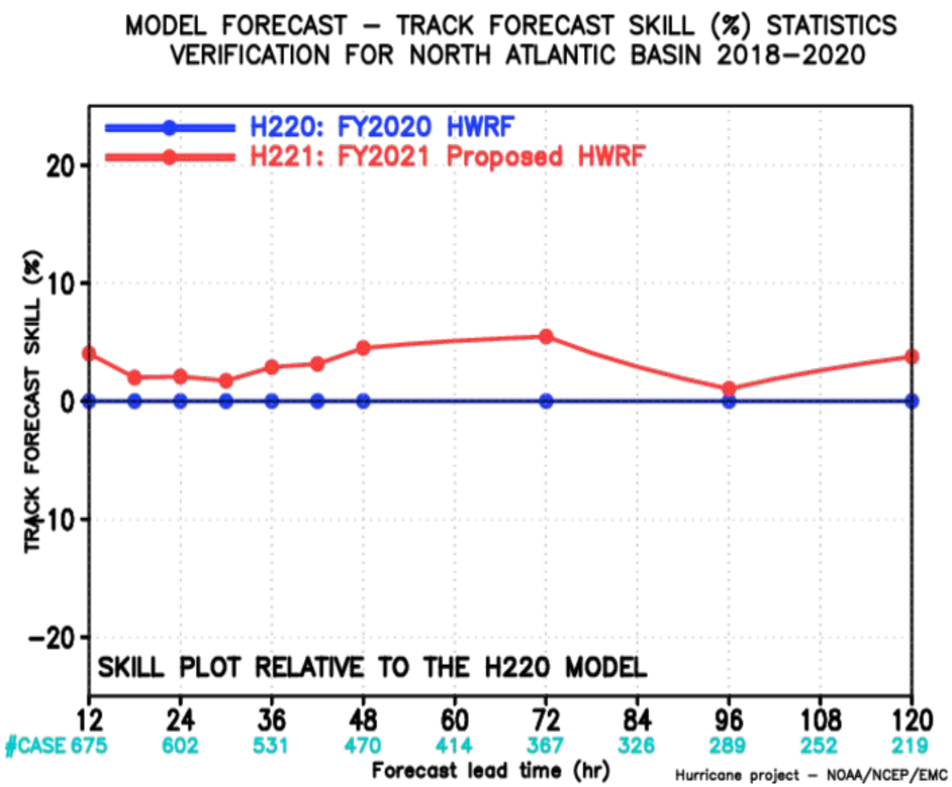
Increase of pgrb2 file size was due to increases in pressure levels, precision, and number of variables requested by users (details)

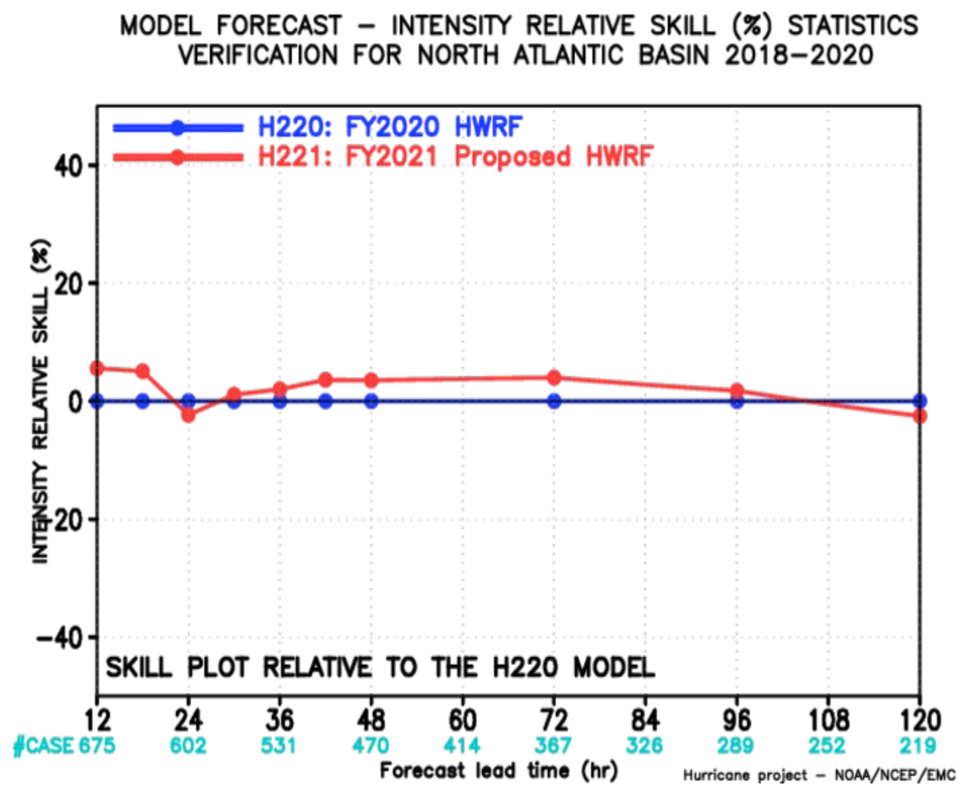
GFS.v16 Retrospective and Real-Time Parallels (1.5 years)

	Machine & Throughput	Period to be covered (total days)	Current Status (8/8/2020)	Wave starting Cycle	CAPE/CIN fix starting cycle	Projected completion Date	Notes
v16retro0e	Mars Dell 3.5 7 cycles/day	05/10/19~05/31/19 (26)	05/31/19	No WAVE	**rerun fcst completed**	July 4	For MEG evaluation of significant weather events.
v16retro1e	Mars Dell 3.5 7 cycles/day	06/1/19~08/31/19 (92)	08/31/19	2019060712	2019081512	July 23	MDL and NCAR need data for JJA 2019 by mid-July; HWRF test will start in June.
v16retro2e	Mars Dell 3.0 4 cycles/day	09/1/19 ~11/30/19 (91)	11/30/19	2019090918	2019102712	August 8	
v16retro3e	HERA 7 cycles/day	12/01/19 ~ ~~03/31/20~~ **05/19/20** (169)	05/19/20	2020013106	2020040112	August 1	MDL and NCAR need data for DJF 2019/20 by mid-July
v16retro5e	Venus Dell 3.5 4 cycles/day	08/31/18~10/12/18 (43)	10/10/18	No Wave	2018091012	August 10	Forecast length is 10 days for all cycles.
v16rt2	Mars Dell 3.0	05/19/20 ~		2020051900	2020071300		

Evaluation of GFS.v16 Downstream Models

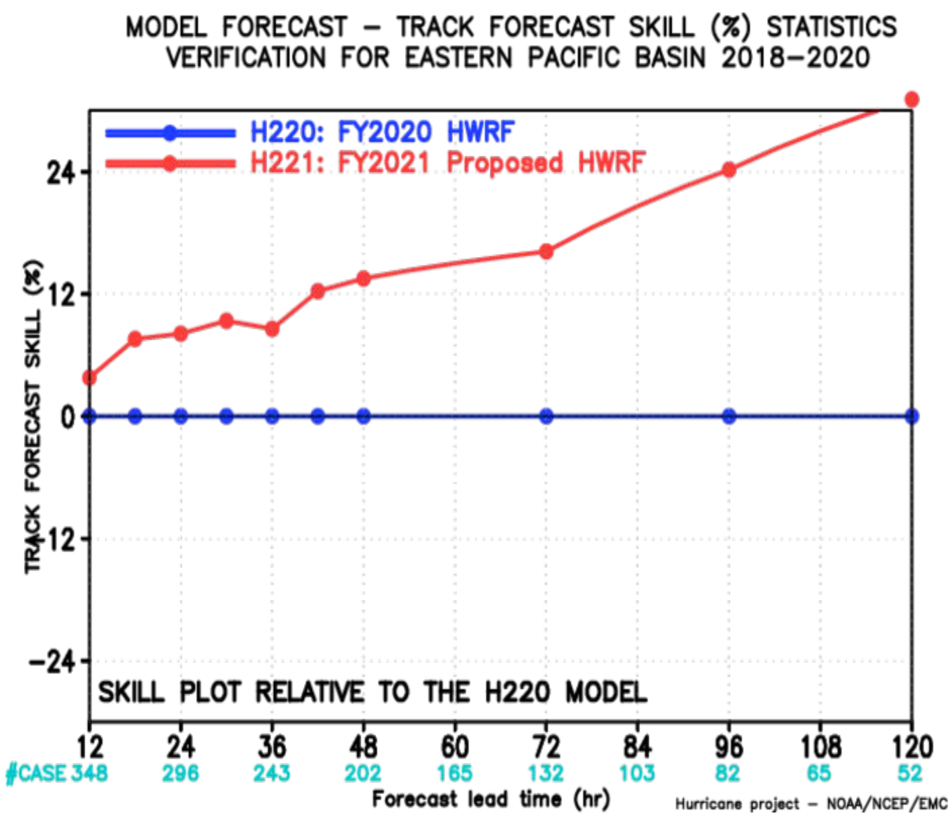
H221: HWRF (with GFS v16) vs H220 Operational HWRF for NATL Basin (Late model)

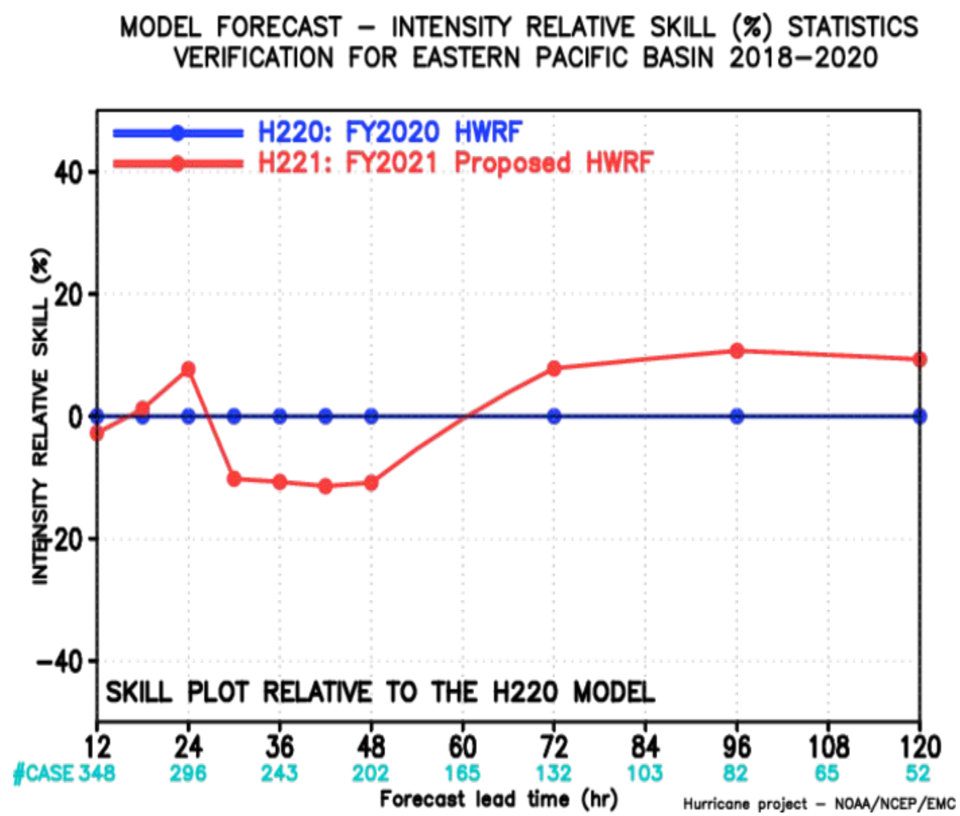




H221 shows improved track skill for all lead times with skill > 5% at Day 3. Intensity skill is also mostly positive except at days 1 and 5 where it is marginally below H220.

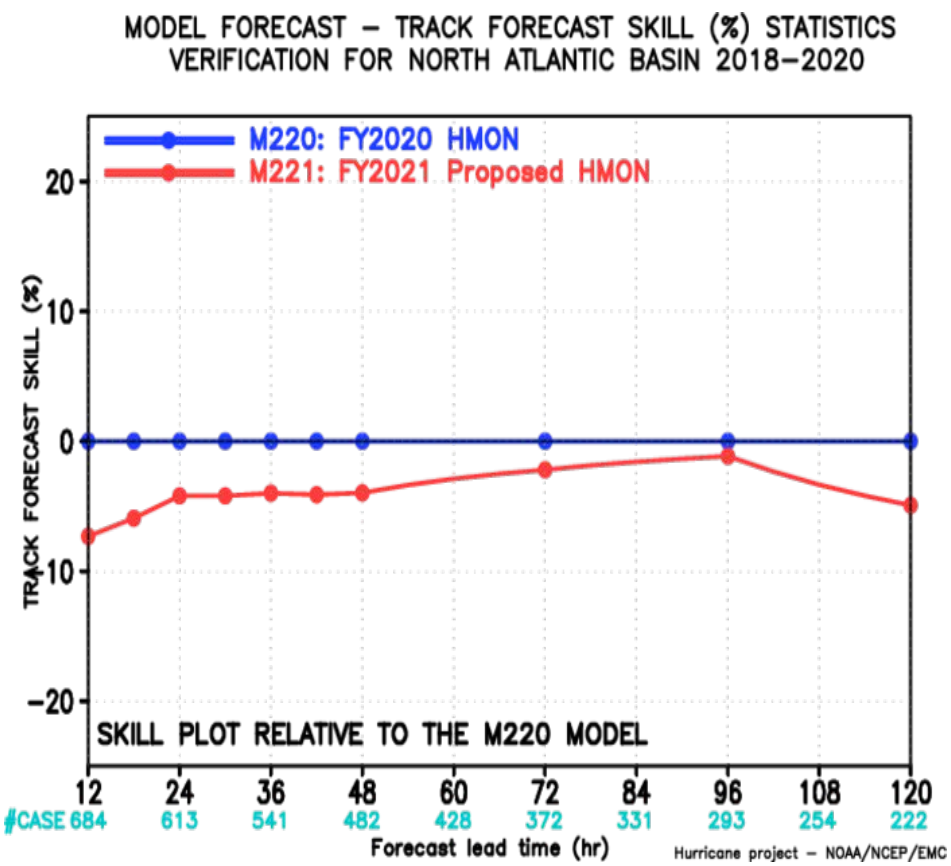
H221: HWRF (with GFS v16) vs H220 Operational HWRF for EPAC Basin (Late model)

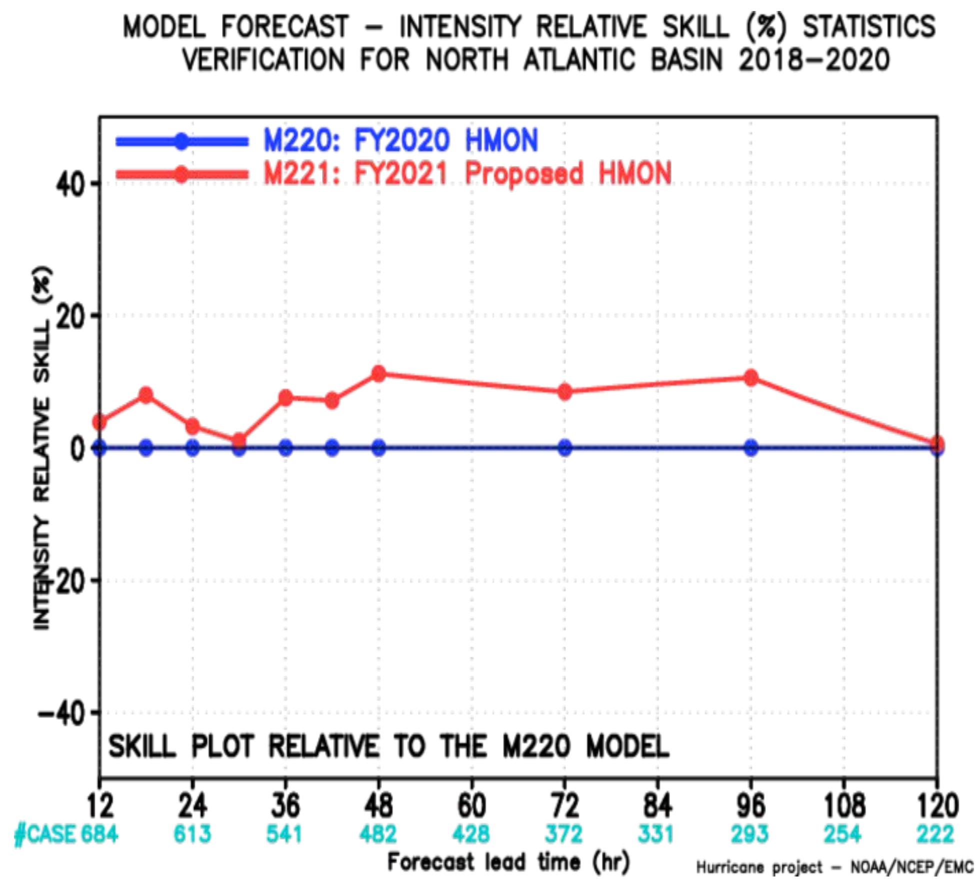




For EPAC basin, H221 shows significantly improved track skill for all lead times with skill > 20% at Days 4 and 5. Intensity skill is negative between hrs 30-60 but positive for longer lead times at Days 3-5. Overall, intensity skill for H221 is neutral.

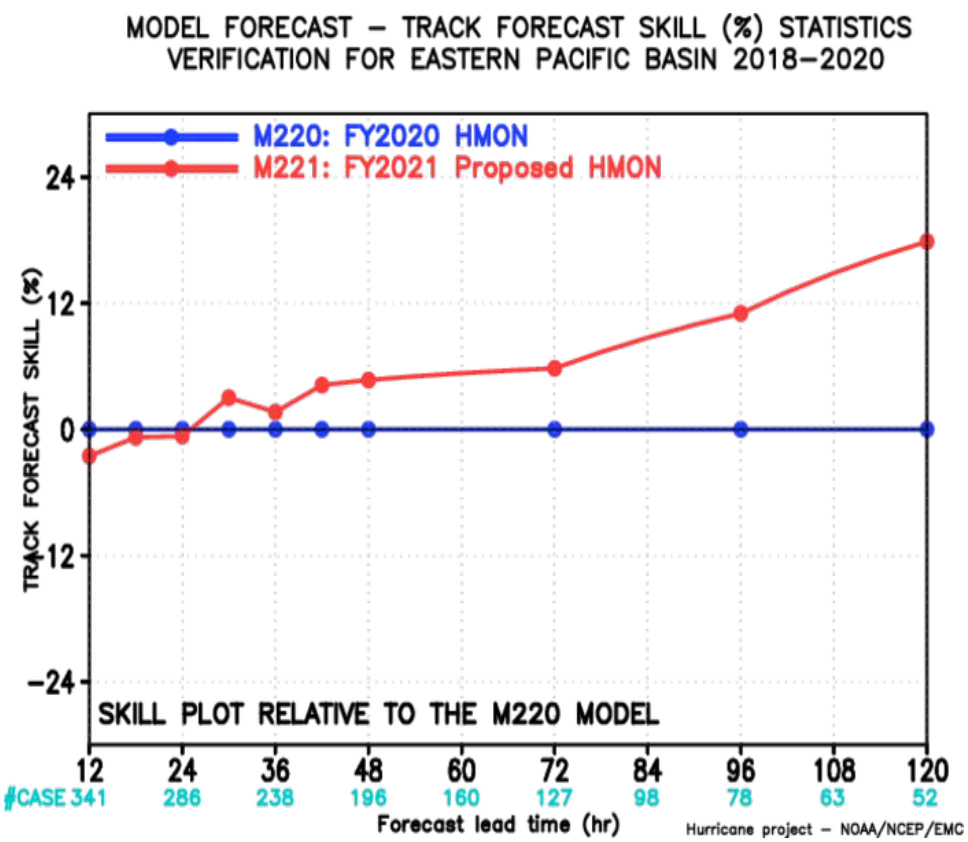
M221: HMON (with GFS v16) vs M220 Operational HMON for NATL Basin (Late model)

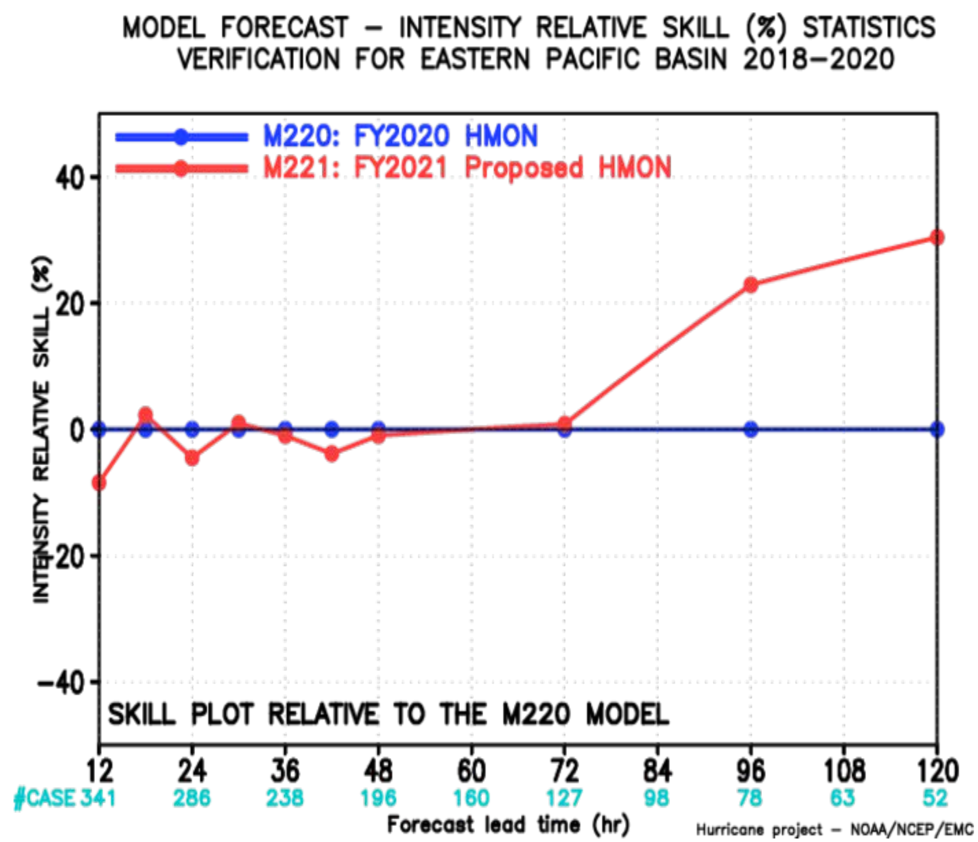




M221 is behind on track skill but within 5% of most lead times after Day 1. Intensity skill is ahead at all lead times as compared to M220 reaching ~10% from Days 2-4.

M221: HMON (with GFS v16) vs M220: Operational HMON for EPAC Basin (Late model)





For EPAC basin, M221 shows **significantly** improved track skill for all lead times after Day 1 with skill reaching > 10% at Days 4 and 5. Intensity skill is neutral up to Day 3 and then much improved for Days 4 and 5.

GEFS.v12 w/GFSv16 Evaluation

Configurations

n_sponge=23

Initial perturbations from EnKF f06, with a 20% reduction globally and all vertical levels (3D)

Experiment periods and verifications

Against own analysis and the GEFSv11 (operation) is the one to compare

Summer month: 7/15 - 8/15/2020

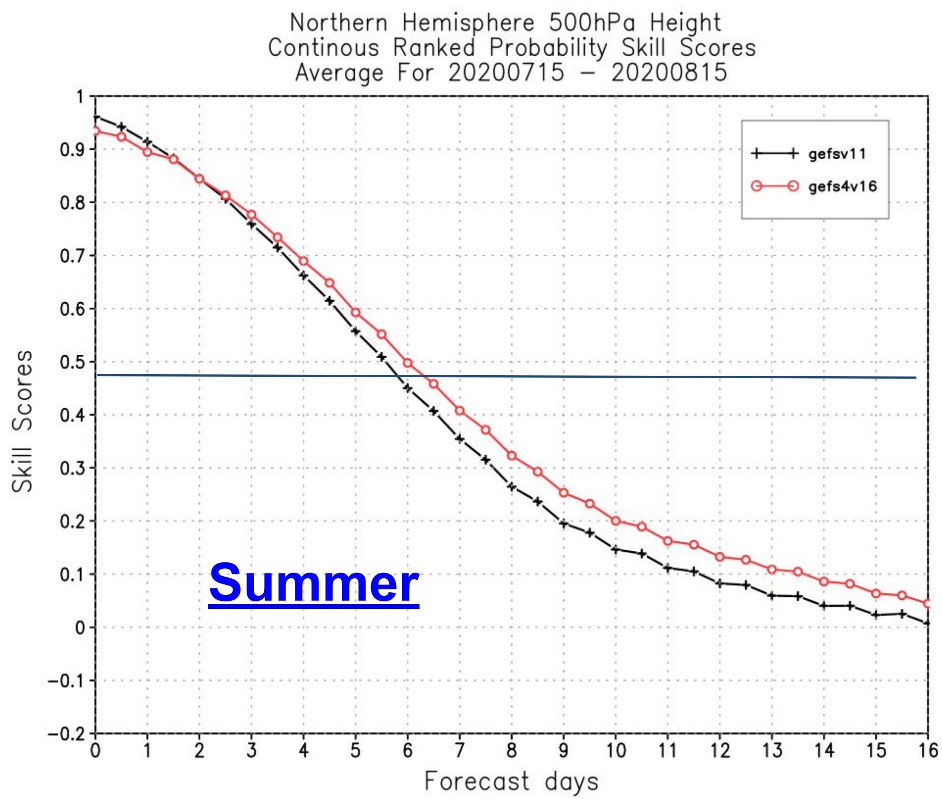
https://www.emc.ncep.noaa.gov/gc_wmb/bfu/nemsfv3gefs/fv3_e75s.html

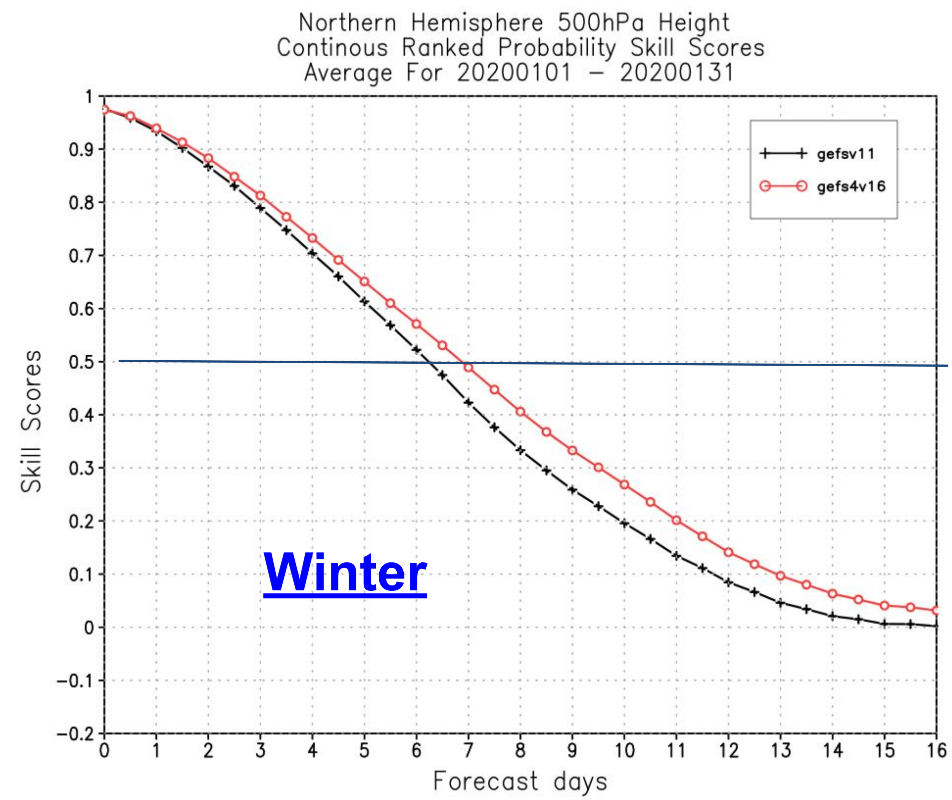
https://www.emc.ncep.noaa.gov/gmb/yluo/GEFS_VRFY/GEFSv12_gfsv16retro_ENSQPFvrfy_sum2020.html

Winter month 1/1 - 1/31/2020

https://www.emc.ncep.noaa.gov/gc_wmb/bfu/nemsfv3gefs/fv3_e75w.html

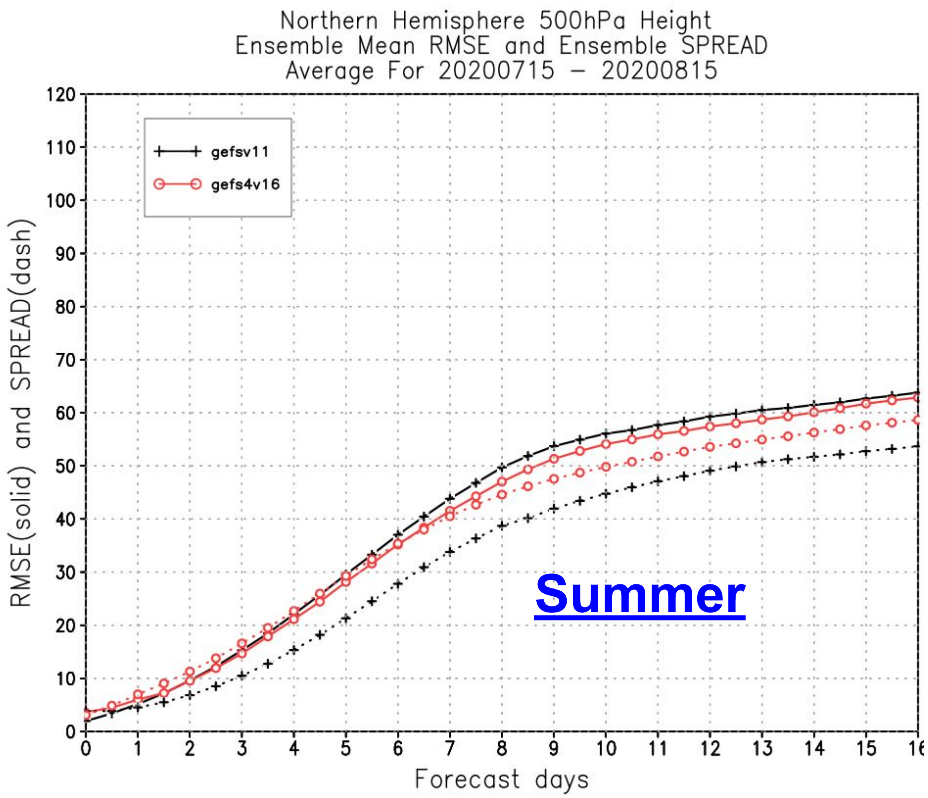
NH 500hPa CRPS

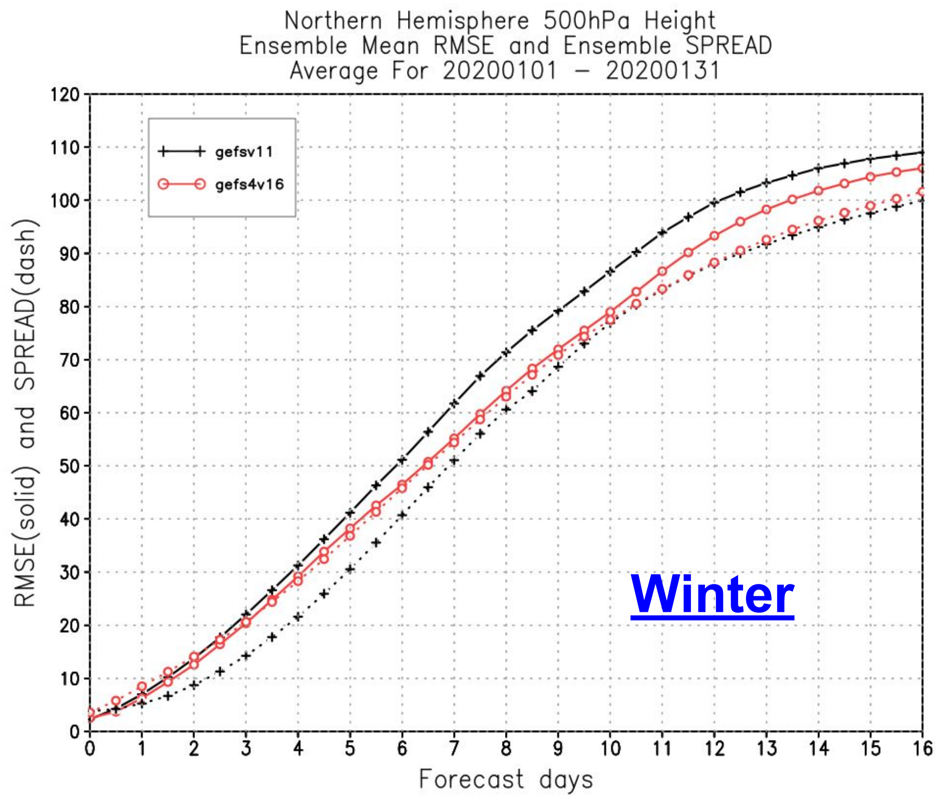




Summary: *CRPS (Continuous Ranked Probability Skill Score (CRPS) is a measure to evaluate ensemble probabilistic forecast (or forecast distribution). Comparing to current operation (GEFSv11), there are about 20 hours improvement for both summer and winter. We have seen some early degradation for summer that may be from different initial analyses. GFSv16 has introduced 127 vertical levels, and GEFSv12 is till running 64 vertical levels, an adjustment may be required for.*

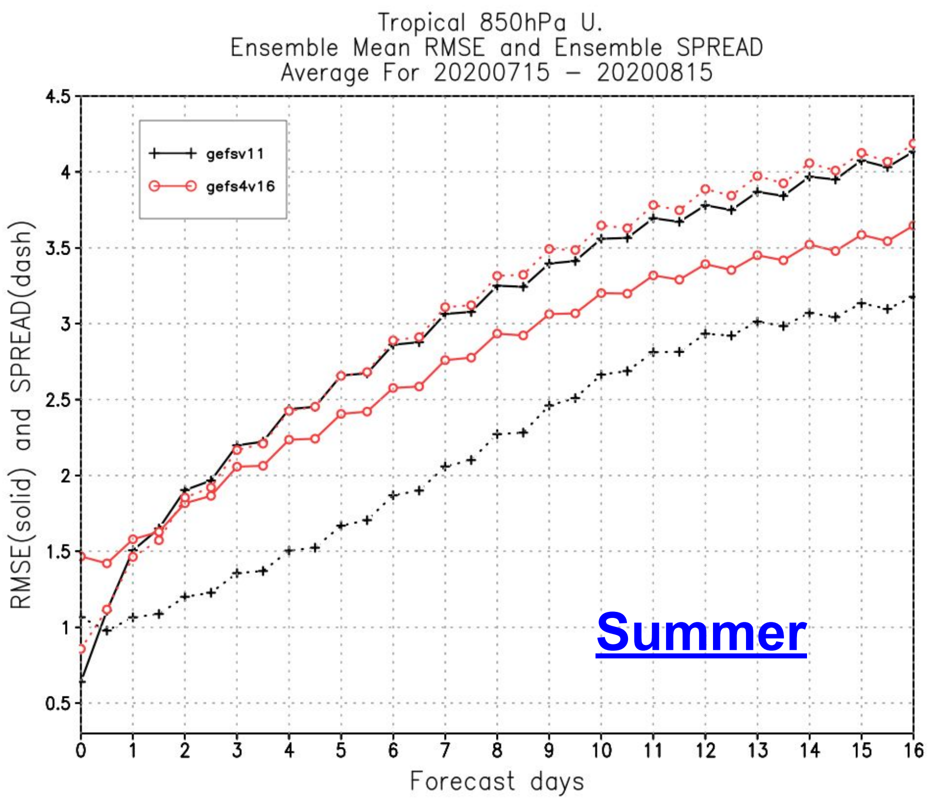
NH 500hPa Ensemble Mean RMS and Ensemble Spread

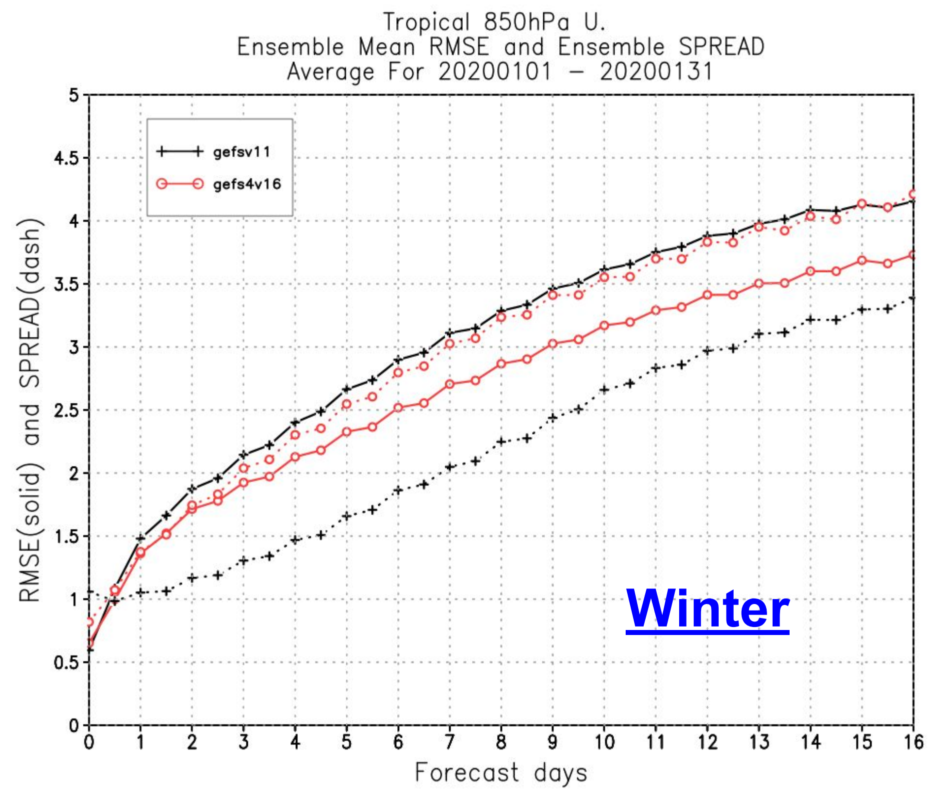




Summary: *After 20% globally reduced initial perturbations from EnKF 6hr forecast, we have seen a reasonable spread for short lead-time, and very comfortable error-spread ratio for all the lead-time. The ensemble mean RMS error are reduced and ensemble spread increased for all extended forecasts.*

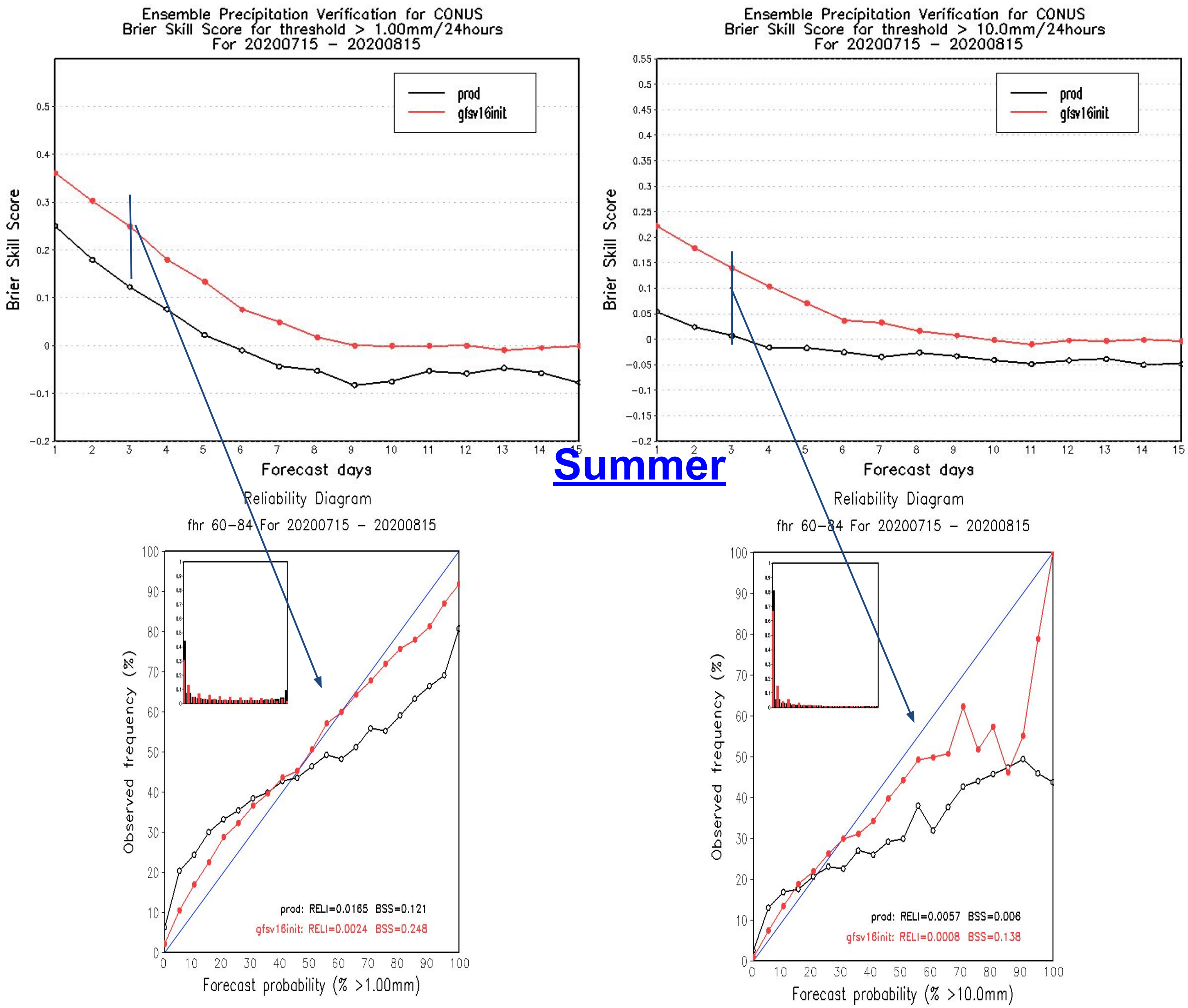
Trop 850hPa Wind RMS and Spread





Summary: *In the tropical area, the overall over-dispersion from GEFSv12 is awared. However, tropical analysis did have large uncertainty. The RMS error could be over/under estimated from single analysis. The consensus analysis (NCEP+CMC+EC+UK) has been used to evaluate the tropical errors (not sure here) which indicated that the GEFSv12 is over-dispersed in the tropical area.*

QPF Brier Skill Score (BSS) and reliability diagram



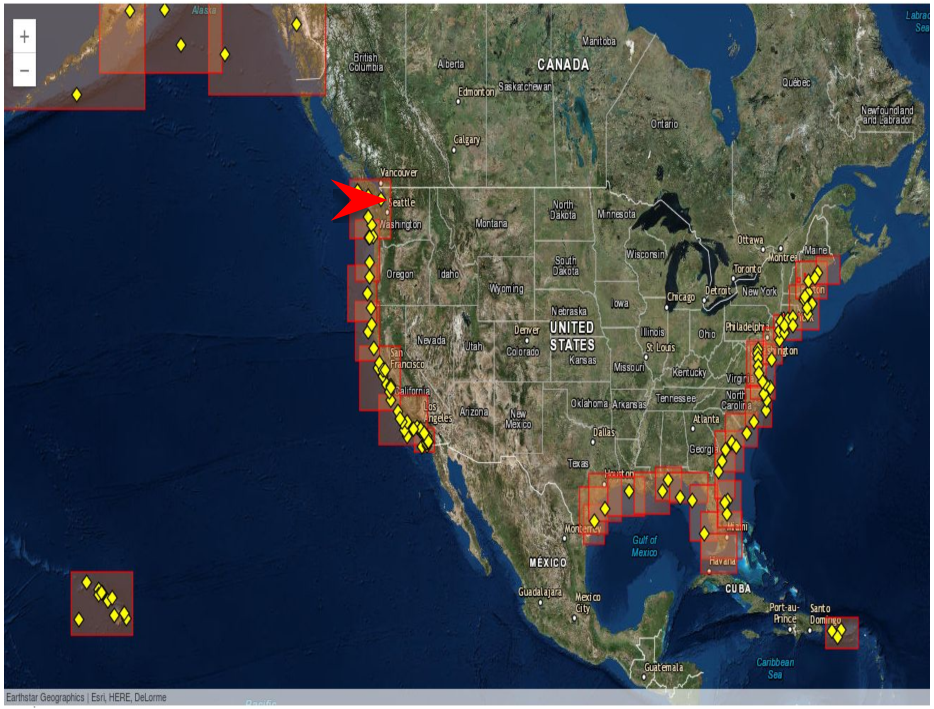
There is **significant improvement** of PQPF forecasts for all threats and all forecast lead-time. The precipitation forecasts are much reliable than current operation (GEFSv11).

Nearshore Wave Prediction System (NWPS) GFSv16 Downstream Evaluation

- Impact of GFSv16-Wave BCs
  - WFO Hawaii example
  - WFO San Juan example

- Impact of GFSv16 U10 atmospheric forcing

[Link](#) to detailed evaluation

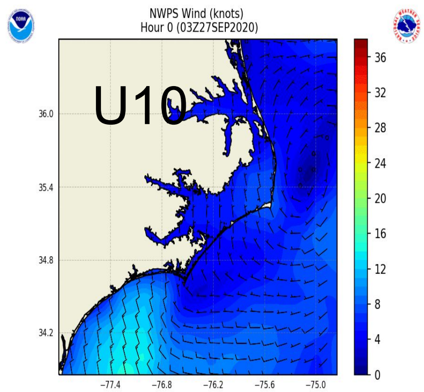


- 36 coastal WFO model domains
- Each uses wave boundary conditions from Global WW3
- U10 wind fields are used as fail-over in case on-demand GFE forecaster wind fields are unavailable

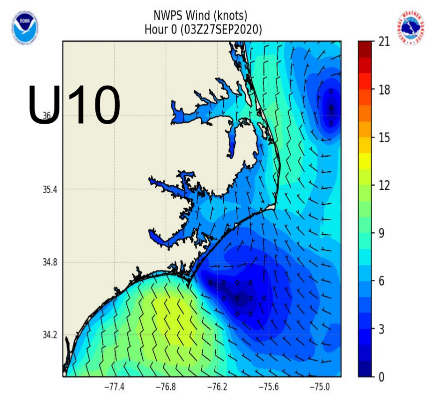
NWPS Evaluation
Summary and Recommendation

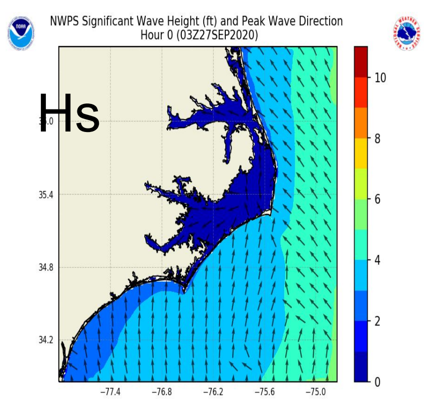
WFO Morehead City: GFS U10 fail-over forcing

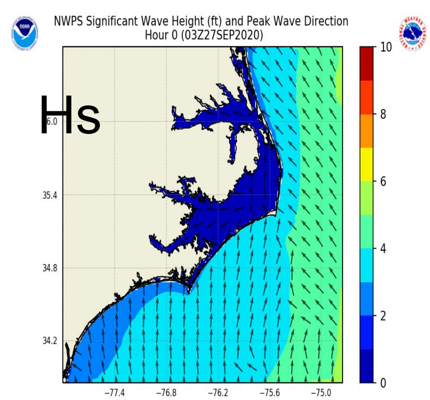
NWPS v1.3 Para with GFSv15



NWPS v1.3 Para with GFSv16

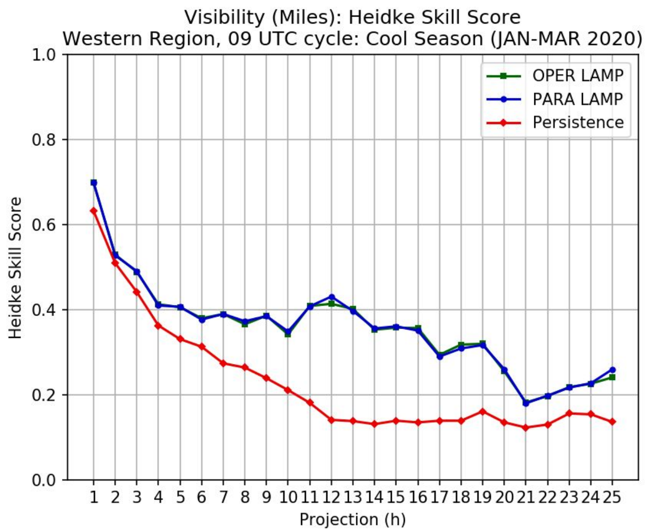






- Changes from GFSv15 to GFSv16 do not adversely affect downstream NWPS.
- New wave boundary conditions from GFSv16-Wave appear realistic, as seen in NWPS swell fields (energy over <0.1 Hz band).
- The wave boundary condition switch from Global WW3 Multi_1 to GFSv16-Wave has only minor impact on downstream NWPS wave results.
- The upgrade of U10 forcing fields from GFSv15 to GFSv16 has only minor impact on downstream NWPS wave results (used in case of fail-over).
- **Recommendation:** Proceed with implementing GFSv16 from the point of view of downstream NWPS.

Localized Aviation MOS Program (LAMP) Evaluation and Recommendation



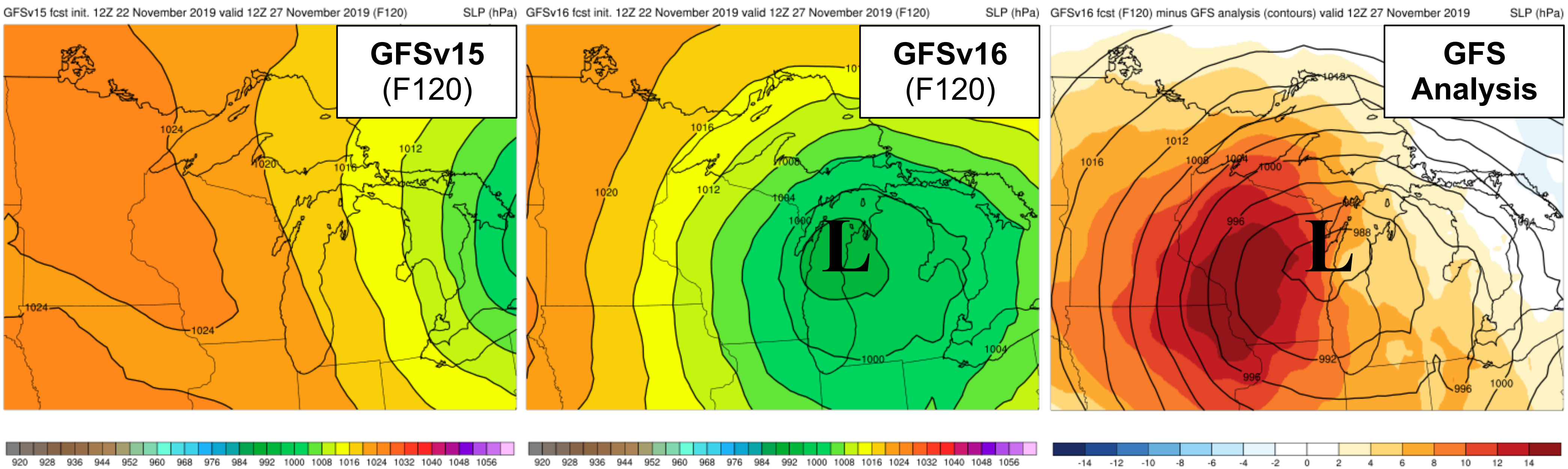
We have completed a somewhat limited impact assessment on LAMP from the GFSv16 upgrade. Regional and overall verification plots for temperature, dew point, wind speed, ceiling height, and visibility are available on google drive here:

https://drive.google.com/drive/folders/1gGJMK7kfWpVMHO-dgwQ1Jk6yEMoYKTex?usp=sharing

While we do see some differences particularly in bias scores, they are not egregious. Our biggest concern was the impact on C&V for the cool season. We're pleased to report that we do not see any alarming differences between OPER LAMP and PARA LAMP for C&V for the sample we verified. We do plan to verify other elements such as convection, lightning, and POP, but do not expect to see a big impact since the GFS is a small component for those systems. We will continue to monitor performance going forward but we do not expect to see major impacts on LAMP from this upgrade. We are planning on revamping the entire system for the RRFS in FY23, so any biases that get introduced with GFSv16 (while small) we should be able to correct in the next implementation. In short, we are **thumbs up re: GFSv16 implementation**.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

The Evaluation of GFSv16



Geoff Manikin, Alicia Bentley, Shannon Shields,
Philippe Papin, Logan Dawson, Chris MacIntosh

EMC Model Evaluation Group Presentation to the EMC Director

30 September 2020

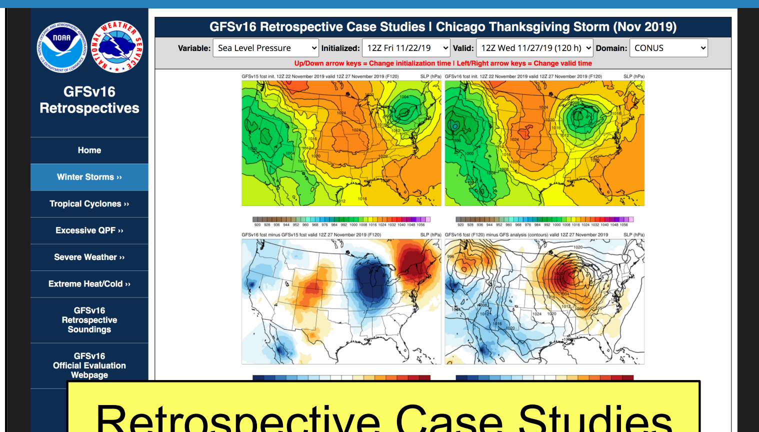
GFSv16 Evaluation Details

- The GFSv16 official evaluation included analyses of:
  - Retrospectives (5/5/19–5/18/20; added 8/31/18–10/12/18)
    - Statistics
    - 50 Case Studies
  - Real-time Parallel (5/19/20–present)
    - Statistics
    - Representative examples

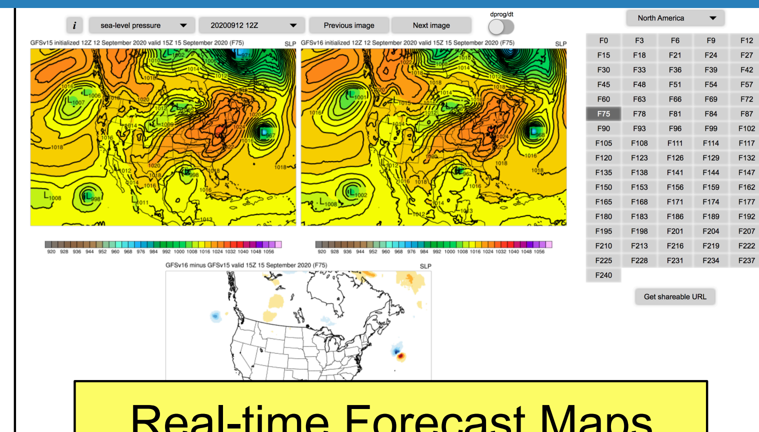
- The GFSv16 official evaluation also incorporated the findings of an STI team of NWS SOOs tasked with analyzing GFSv16 forecasts in a testbed format

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

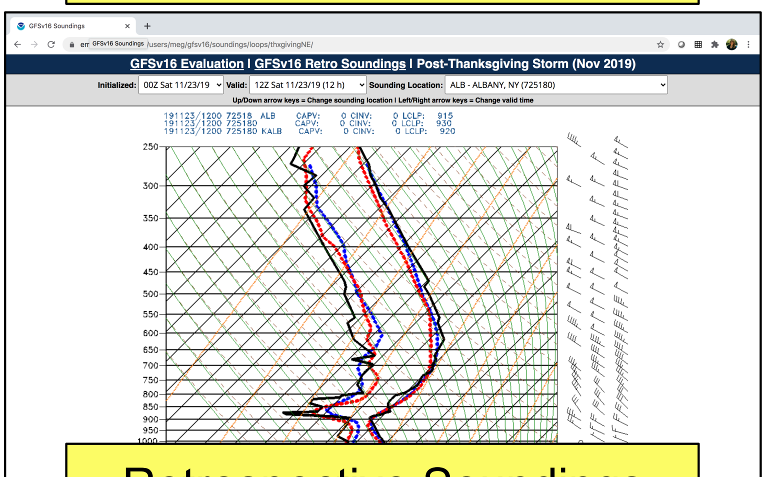
GFSv16 Evaluation Resources (VPPPG)



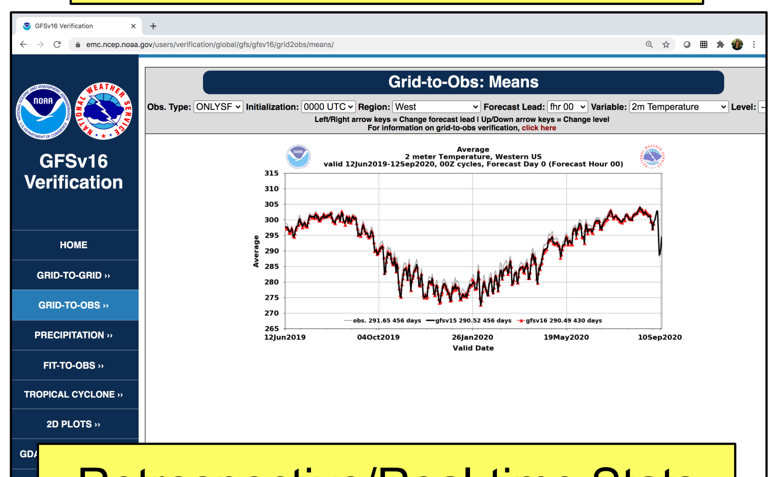
Retrospective Case Studies



Real-time Forecast Maps



Retrospective Soundings



Retrospective/Real-time Stats

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Summary of GFSv16 Verification Statistics

Improvement Neutral Degradation

Parameter	Remarks
200/250-hPa Winds	Existing NH & SH low bias in GFSv15 worse in GFSv16, but mitigated in tropics & over U.S. RMSE higher initially and in tropics, but reduced in medium range for NH & SH.
500-hPa Height	Improved NH & SH 500-hPa AC scores in medium-range. Lower NH & SH RMSE at most lead times.
850-hPa Winds	Existing NH & SH low bias in GFSv15 made worse in GFSv16, especially in tropics. RMSE higher initially, but significantly reduced in short-to-medium range for NH & SH.
850-hPa Temp.	Mitigated cold bias seen in GFSv15 during NH cool season in medium-range. Lower NH & SH RMSE in the medium-range. Colder temperatures in short-range, especially in NH cool season short-range forecast.
1000-hPa Height	Improved NH & SH 1000-hPa AC scores at most lead times. Lower NH & SH RMSE at most lead times.
10-m Winds	Little change in bias overall over the U.S. Slight but significant decrease in RMSE in short-to-medium range.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Summary of GFSv16 Verification Statistics

Parameter	Remarks
2-m Temp.	Mitigated cold bias over CONUS during cool season at longer lead times. Lower RMSE over CONUS during cool season at longer lead times. Inconsistent CONUS bias signal comparing 2019 & 2020 warm seasons. Higher RMSE over Eastern U.S. during warm season at most lead times.
Low-Level Moisture/ 2-m Dewpoint	Introduced significant low bias over CONUS at short lead times (likely related to soil moisture); bias worse in summer, more negligible in winter. RMSE increased in short-range but reduced in medium-range.
Precipitation	U.S. ETS significantly improved at most 24h precipitation thresholds at most lead times. U.S. high bias reduced at low thresholds; low bias reduced at medium to high thresholds.
CAPE (real-time)	RMSE increased; existing low GFSv15 bias made significantly worse due to feedback mechanism from drier soil.
TC Track	Reduced errors overall for strong TCs but slow & right-of-track biases at long lead times.
TC Size	Reduced low bias in 34-kt wind radii as GFSv16 produces larger and stronger TCs.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Summary of GFSv16 Verification Statistics

Parameter	Remarks
TC Intensity	Lower intensity error at almost all lead times in North Atlantic. Less of a weak bias at longer lead times in North Atlantic. Similar intensity error at almost all lead times in East Pacific. Slightly less of a weak bias at longer lead times in East Pacific. Tendency to strengthen most TCs in the long-range.
TC Genesis	POD improved in both NATL & EPAC, but FAR also increased. Overall increase in CSI. Increased TC genesis lead time, but many TCs still completely missed by both models. Too many false alarms from 50°-70°W.
HWRF	Improved track & intensity forecast when initialized with GFSv16.
HMON	Degraded track forecast for NATL when initialized with GFSv16 but improved for EPAC. Intensity forecast improved.
Waves	Lower globally-averaged RMSE & bias for Sig. Wave Height in GFS-Wave. Some regional degradation of waves forecasts where the high-resolution Multi-1 output grids are not available in GFS-Wave

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Overall Atmospheric Impressions of GFSv16

Region	Recommendation	Key Remarks
Eastern Region	**Implement**	Improvement with PBL structure, cold-air damming, and medium-range Ptype. Low CAPE bias and low-level cold bias in cool season (short-term) concerns.
Central Region	**Implement**	Notable synoptic improvements in medium-range. Small improvement in low-level PBL in cold season; cases of overmixed PBL in warm season. Improvement in low-level temps (reduction in 2-m cold bias).
Southern Region	**Implement**	Evidence of improved temp. profile in shallow, cold air masses. Larger TC FAR and right-of-track bias. Low warm season CAPE.
Western Region	**Implement**	Better details in QPF, winds, temps/RH. Improvement in low-level temps. Higher resolution. Noticeable improvement in the mid-range.
Alaska Region	**Implement**	More significant cyclones, improved PBL & better cold low-level temp. Large 2-m temp. change flipping from warm to cold bias.
Pacific Region	**Implement**	Significantly better TC genesis; some increase in false alarms. With stronger TCs, GFSv16 has better track, size, and intensity.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Overall Atmospheric Impressions of GFSv16

Center	Recommendation	Key Remarks
WPC	Implement	Progressive bias in GFSv15 appears mitigated (handles synoptic scale better). Better captured higher QPF and cold-air damming in East. Tendency to over-forecast sleet and low CAPE bias concerns.
SPC	Neutral	Improved forecasts of frontal boundaries for a few cases. Degradation in low-level temps during warm season. Drier soil moisture exacerbates the 2-m dewpoint and low instability biases when coupled with overmixing bias.
NHC	Neutral	Improvements to TC intensity/wind radii and increased lead time and false alarms for genesis. Increased right-of-track & along-track bias.
AWC	Implement	Mitigation of progressive bias seen in GFSv15. Better ability to capture cold-air damming events. Better jet stream forecasts. Slightly better PBL. Improvement in low-level temps.
CPC	Implement	Slightly improved 500-hPa heights. Temps get warmer with forecast time in winter latitudes. Winter (summer) zonal winds decrease (increase) with lead time. Ozone in polar night issue.
First Energy Corp	Implement	Many parameters behaved similarly between the two versions. More realistic surface pressure intensity. Views GFSv16 as a foundational component of the UFS

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Overall Impressions of GFSv16 Waves

Center/Region	Recommendation	Key Remarks
Ocean Prediction Center	**Do NOT Implement**	No improved reliability in forecasts of ocean waves. Improvement of wave height bias over the north Pacific, but a consistent lack of improvement over the north Atlantic. Period and direction seemed better, but wind speed and wave height are worse over North Atlantic and west coast which potentially hurts OPC ops. Low bias for the highest wave heights.
National Hurricane Center	**Neutral**	Unsure about improved reliability in forecasts of ocean waves. Slightly higher Hs RMSE 00-48 hr and a slight low wave height bias. GFS-Wave lowest in 95th quantile stats vs. obs and Multi_1. Slightly lower long distance-traveled swell.
Eastern Region	**Implement**	Improved reliability in forecasts of ocean waves. Increase in error/ bias in some areas is smaller than overall improvement. Some concern of operational issues with coarser grids in GFS-Wave.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Waves Concerns

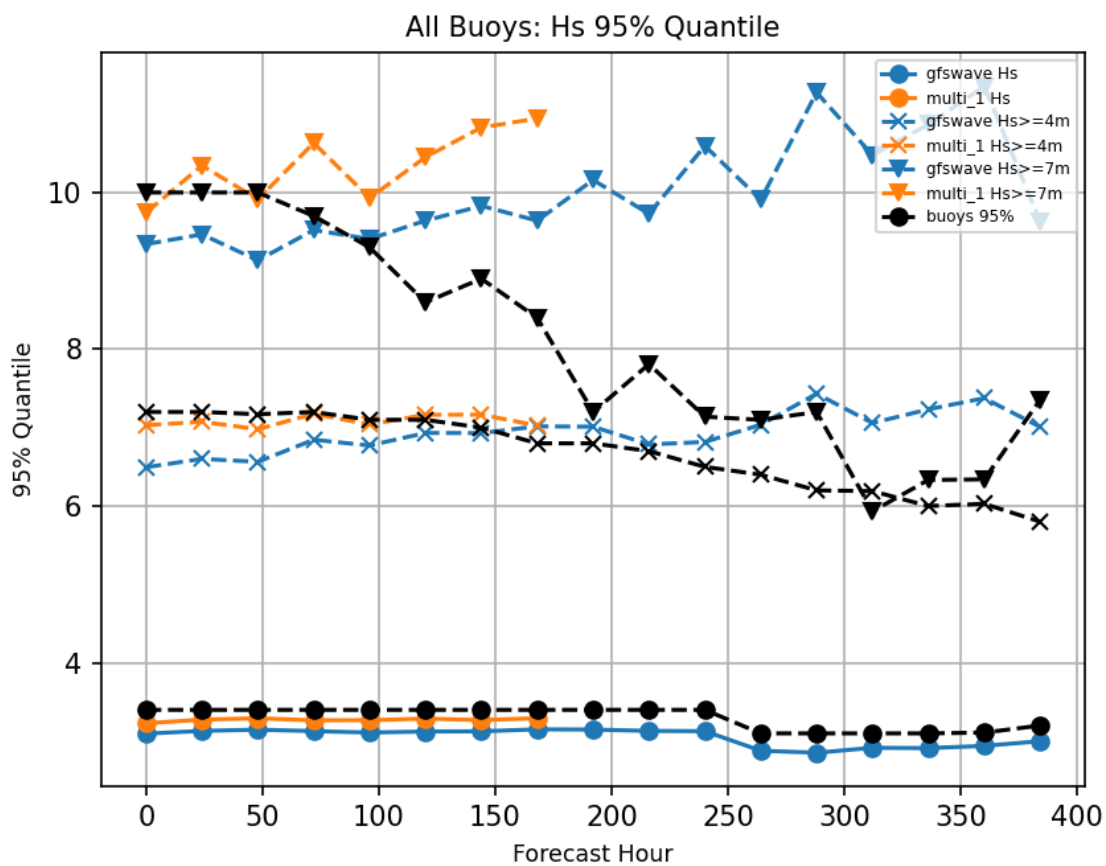
Buoys Overall	Sig Wave Height	Wind Speed	Peak Period
Subjective	Multi-1	GFS Wave	GFS Wave
Objective/month	Multi-1	GFS Wave	GFS Wave
Objective/fcst	Tie	Tie	GFS Wave

Satellite Overall	Sig Wave Height	Wind Speed
Global	GFS Wave	GFS Wave
North Atlantic (at_10m)	Multi-1	Multi-1
US West Coast (wc_10m)	Multi-1	Multi-1

From Deanna Spindler

- Buoys (coastal areas) and satellite data (open ocean) were used to compare the existing Multi-1 global wave model with GFSv16 (GFS Wave - wave component coupled to atmosphere)
- Regionally, Multi-1 has better sig wave heights and wind speeds
- Near the coastal US, Multi-1 performs better, as expected due to the loss of the 4 arcmin grids

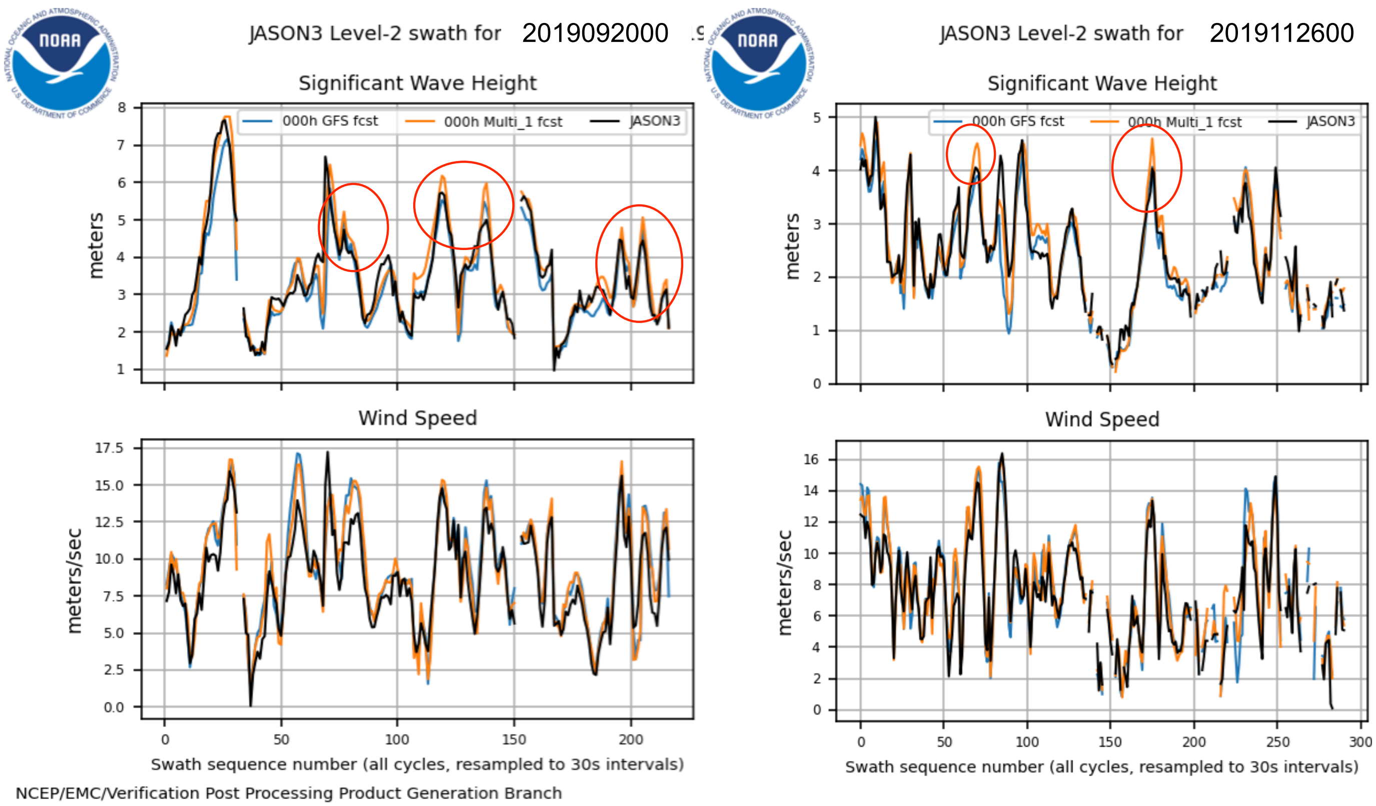
Waves Concerns



- 95% Quantile Wave Height is around 3m, and Multi-1 holds a slight edge
- For larger waves, sample size is quite small, and both Multi-1 and GFS-Wave overpredict “dangerous seas” (according to buoy data), but Multi-1 does better with the larger waves in the early forecast hours
- OPC deems better detection of large waves as critical to their operations

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Waves Concerns



- These examples show that Multi-1 overestimates wave height on the open seas and that GFSv16 may be overall better with the largest waves

- It’s tough to discern this in the stats, since smaller waves dominate

Common Strengths From All Evaluations

- Notable improvements in synoptic-scale performance in the medium-range
  - Progressive bias in GFSv15 appears mitigated with better consistency catching correct solutions earlier
  - Improved frontal positions and QPF

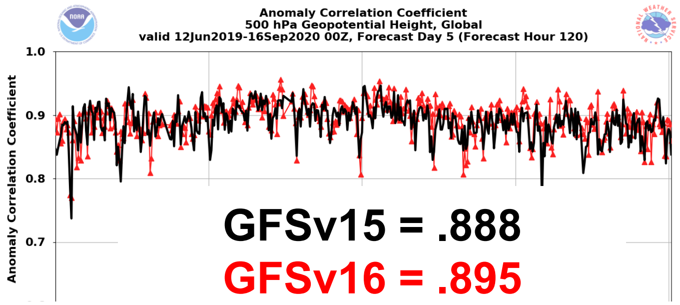
- Improvement in low-level temperature forecasts (mitigation of the winter low-level cold bias)

- Better ability to resolve shallow, cold air masses and some associated cold air damming events

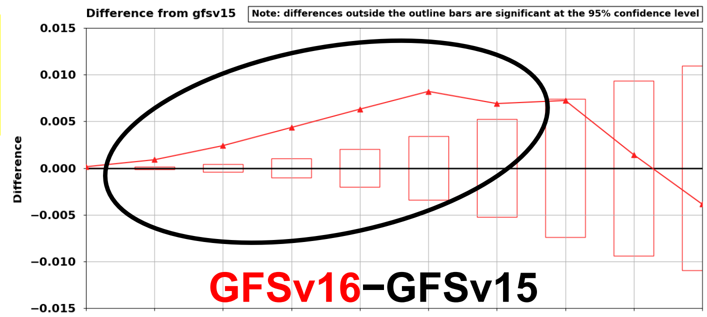
- Improvements to TC intensity and increased lead time for genesis
  - With stronger TCs, GFSv16 has overall better track, size, and intensity

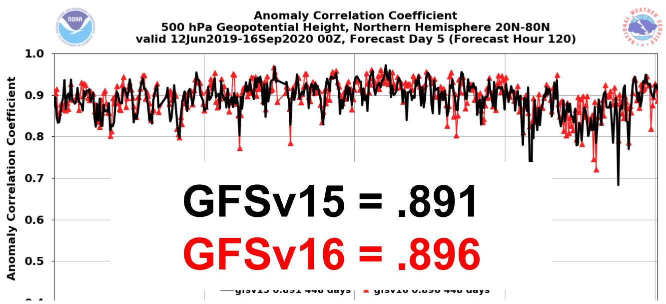
Strengths: 500-hPa AC Scores (Global)

Valid: 6/12/19–9/16/20 (Day 5)

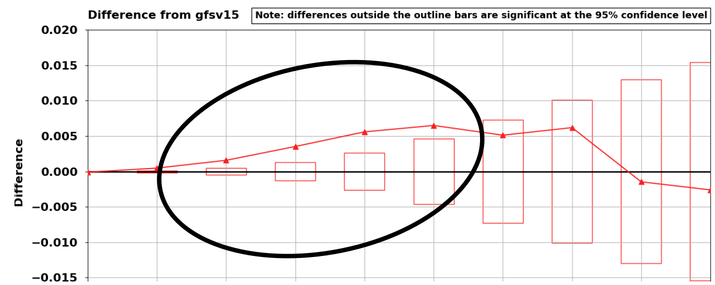


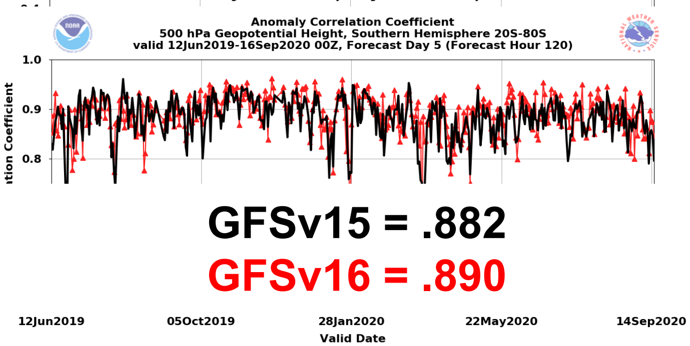
Global



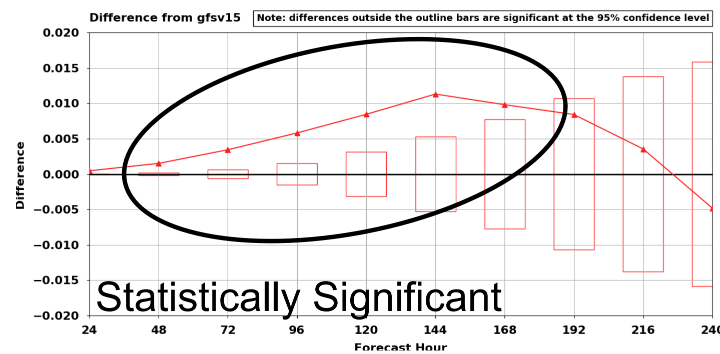


NH





SH



GFSv16 AC Scores (NH 500-hPa Z at Day 5)

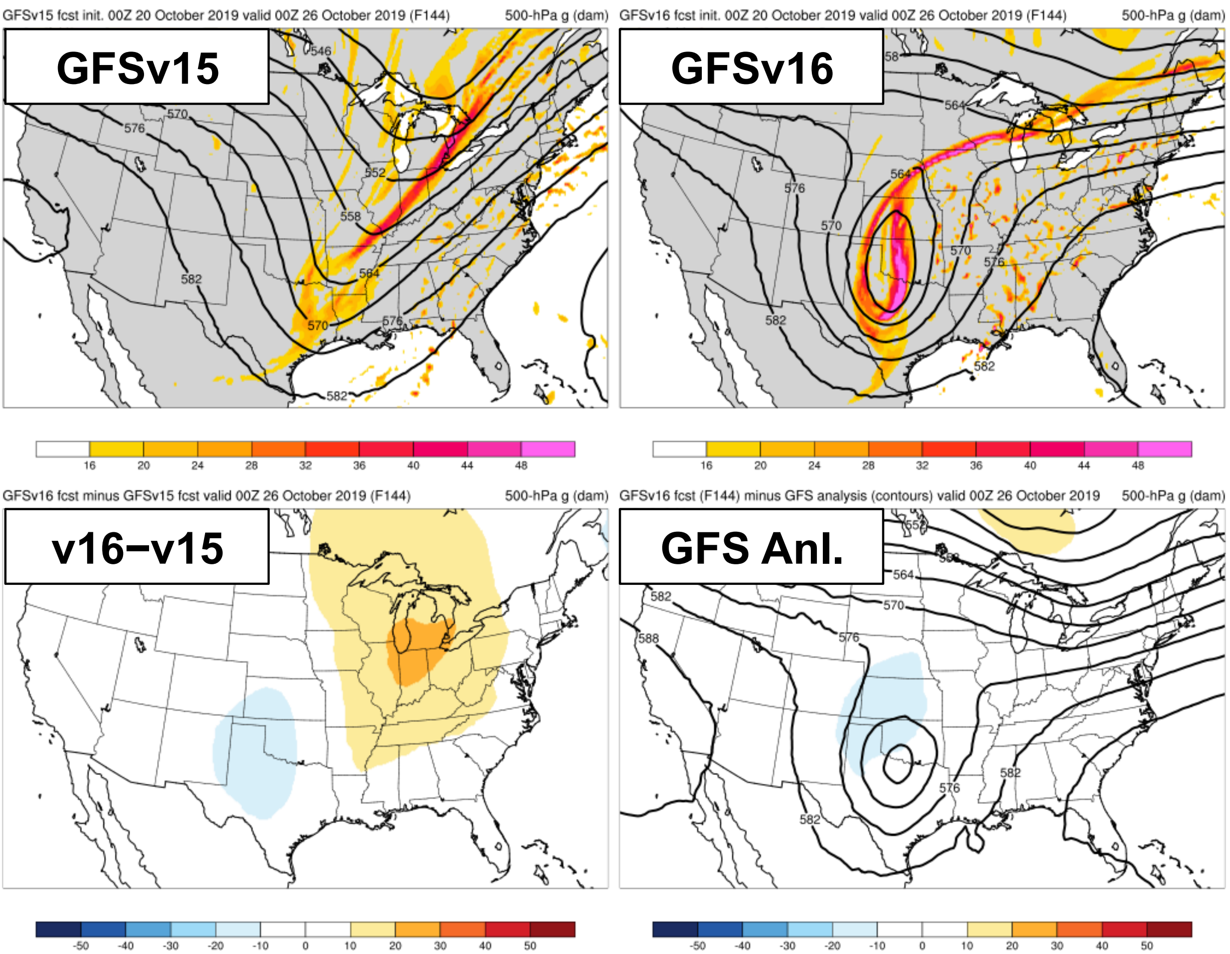
	GFSv15 (OPS)	GFSv16 (RETRO)
Fall 2018	0.916	**0.916**
May 2019	0.880	**0.897**
Summer 2019	0.880	**0.888**
Fall 2019	0.897	**0.901**
Winter/Spring 2020	0.909	**0.913**
Real-Time Parallel	0.864	**0.871**
Full Retro Period	0.890	**0.896**

GFSv16 NWS SOO Team Synoptic Ratings

	GFSv16 vs. GFSv15 Mean Rating (−3 to +3)	% GFSv16 was as good or better than GFSv15	% GFSv16 was worse than GFSv15
Extended Range	**0.35**	**78%**	**22%**
Medium Range	**0.59**	**83%**	**17%**
Short Range	**0.07**	**85%**	**15%**

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Captures Synoptic Pattern Better

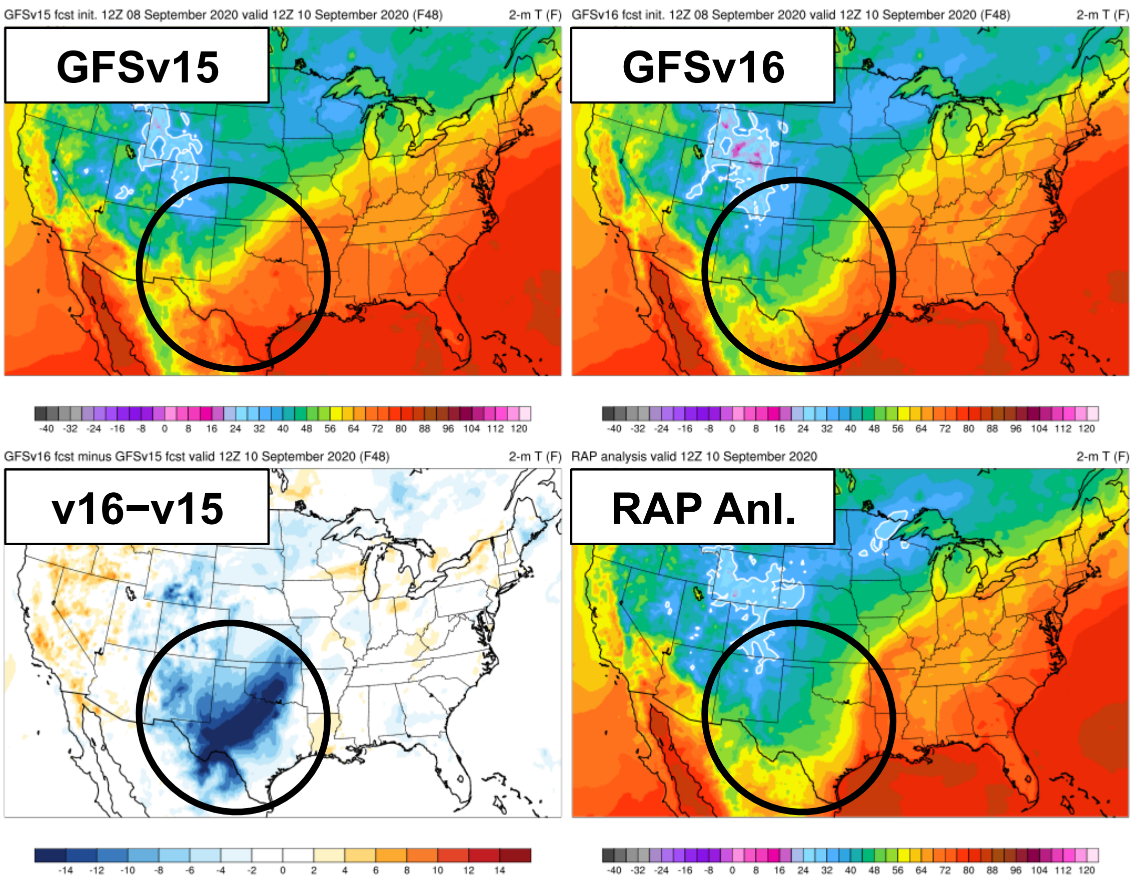


TC Olga Case
Fcst: 00z 10/20/20 (F144)
Valid: 00Z 10/26/20

- GFSv16 forecasted the location of this and other cutoff lows earlier and more consistently than GFSv15, with some mitigation of the progressive issue noted in the GFSv15 evaluation

- Several evaluators noted that GFSv16 showed more run-to-run continuity than GFSv15

Strengths: Position of Frontal Boundaries



Cold Front Example
Fcst: 12z 09/08/20 (F048)
Valid: 12Z 09/10/20

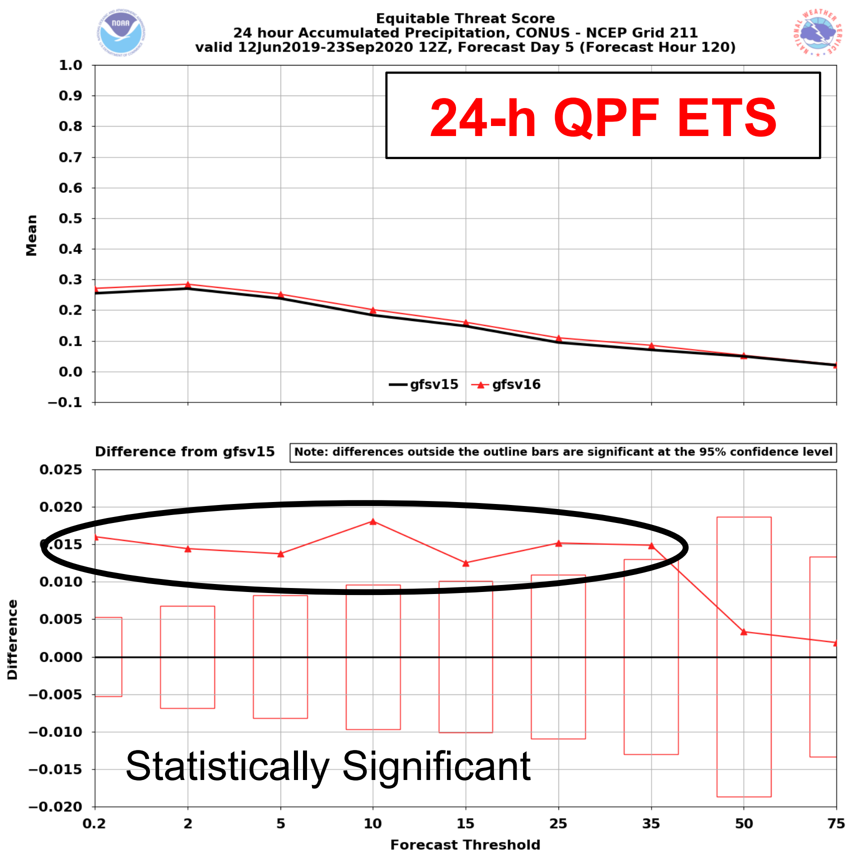
- GFSv16 forecasted the position of the cold front more correctly and consistently than GFSv15

Thanks to Steverino Silberberg (AWC) from NWS SOO Team

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Improved QPF ETS and Bias

Valid: 6/12/19–9/23/20 (F120)

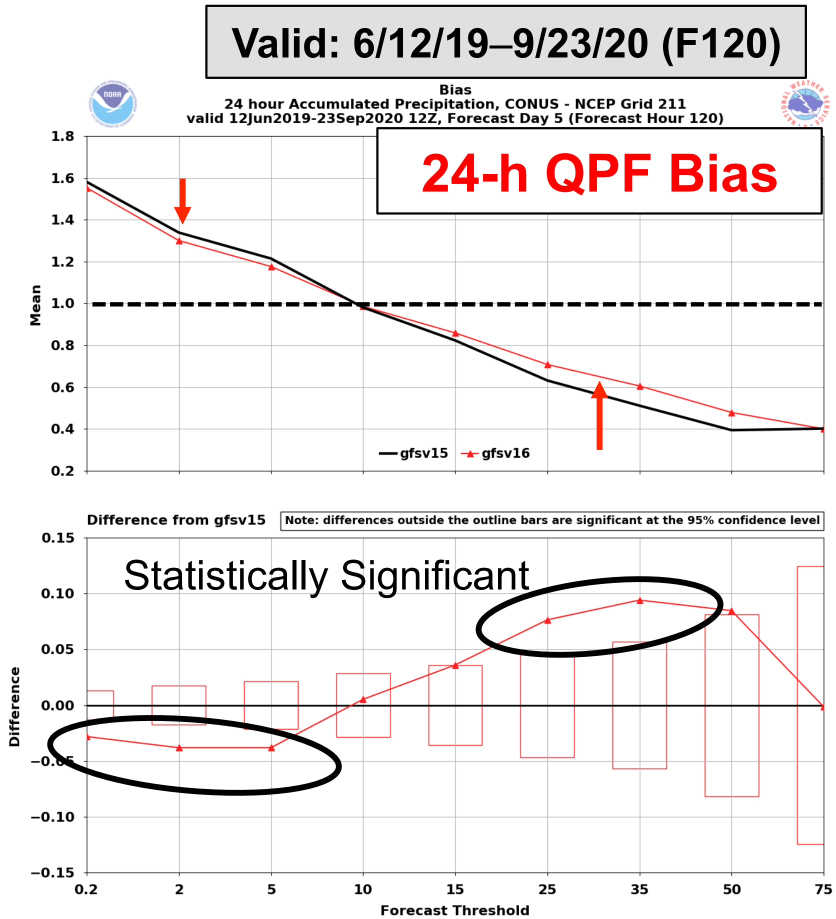


Equitable Threat Score (ETS)

- 24-h QPF improvements appear most pronounced in the medium range, which is consistent w/ improved 500-hPa AC scores
  - **F120:** Statistically significant improvement at 0.2–35 mm thresholds

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Improved QPF ETS and Bias

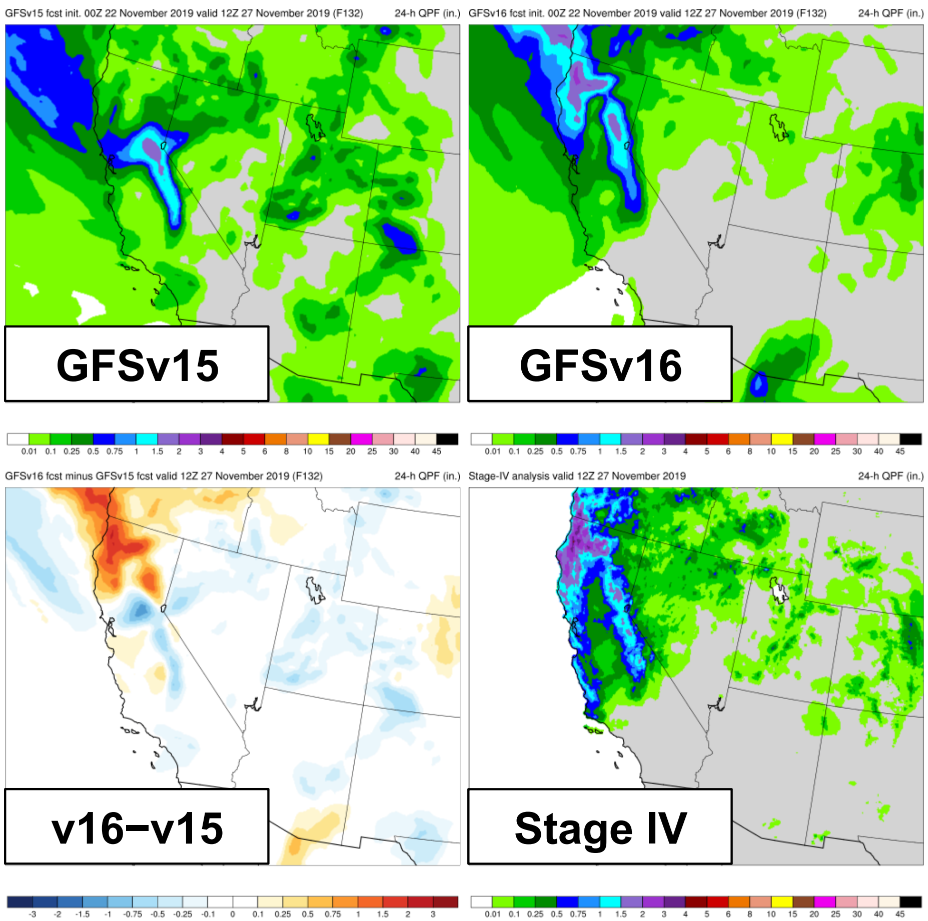


Bias

- 24-h QPF bias improvements also most pronounced in the medium range
- Reduction of the high bias at lower QPF thresholds is statistically significant
- Reduction of the low bias at medium-to-high QPF thresholds is statistically significant
- Overall bias improvement is seen in the short range as well

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Improved QPF ETS and Bias



West Coast Bomb Cyclone Case
Fcst: 00z 11/22/19 (F132)
Valid: 12Z 11/27/19

- GFSv16 consistently had (correctly) higher QPF amounts inland over N California and Oregon for this case

Common Strengths From All Evaluations

- Notable improvements in synoptic-scale performance in the medium-range
  - Progressive bias in GFSv15 appears mitigated with better consistency catching correct solutions earlier
  - Improved frontal positions and QPF

- Improvement in low-level temperature forecasts (mitigation of the winter low-level cold bias)

- Better ability to resolve shallow, cold air masses and some associated cold air damming events

- Improvements to TC intensity and increased lead time for genesis
  - With stronger TCs, GFSv16 has overall better track, size, and intensity

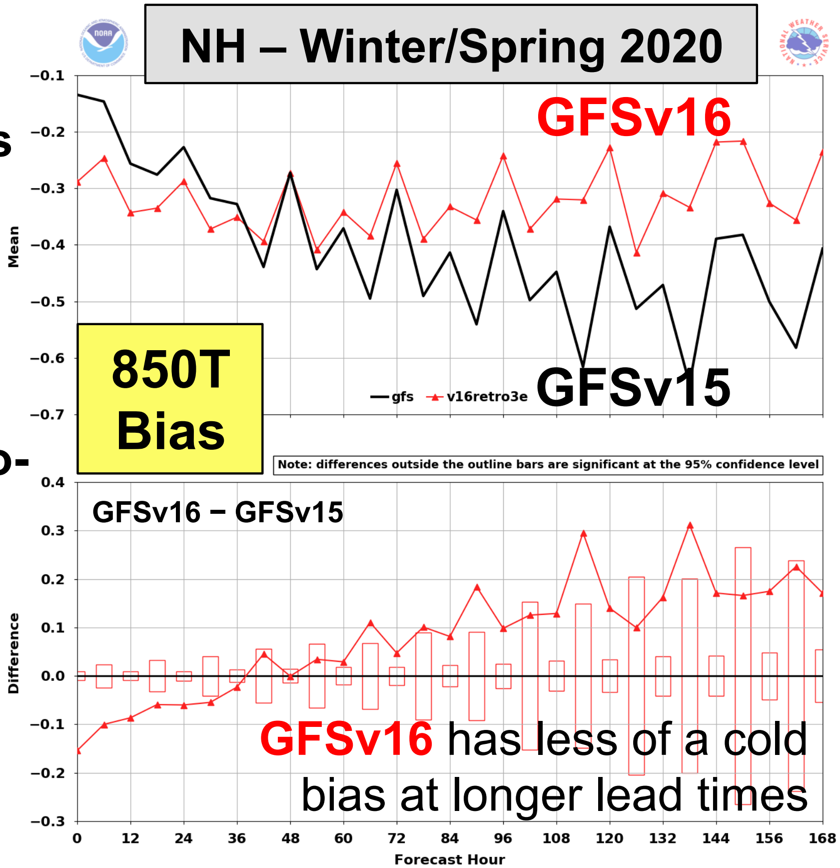
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

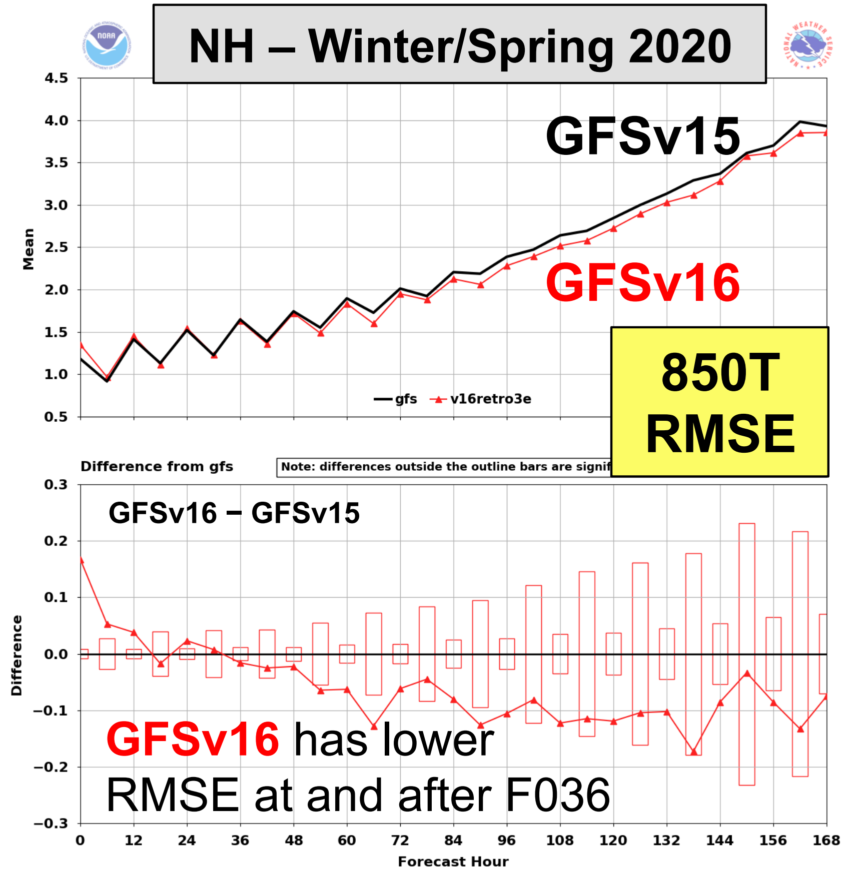
Strengths: Mitigated Low-Level Cold Bias

GFSv15 has a known <u>low-level cold bias</u> that gets worse with lead time

00Z cycles

Grid-to-Obs

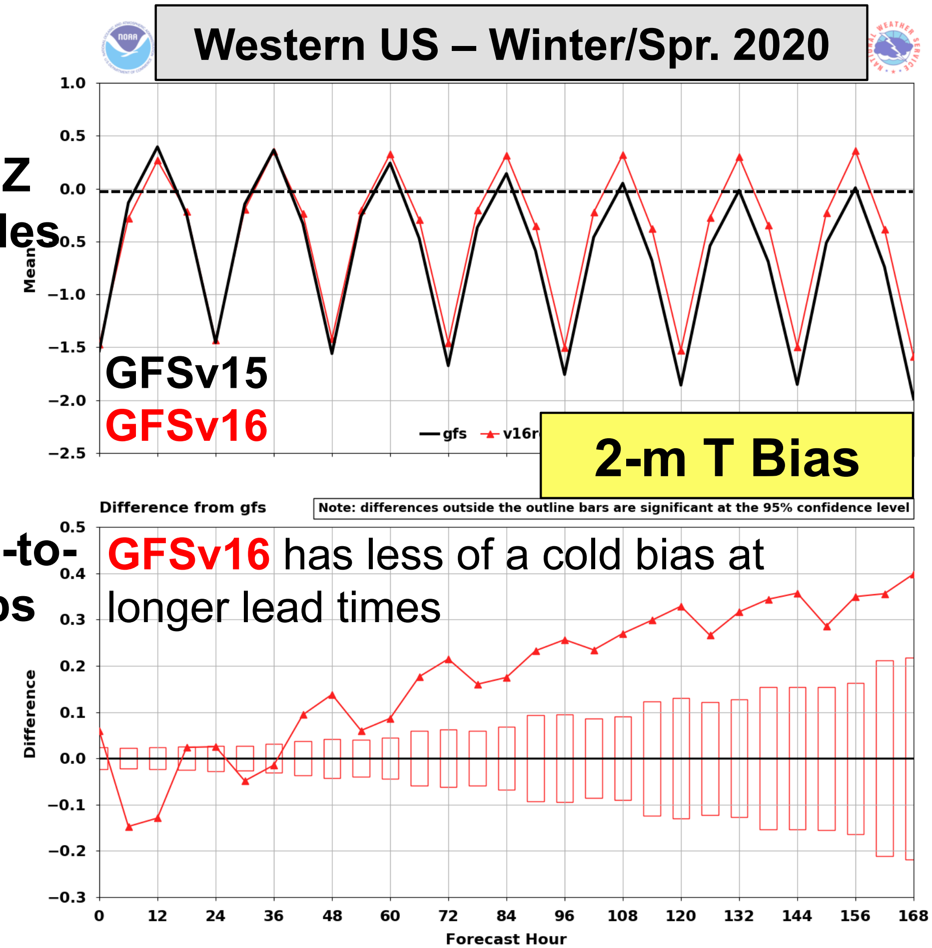


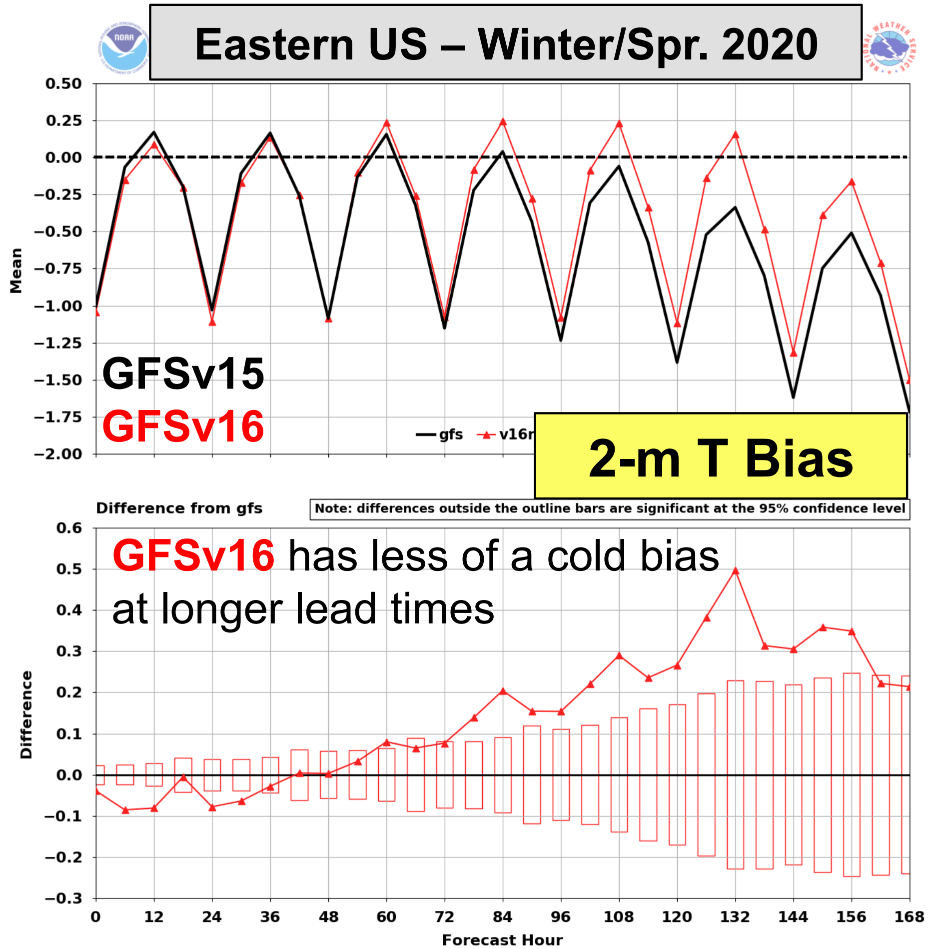


Strengths: Mitigated Low-Level Cold Bias

00Z cycles

Grid-to-Obs





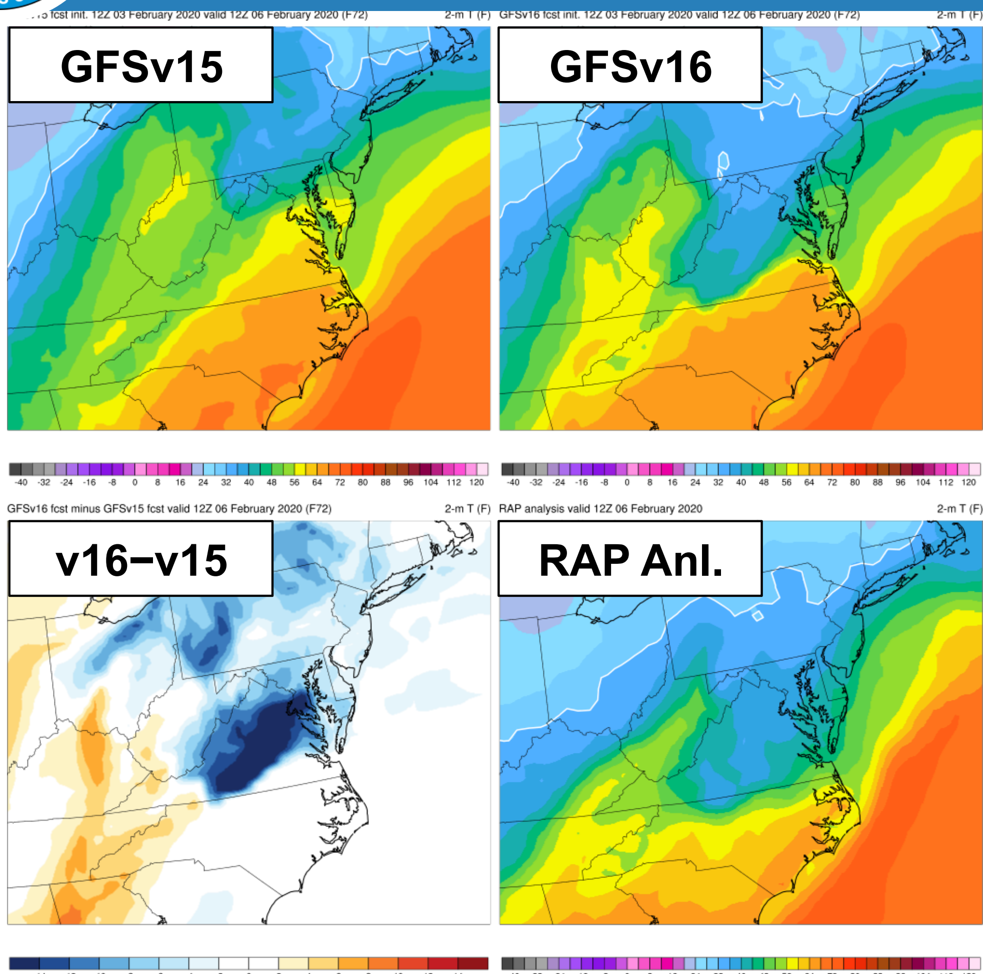
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Common Strengths From All Evaluations

- Notable improvements in synoptic-scale performance in the medium-range
  - Progressive bias in GFSv15 appears mitigated with better consistency, catching correct solutions earlier
  - Improved frontal positions and QPF
- Improvement in low-level temperature forecasts (mitigation of the winter low-level cold bias)
- Better ability to resolve shallow, cold air masses and some associated cold air damming events
- Improvements to TC intensity and increased lead time for genesis
  - With stronger TCs, GFSv16 has overall better track, size, and intensity

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Temps in Shallow, Cold Air Masses



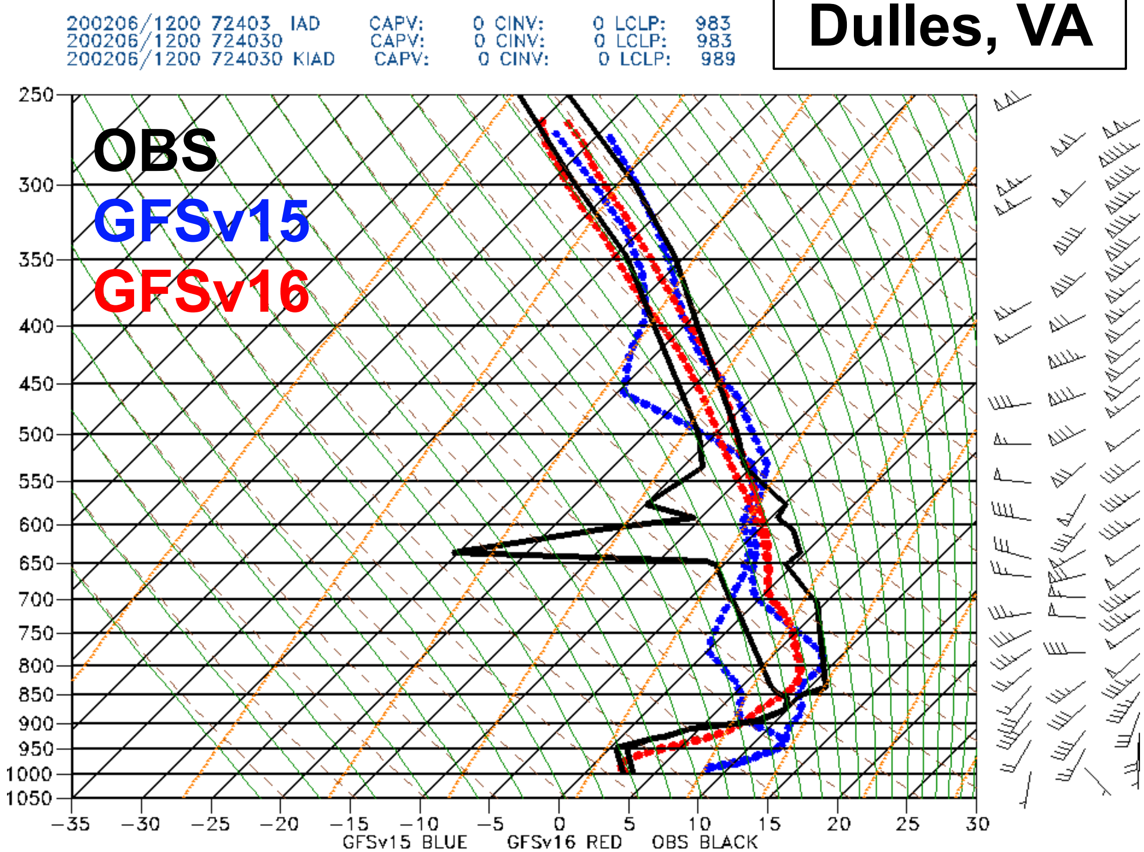
Mid-Atlantic Severe Case
Fcst: 12z 02/03/20 (F072)
Valid: 12Z 02/06/20

- GFSv16 was correctly colder than GFSv15 over VA/MD area, where cold air damming is occurring along the eastern Appalachians
- Improved 2-m T forecasts in shallow, cold air masses may be tied to a better handling of low-level clouds
- This is a long-standing GFS issue for which there seems to be some v16 improvement

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Temps in Shallow, Cold Air Masses

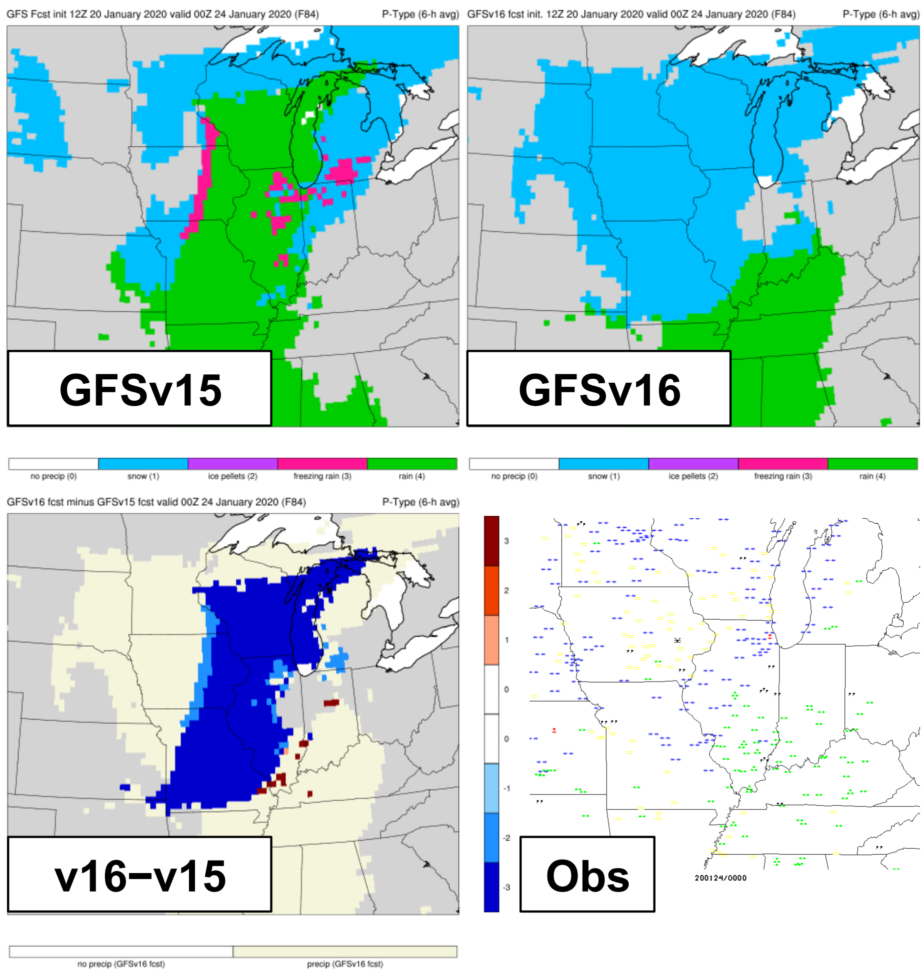
Dulles, VA



Mid-Atlantic Severe Case
Fcst: 12z 02/03/20 (F072)
Valid: 12Z 02/06/20

- **F072:** GFSv16 was correctly colder than GFSv15 over VA/MD area, where cold air damming is occurring along the eastern Appalachians
- Improved 2-m T forecasts in shallow, cold air masses may to be tied to a better handling of low-level clouds

Strengths: Resolved Low-level Warming Issue



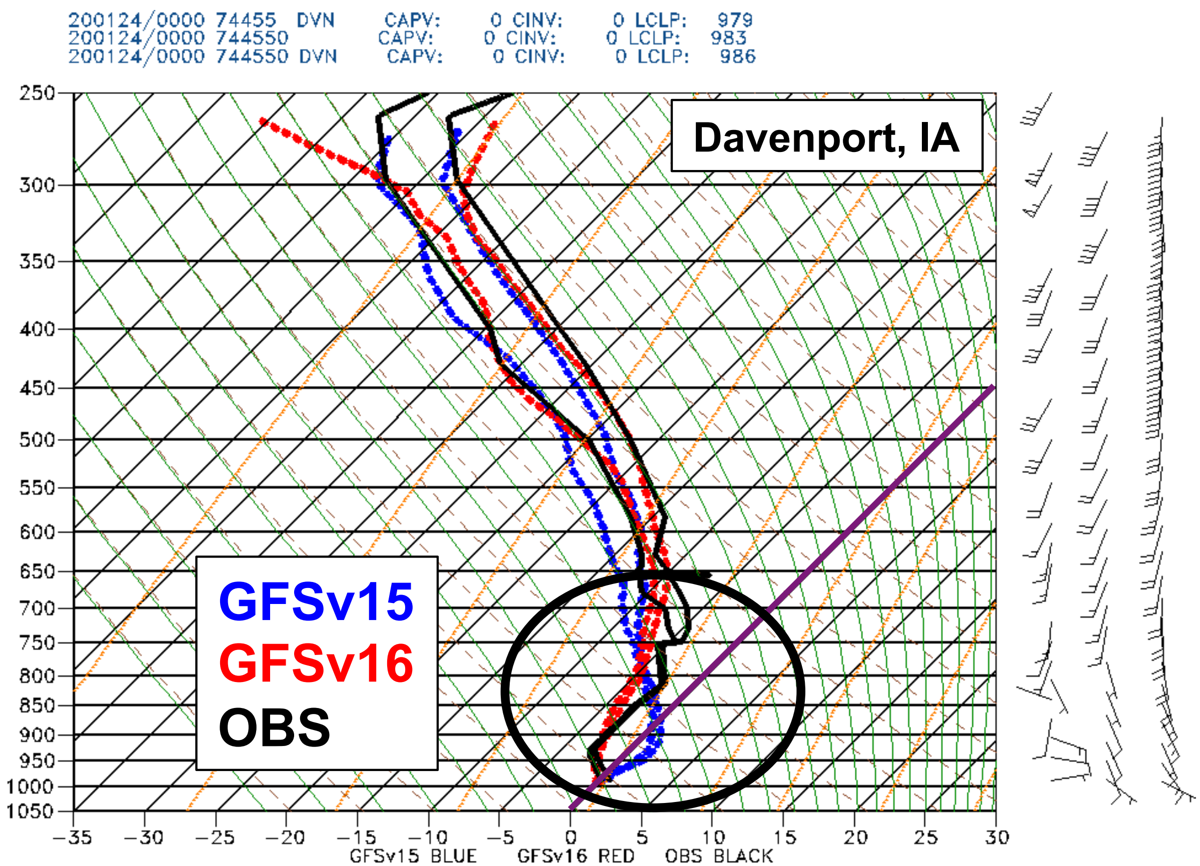
Midwest Ptype Event
Fcst: 12z 01/20/20 (F084)
Valid: 00Z 01/24/20

- An odd GFSv15 low-level warming issue that was seen a few cases last winter in GFSv15 appears to be resolved in GFSv16. In this example, GFSv15 forecasts rain over IA/IL/WI/MO where snow occurred; GFSv16 forecast is much improved

Thanks to Ray Wolf (WFO DVN)
See 2/6/20 MEG Presentation

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Resolved Low-level Warming Issue



Midwest Ptype Event
Fcst: 12z 01/20/20 (F084)
Valid: 00Z 01/24/20

- The improvement is shown in this example in which GFSv15 shows erroneous low-level warming that did not occur. GFSv16 has a correctly colder profile

Common Strengths From All Evaluations

- Notable improvements in synoptic-scale performance in the medium-range
  - Progressive bias in GFSv15 appears mitigated with better consistency catching correct solutions earlier
  - Improved frontal positions and QPF

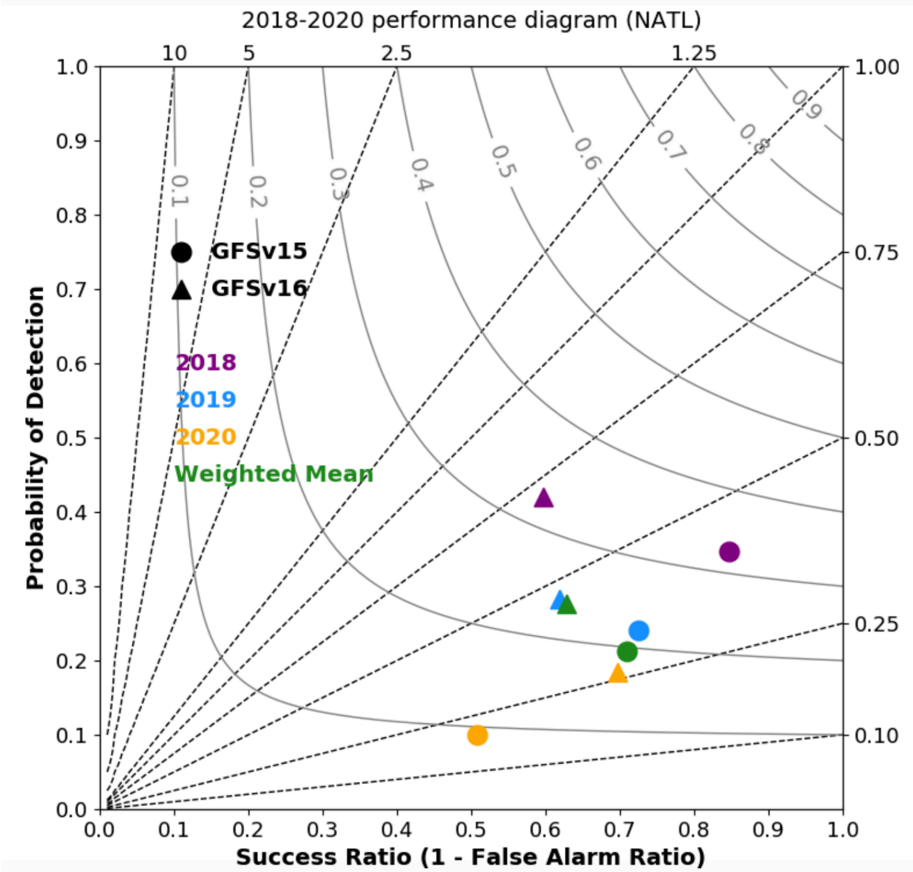
- Improvement in low-level temperature forecasts (mitigation of the winter low-level cold bias)

- Better ability to resolve shallow, cold air masses and some associated cold air damming events

- Improvements to TC intensity and increased lead time for genesis
  - With stronger TCs, GFSv16 has overall better track, size, and intensity

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Identifies TCs More Often & Earlier

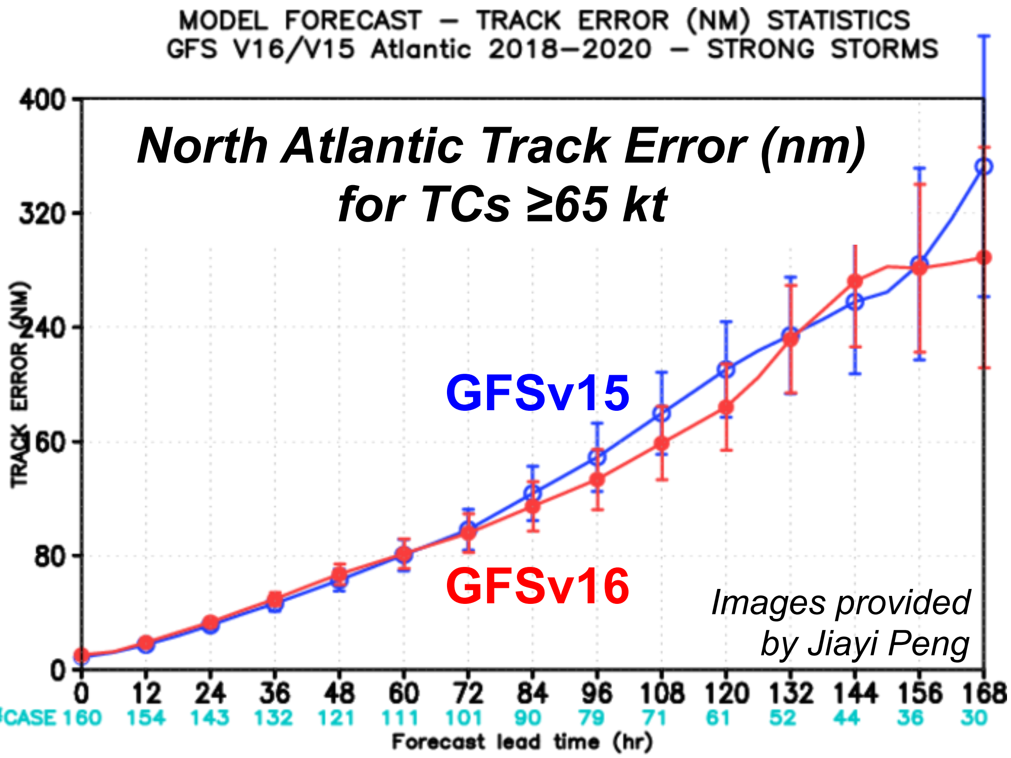


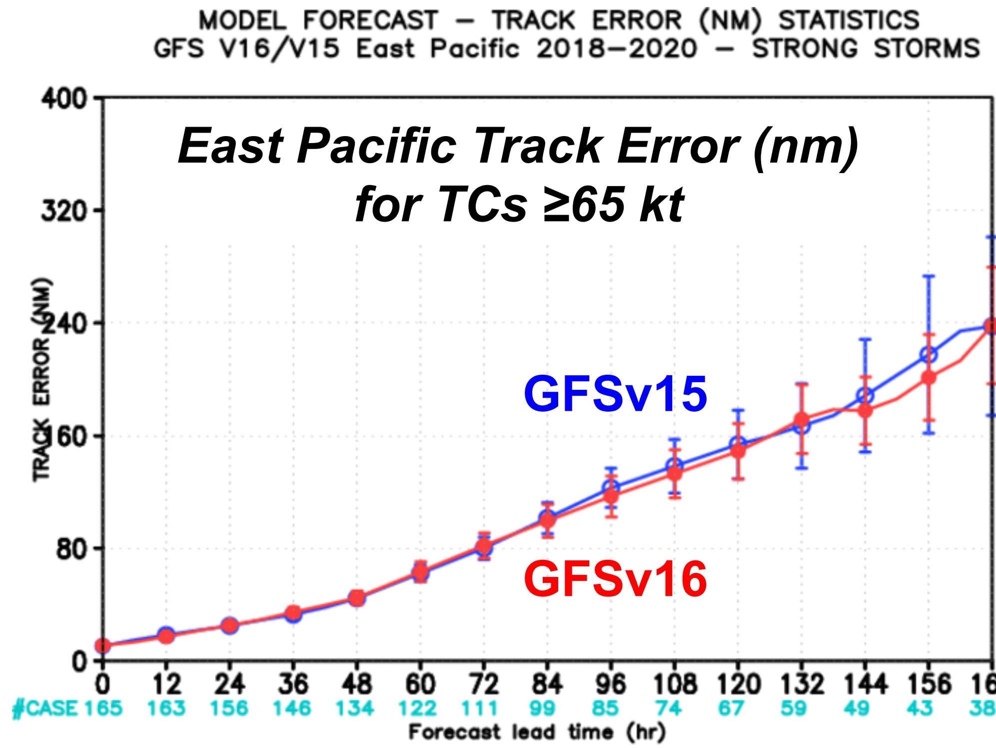
Thanks to Dan Halperin (ERAU)

- **Legend:**
  - x-axis: Success Ratio (1−FAR)
  - y-axis: Probability Of Detection (POD)
  - dashed lines: Frequency Bias
  - solid lines: Critical Success Index (CSI)
- All values would equal 1 in a perfectly performing model
- On average, GFSv16 exhibits:
  - Larger POD and CSI (closer to 1)
  - Frequency Bias is closer to 1
  - Smaller Success Ratio (FAR too high)
- GFSv16 is more cyclogenetic than GFSv15, and it identifies genesis with more lead time

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Improved Medium-Range Track Error

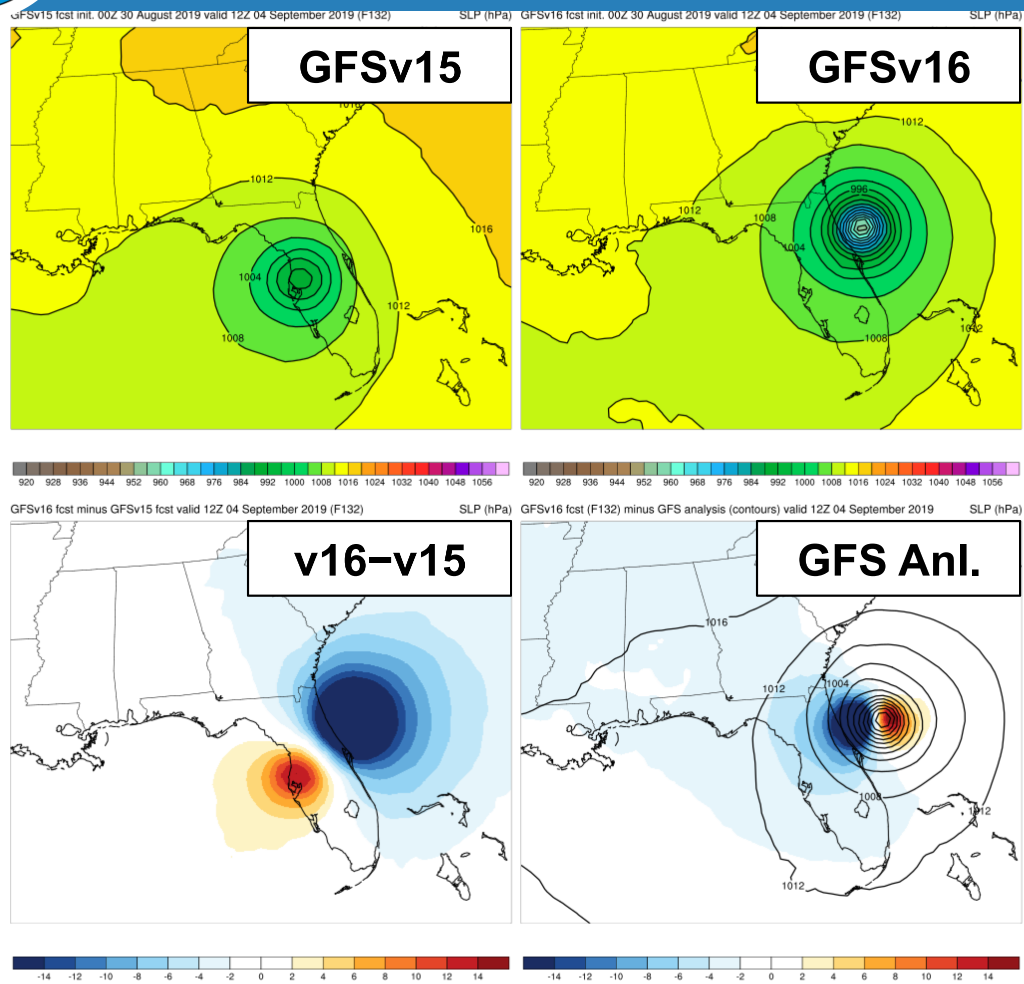




GFSv16 has lower track error than **GFSv15** for strong TCs (≥65 kt) during most of the medium range in both the North Atlantic and East Pacific

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Improved Medium-Range Track Error

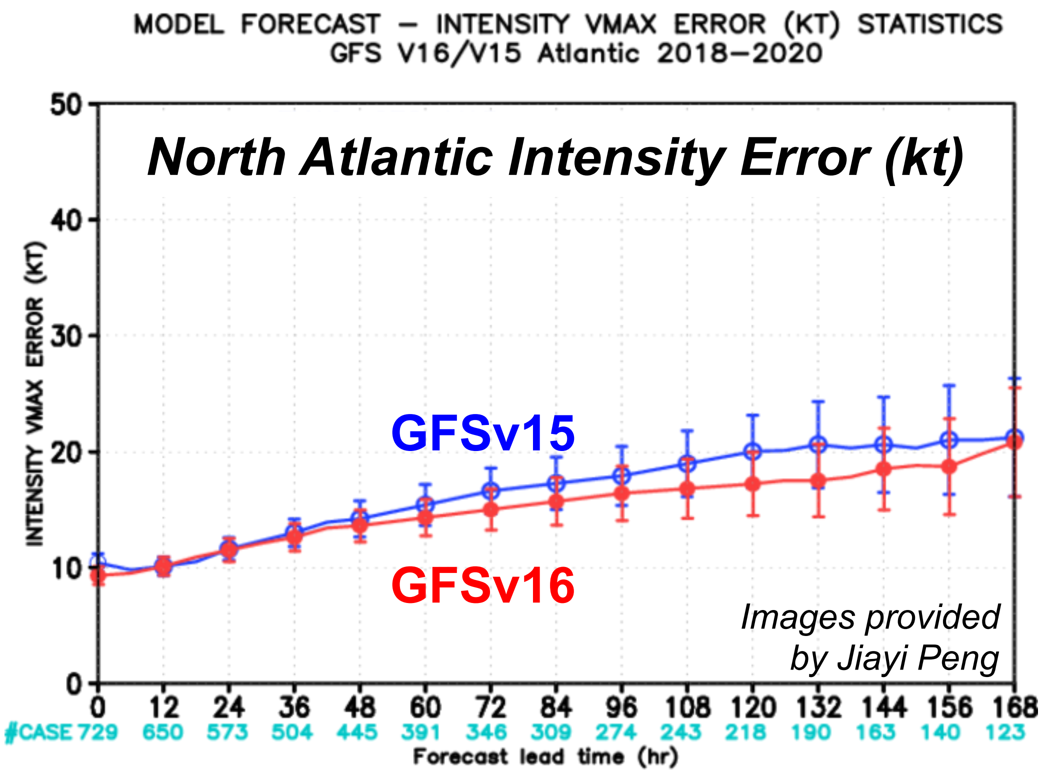


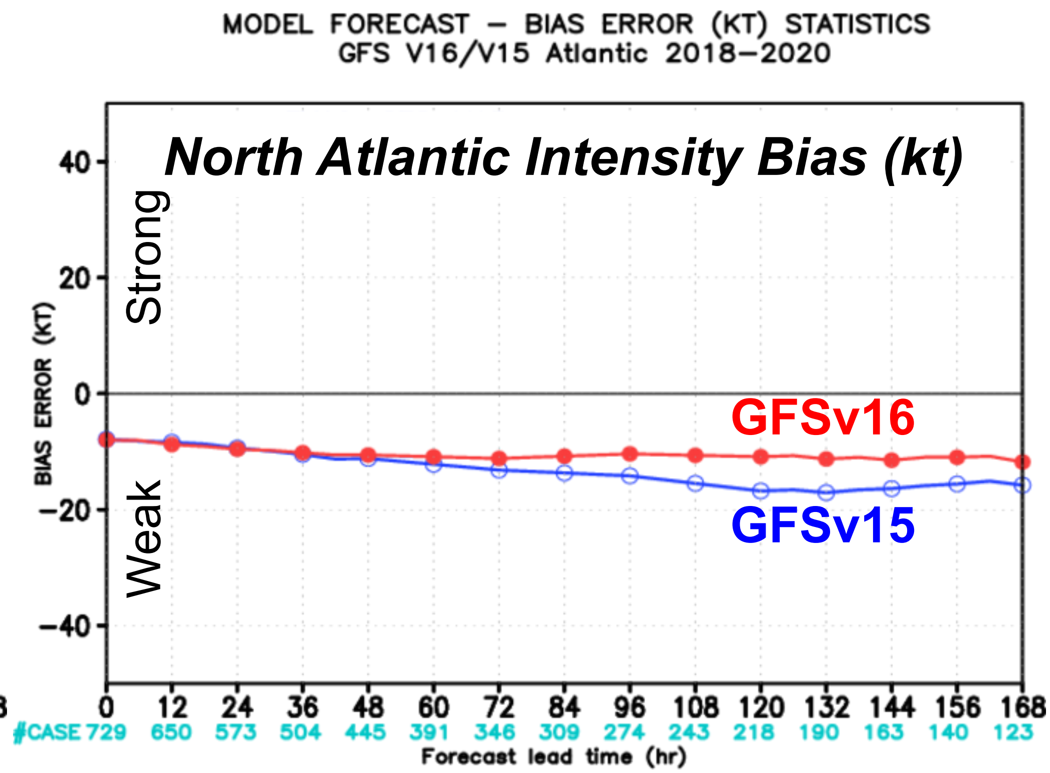
TC Dorian
Fcst: 00z 08/30/19 (F132)
Valid: 12Z 09/04/19

- GFSv16 forecasted Dorian to track north of Puerto Rico more than 24 h earlier than GFSv15 (not shown)
- Shown here, GFSv16 forecasted Dorian to turn right and skim the Florida coast 36 h earlier than GFSv15

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Improved TC Intensity in N Atlantic



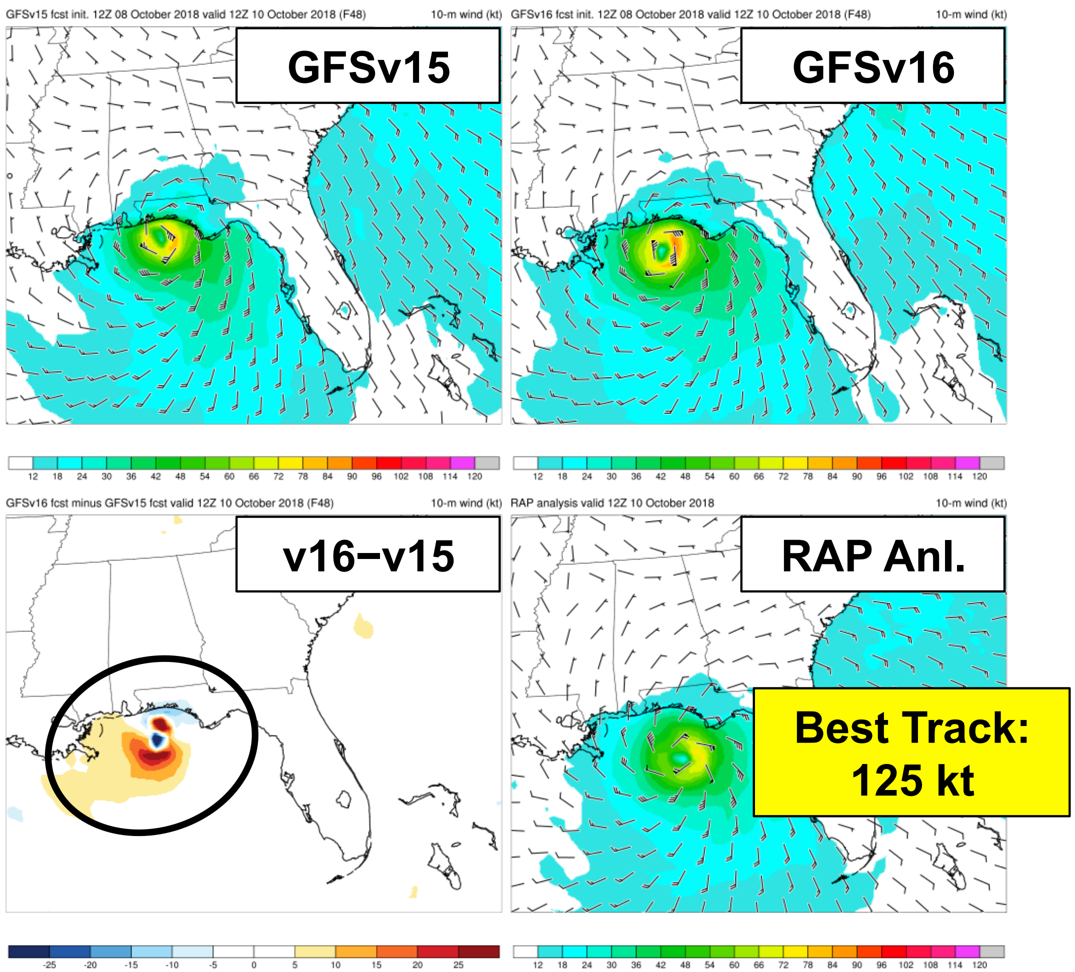


GFSv16 has lower intensity error than **GFSv15** at almost all lead times in the N Atlantic

GFSv16 has less of a weak bias than **GFSv15** at longer lead times

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Strengths: Improved TC Intensity in N Atlantic



TC Michael
Fcst: 12z 10/08/18 (F048)
Valid: 12Z 10/10/18

- **Michael:** GFSv16 consistently (and correctly) forecasted a stronger TC than GFSv15

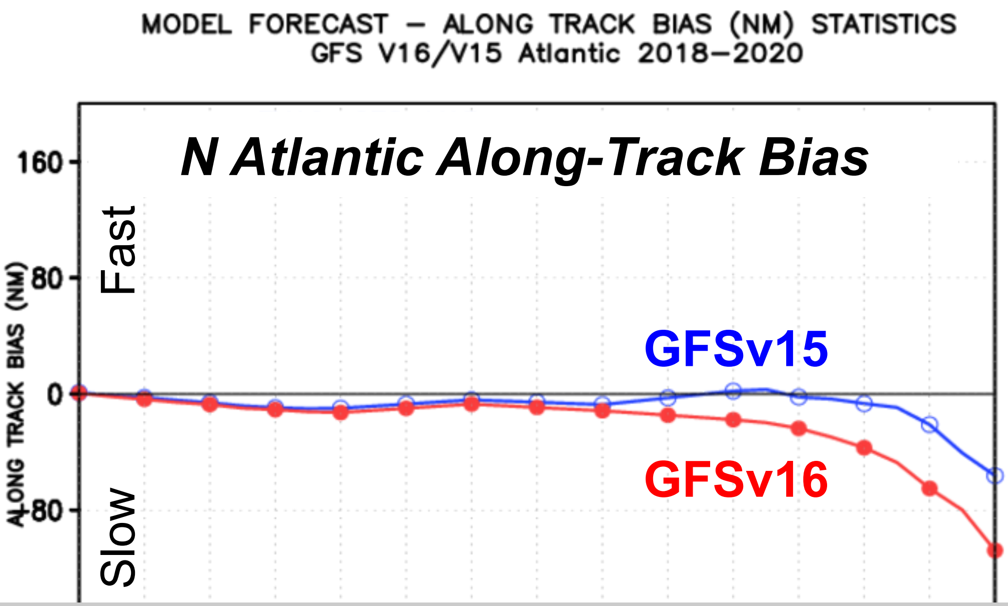
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

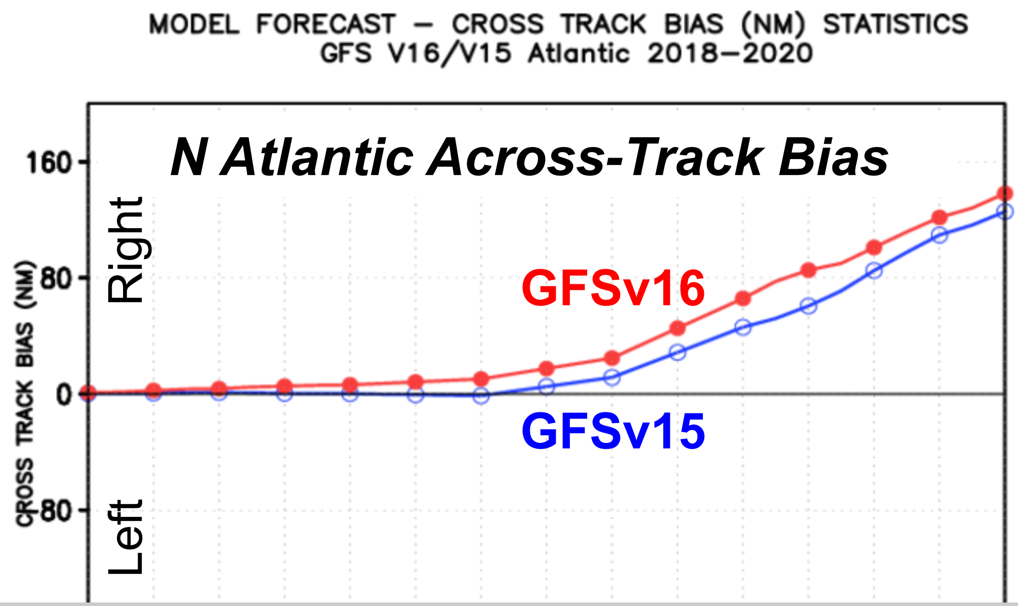
Common Concerns Across the Evaluations

- Increased right-of-track bias at longer lead times for North Atlantic TCs
- Larger TC False Alarm Rate (FAR) in the western North Atlantic (70°W–50°W)
- Tendency to strengthen all TCs in the long range (pre-formation, not in stats)
- Exacerbation of low instability (i.e., CAPE) bias that already existed in GFSv15, driven largely by dry soil moisture
- Lack of considerable improvement in forecasting radiation inversions

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Concerns: Increased Right-of-Track Bias



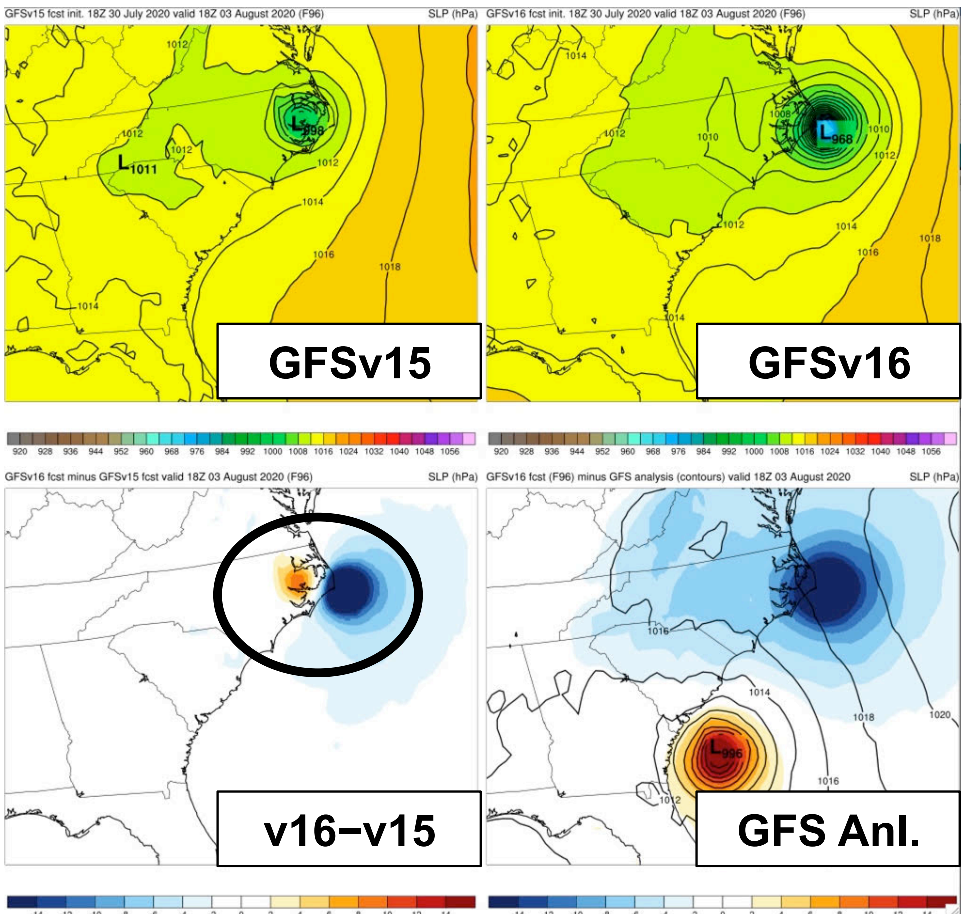


A slower and right-of-track bias at longer lead times suggests that GFSv16 may be recurving TCs earlier than GFSv15

GFSv16 has a larger slow bias than **GFSv15** that grows with forecast length in the N Atlantic

GFSv16 has a larger right-of-track bias than **GFSv15** that is largest at longer lead times

Concerns: Increased Right-of-Track Bias



TC Isaias
Fcst: 18z 07/30/20 (F096)
Valid: 18Z 08/03/20

- GFSv16 was often further right of track than v15 in the short and medium ranges. Both are also too fast in this example.

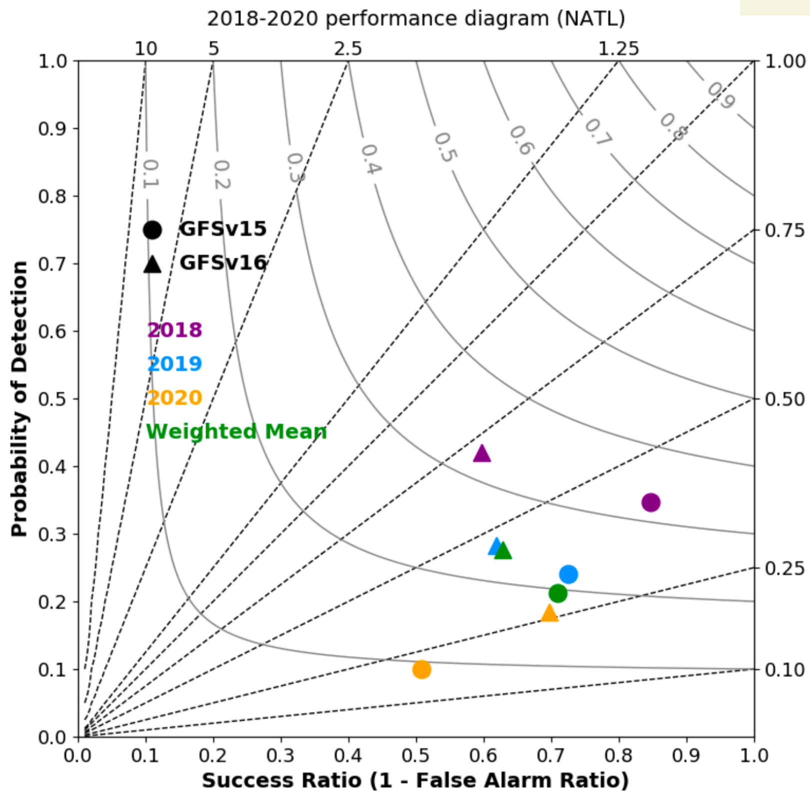
Common Concerns Across the Evaluations

- Increased right-of-track bias at longer lead times for North Atlantic TCs
- Larger TC False Alarm Rate (FAR) in the western North Atlantic (70°W–50°W)
- Tendency to strengthen all TCs in the long range (pre-formation, not in stats)
- Exacerbation of low instability (i.e., CAPE) bias that already existed in GFSv15, driven largely by dry soil moisture
- Lack of considerable improvement in forecasting radiation inversions

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Larger TC False Alarm Rate

From Dan Halperin, ERAU



- Forecast verification by year for each model configuration.
  - x-axis: success ratio
  - y-axis: probability of detection
  - Dashed lines: frequency bias
  - Curved lines: critical success index
- All values would equal 1 for a perfect performing model.
- Compared to GFSv15, GFSv16 exhibits on average:
  - Larger probability of detection
  - Smaller success ratio
  - Larger critical success index
- Overall, GFSv16 is more cyclogenetic than GFSv15.

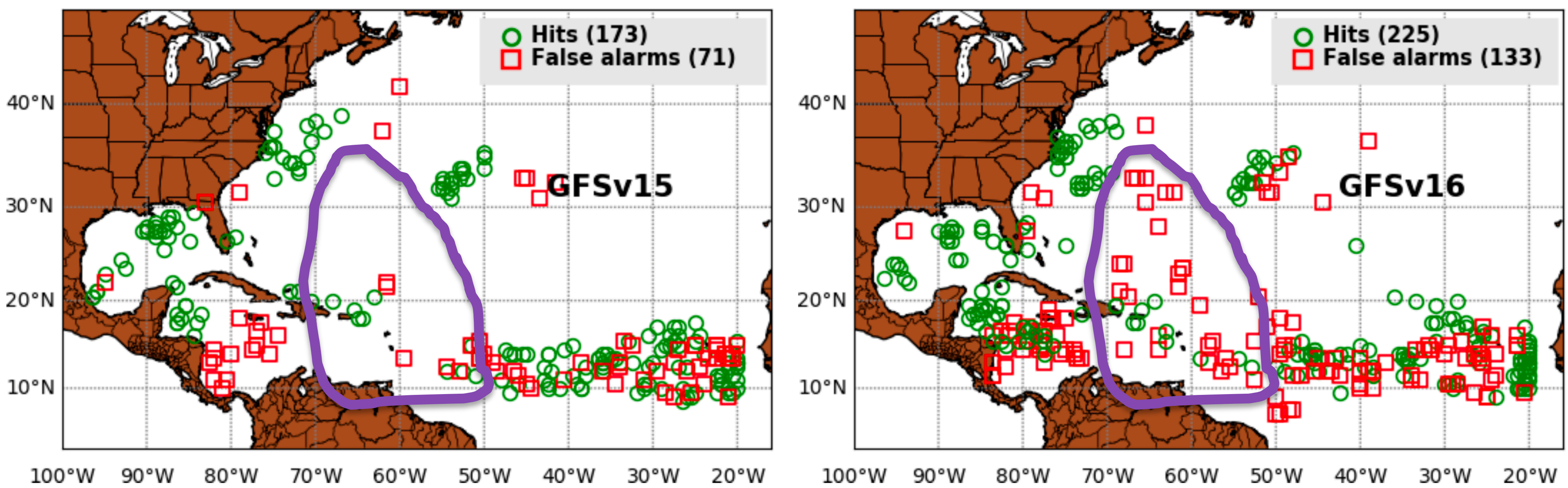
A ▲ further left than the ● of the same color indicates that v16 has a higher false alarm rate for that season

While preliminary 2020 numbers look good for v16, the weighted mean for the three TC seasons shows that v16 has a larger FAR

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Larger TC False Alarm Rate

From Dan Halperin, ERAU

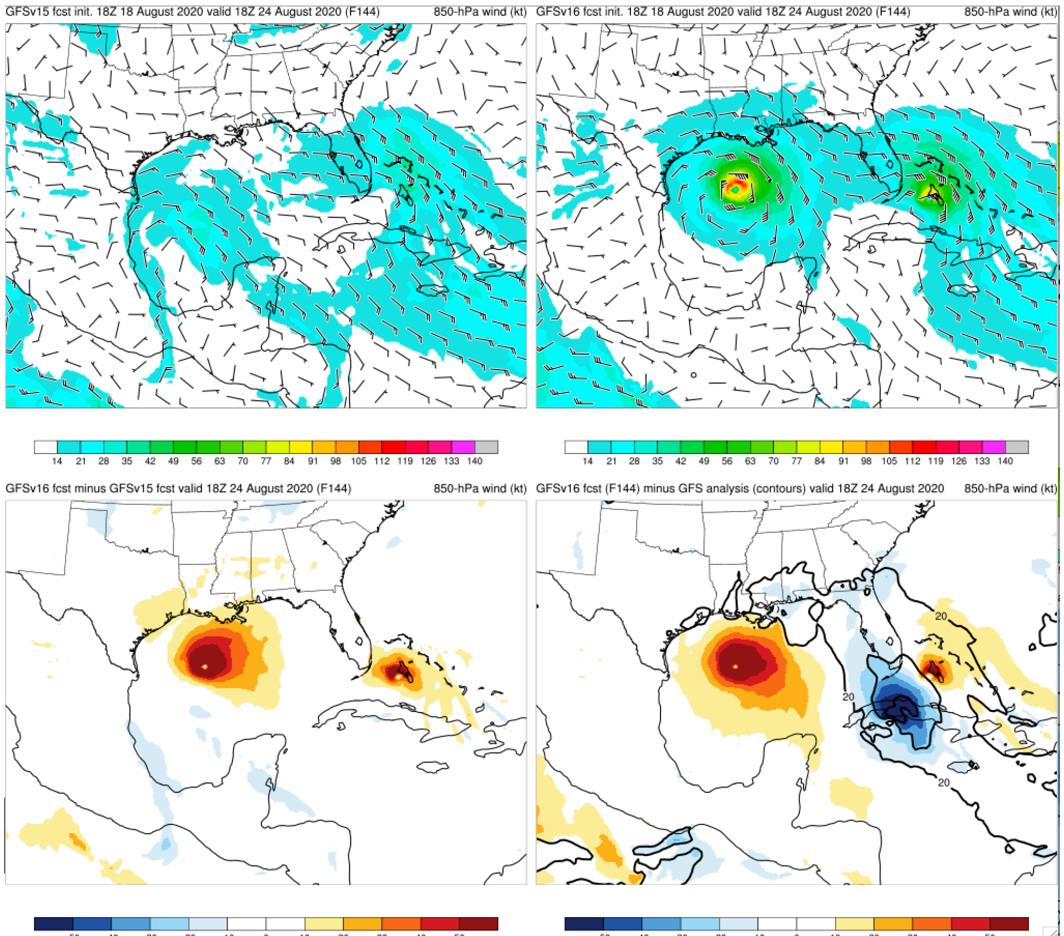


Large number of false alarms in GFSv16, relative to v15, between 50º and 70º W

Common Concerns Across the Evaluations

- Increased right-of-track bias at longer lead times for North Atlantic TCs
- Larger TC False Alarm Rate (FAR) in the western North Atlantic (70°W–50°W)
- Tendency to strengthen all TCs in the long range (pre-formation, not in stats)
- Exacerbation of low instability (i.e., CAPE) bias that already existed in GFSv15, driven largely by dry soil moisture
- Lack of considerable improvement in forecasting radiation inversions

Concerns: Strengthens Too Many TCs



TCs Laura/Marco
Fcst: 18z 08/18/20(F144)
Valid: 18Z 08/24/20

- **Marco**: GFSv16 had better track forecasts, but 15 consecutive v16 cycles had Marco as a sub 982 low (with many in the 950s and 960s); no GFSv15 cycle was that intense
- **Laura**: GFSv16 did well with many aspects of the intensity forecast, but this example shows a major threat to south FL that did not materialize

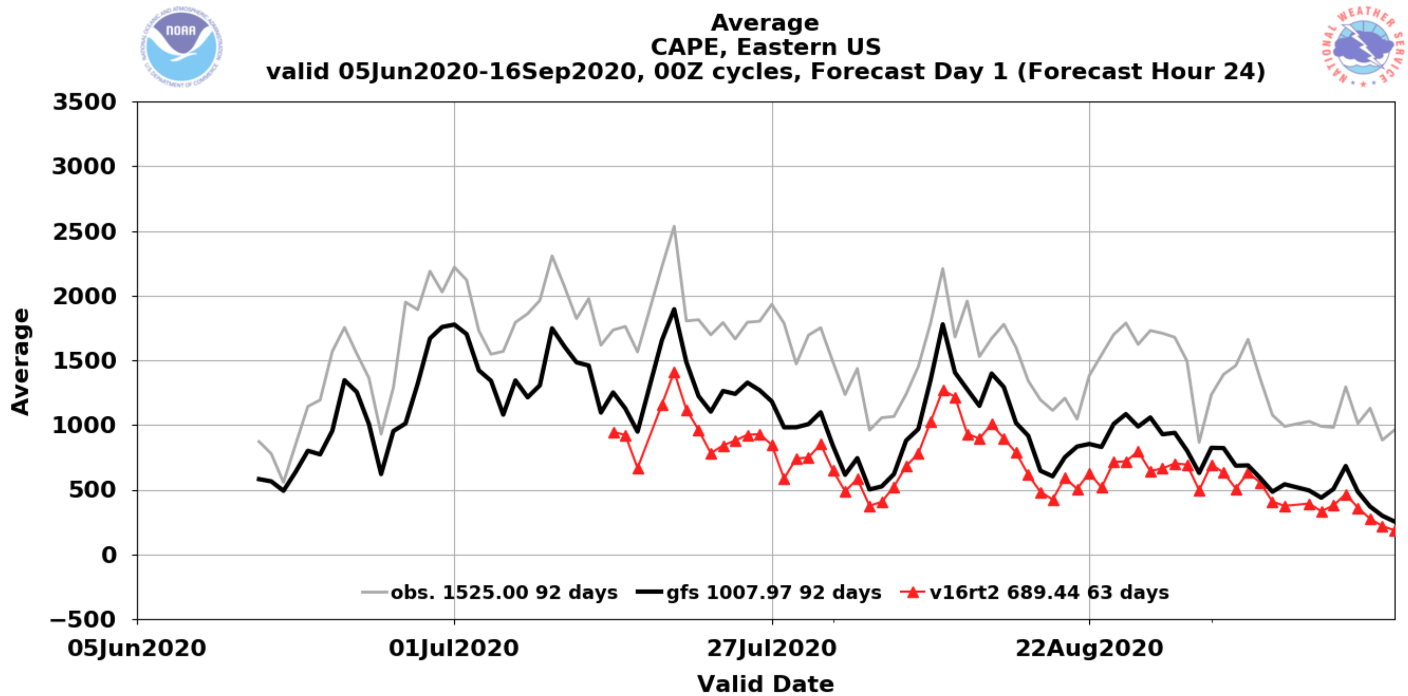
Common Concerns Across the Evaluations

- Increased right-of-track bias at longer lead times for North Atlantic TCs
- Larger TC False Alarm Rate (FAR) in the western North Atlantic (70°W–50°W)
- Tendency to strengthen all TCs in the long range (pre-formation, not in stats)
- Exacerbation of low instability (i.e., CAPE) bias that already existed in GFSv15, driven largely by dry soil moisture
- Lack of considerable improvement in forecasting radiation inversions

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

CAPE Magnitudes Are Reduced in GFSv16

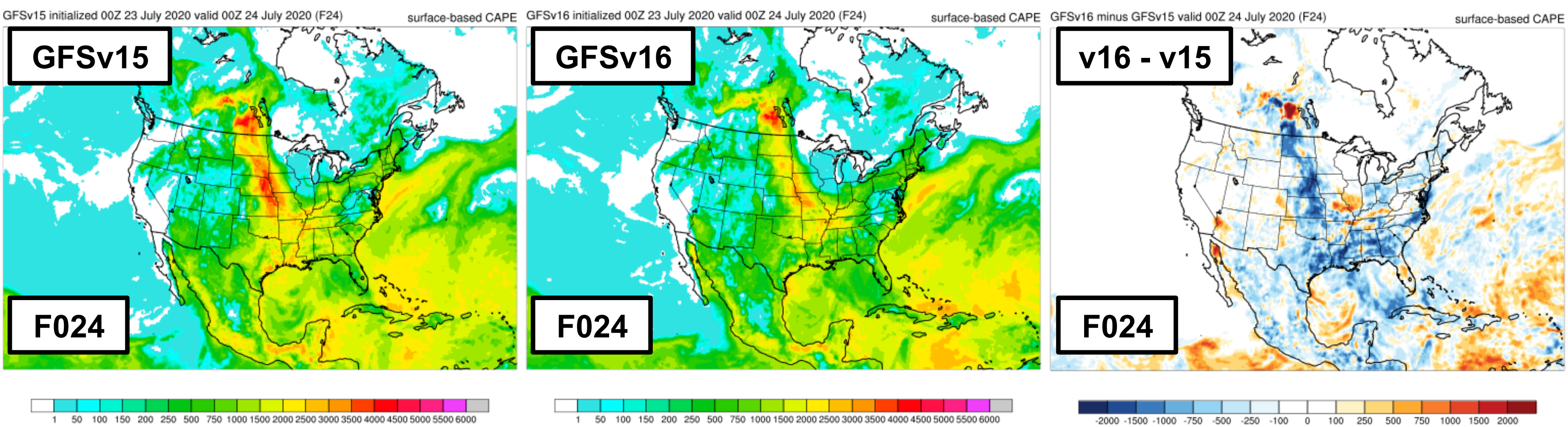
GFSv15
GFSv16
Obs



- Operational **GFSv15** CAPE analyses/forecasts are consistently lower than **obs**
- CAPE magnitudes in **GFSv16** analyses/forecasts are consistently lower than those from **GFSv15**

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

CAPE Magnitudes Are Reduced in GFSv16

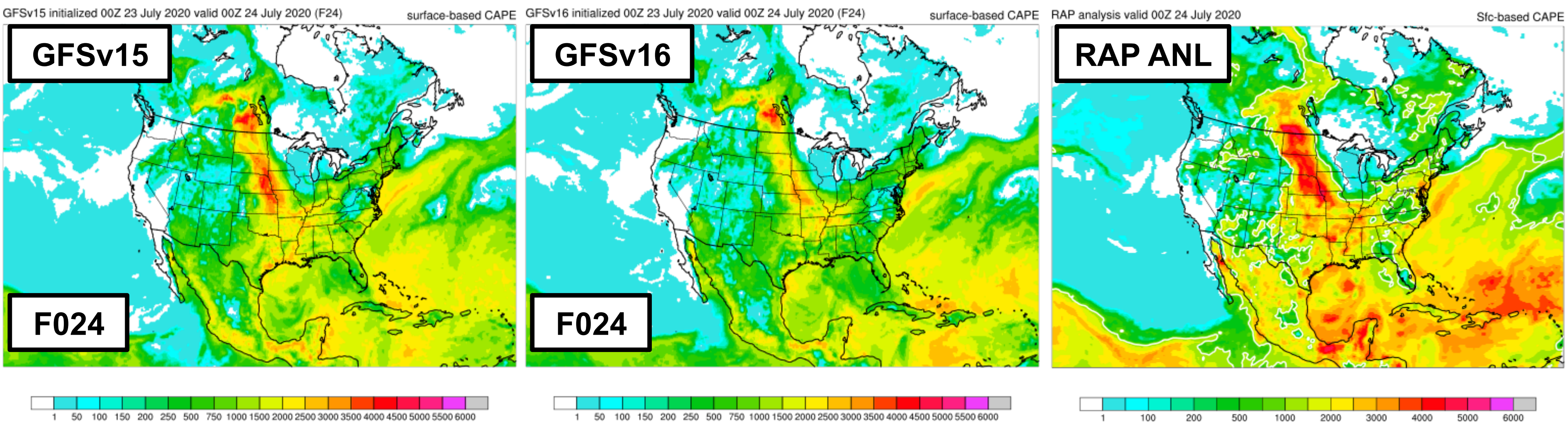


Surface-Based CAPE Forecasts (left and middle) and Forecast Differences (right)
Init: 00Z 07/23/20 Valid: 00Z 07/24/20 (F024)

- GFSv16 CAPE was notably lower across the Northern and Central Plains, as well as over the Gulf Coast region and southeast; smaller reductions over the northeast, Ohio Valley, and Mexico

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

CAPE Magnitudes Are Reduced in GFSv16



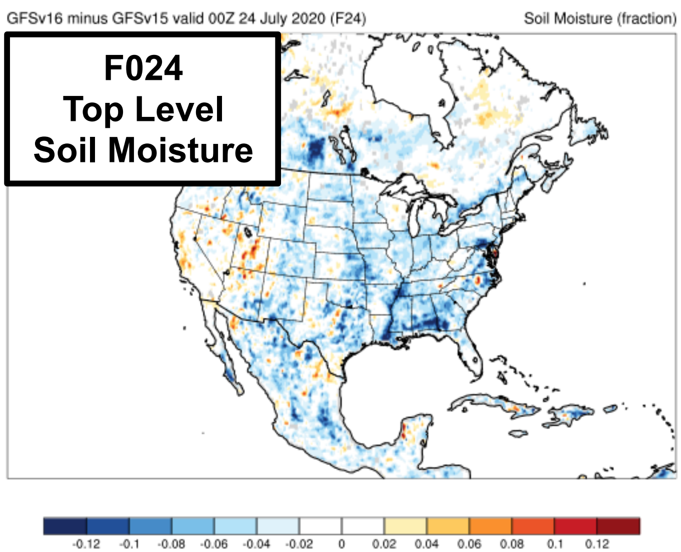
Surface-Based CAPE Forecasts (left and middle) and Analysis (right)
Init: 00Z 07/23/20 Valid: 00Z 07/24/20 (F024)

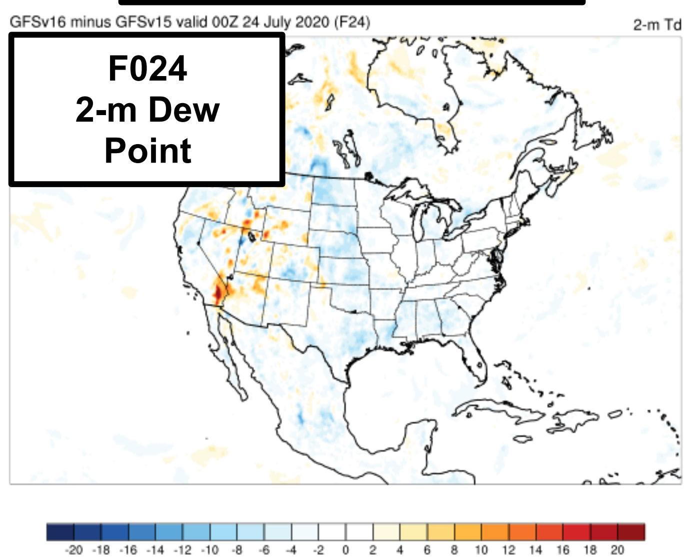
- The higher CAPE values in GFSv15 are almost always better, and even those are too low
- The only cumulative negative SOO team rating (across their complete set of ratings) was for short-range forecasts of CAPE

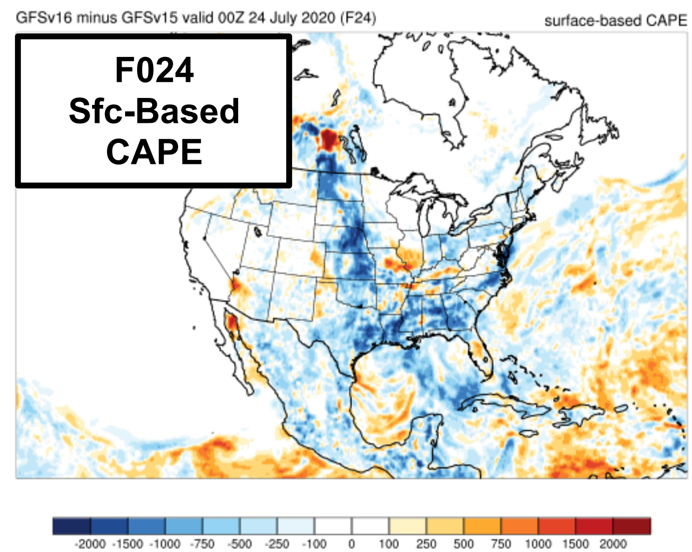
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Drier Soil Reduces Boundary Layer Moisture

v16 - v15 Differences







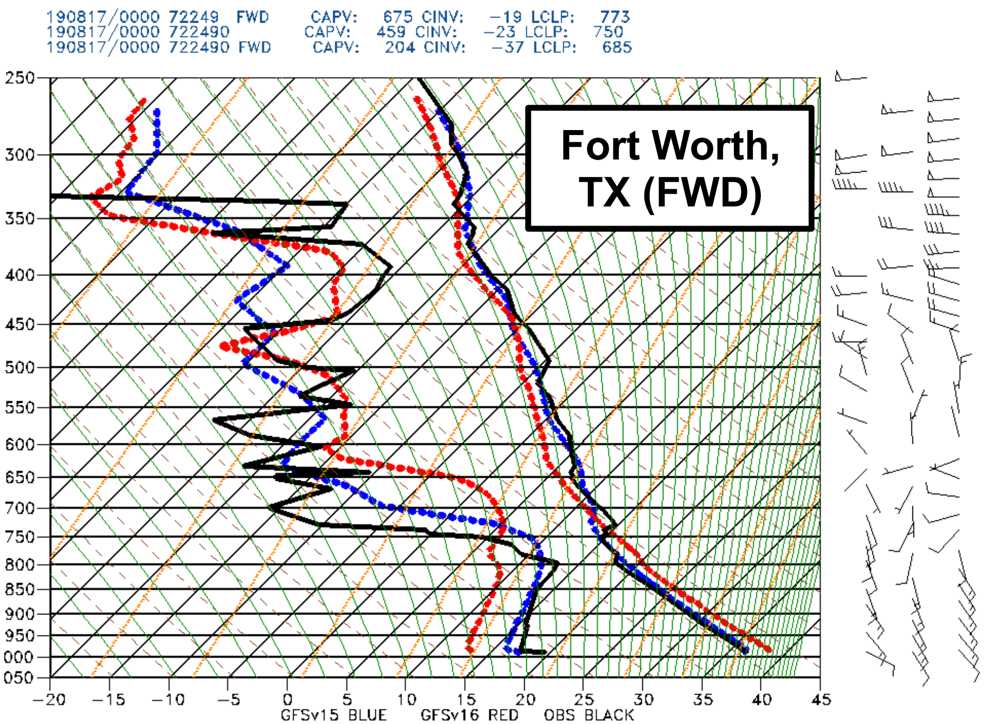
Init: 00Z 07/23/20 Valid: 00Z 07/23/20 (F024)

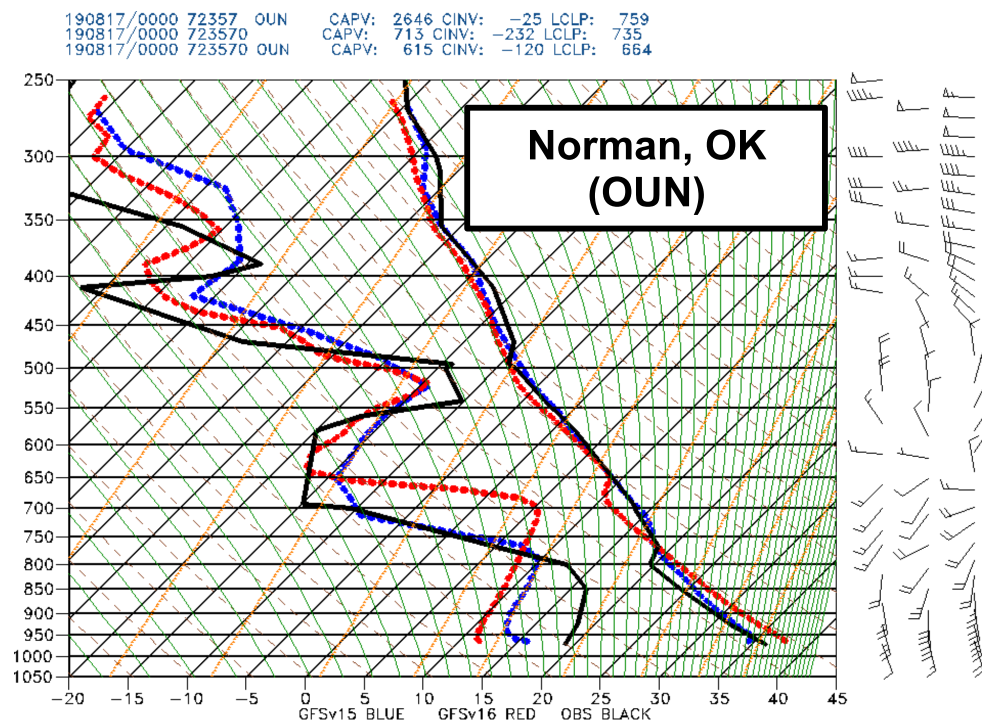
- GFSv16 top level soil moisture is considerably drier than in v15
- Good alignment between lower 2-m dew points and largest areas of reduced CAPE

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Tendency to Overmix the Boundary Layer

GFSv15 GFSv16 Obs





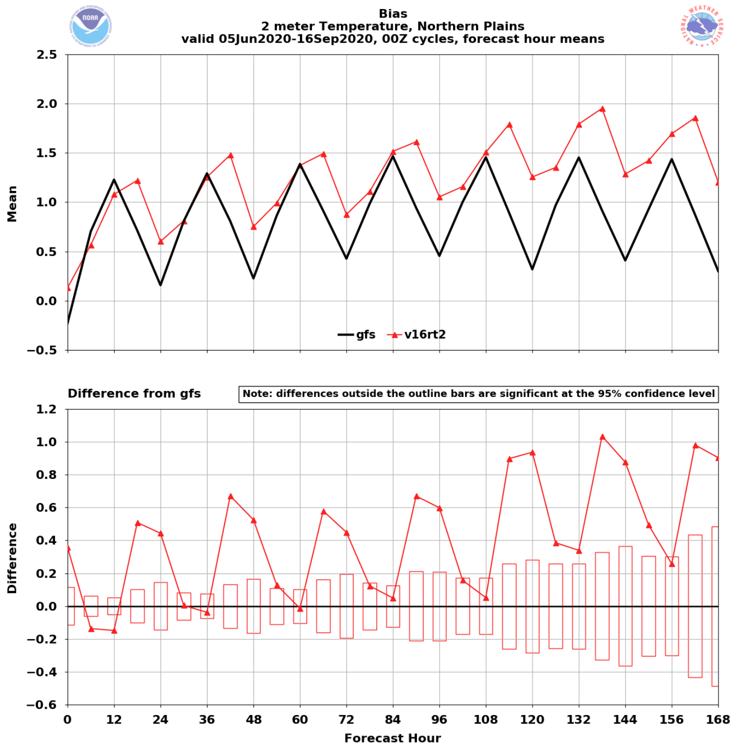
Init: 12Z 08/16/20 Valid: 00Z 08/17/20 (F012)

- **GFSv16** PBL was drier/warmer/deeper than **GFSv15** and **obs** in the unstable air

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

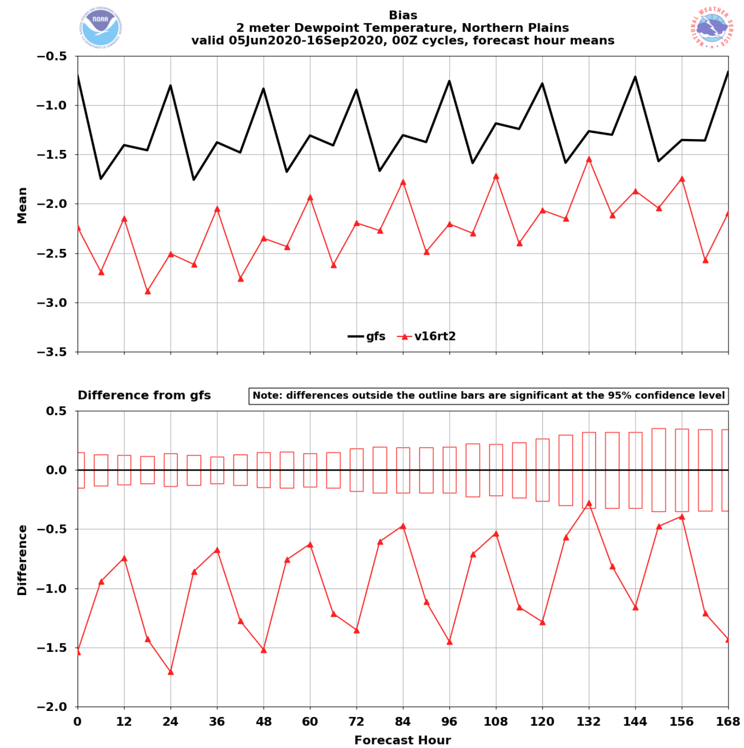
Warm/Dry Bias Exacerbated Across the Great Plains

2-m T



GFSv15

GFSv16



2-m Td

Northern Plains 2-m T (left) and 2-m Td (right) Bias as a Function of Forecast Lead

- Similar bias exists, but to a lesser extent, in neighboring regions

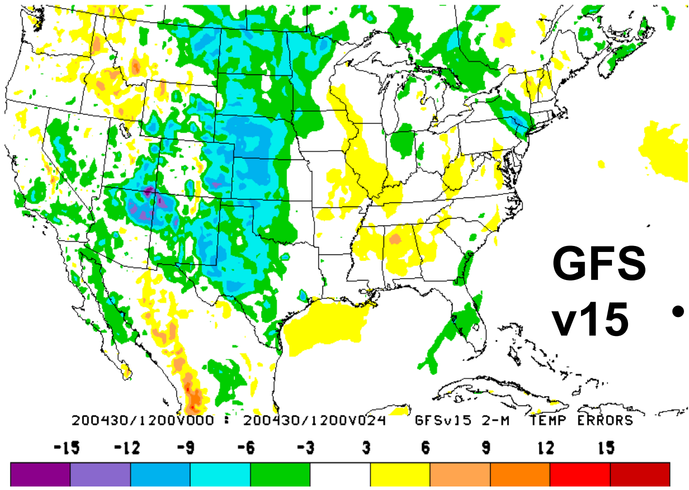
Common Concerns Across the Evaluations

- Increased right-of-track bias at longer lead times for North Atlantic TCs
- Larger TC False Alarm Rate (FAR) in the western North Atlantic (70°W–50°W)
- Tendency to strengthen all TCs in the long range (pre-formation, not in stats)
- Exacerbation of low instability (i.e., CAPE) bias that already existed in GFSv15, driven largely by dry soil moisture
- Lack of considerable improvement in forecasting radiation inversions

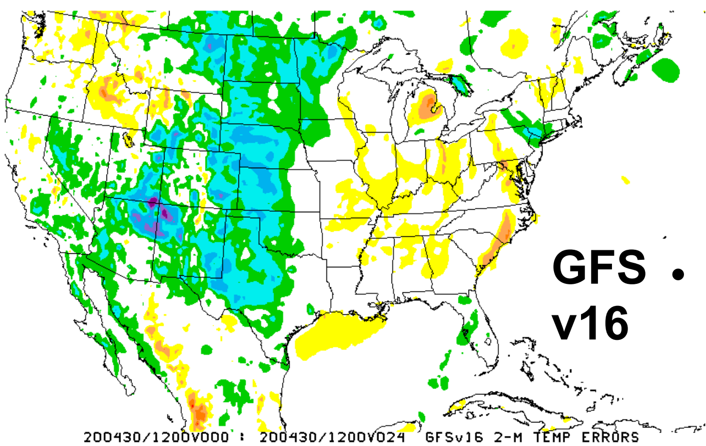
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Inversions

2m Temp Errors   Analysis - Forecast



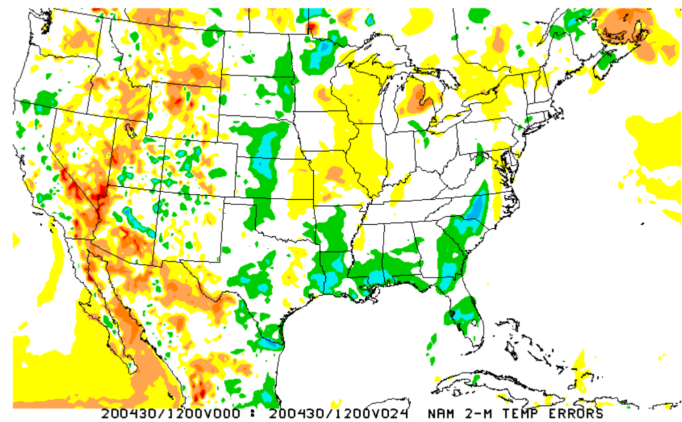
GFS v15



GFS v16

Fcst: 12Z 04/29/20 (F024)
Valid: 12Z 04/30/20

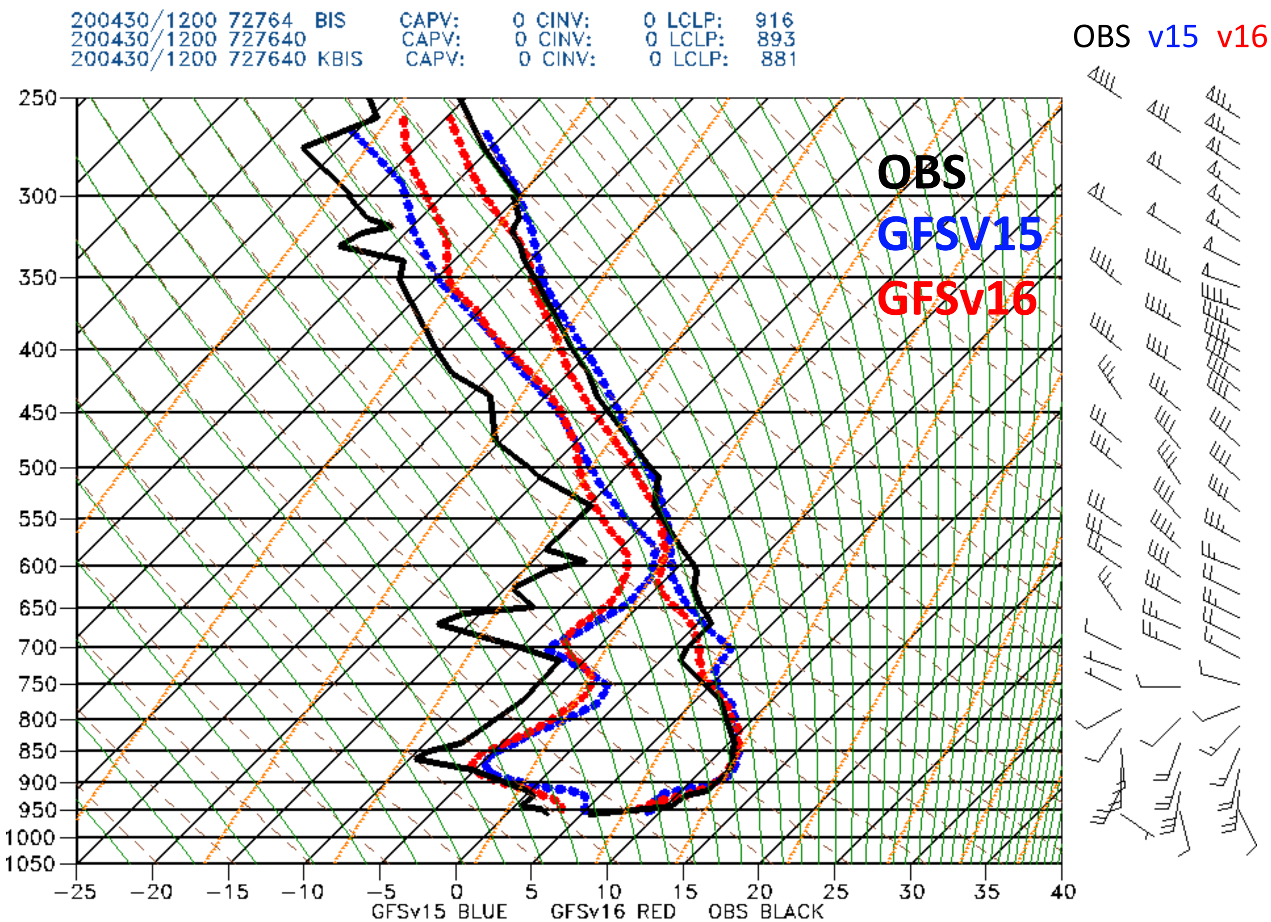
- Very large errors for a short range forecast. **GFSv16 appears to offer very slight improvement**, but both forecasts are **far too warm** over the Plains and upper Midwest
- The **lack of a sufficiently strong inversion** shows up well in the forecast soundings



NAM error map for same case shows that this type of case (strong radiational cooling) can be handled better

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Inversions - BIS soundings



Bismarck, ND BIS

Fcst: 12Z 04/29/20 (F024)
Valid: 12Z 04/30/20

- GFSv15 and v16 both fail to capture the strength of the low-level inversion and end up way too warm at the lowest levels
- GFSv16 shows very modest improvement over v15
- Note how the observed winds are weak at the lowest level; both GFS versions have winds that are too strong

Final Thoughts

- There were 12 recommendations submitted for the atmospheric component of the evaluation: 10 recommended implementation, and 2 were neutral
- The biggest overall positives are the improved medium range synoptic performance of GFSv16 and the mitigated low-level cold bias in the cool season
- The biggest negative is the reduction of warm season CAPE values that are already too low in GFSv15
- Tropical performance has a mix of improvements and degradation
- There were 3 recommendations submitted for the waves component of the evaluation: 1 recommended implementation, 1 was neutral, and 1 did not recommend implementation
- The biggest waves concerns are along the U.S. West Coast and over the North Atlantic, where users like the higher-resolution Multi-1 output grids