Evaluating Model Physics in the Unified Forecast System (UFS) Medium-Range Weather Application

DTC Global T&E Team:

Weiwei Li[1], Ligia Bernardet[2], Michelle Harrold[1], Dan D'Amico[1], Xia Sun[2,3], Lulin Xue[1], Judy Henderson[2], Jimy Dudhia[1], Grant Firl[1], Brianne Nelson[1], and Michael Ek[1]

1. *National Center for Atmospheric Research (NCAR), and Developmental Testbed Center (DTC)*
2. *NOAA/Global Systems Laboratory (GSL), and Developmental Testbed Center (DTC)*
3. *University of Colorado/Cooperative Institute for Research in Environmental Sciences (CIRES), NOAA/Global Systems Laboratory (GSL), and Developmental Testbed Center (DTC)*

Outline

Part I. Evaluating Model Physics in the UFS Medium-Range Weather (MRW) Model - global runs

Part II. An Exercise of Using Common Community Physics Package (CCPP) Single-Column Model (SCM) under the Hierarchical Testing Framework *by Dan D'Amico et al. (tomorrow @ 11:45 AM EDT)*

DTC's global runs

- Free forecasts (cold start) were initialized from the GFS analysis
- Model configurations:
  - C768L64 (~13 km/64 Layers) using public release v1.0 of UFS Weather Model with GFSv16beta physics suite in CCPP
  - 32 runs, 5-day forecast: Oct 17, 2019 - Mar 15, 2020, initialized every 5 days, alternating 00 and 12 UTC initializations + a LASSO case (on 06/11/2016) to compare with the single-column model (SCM) test
- T&E: an array of analyses including Model Evaluation Tools (MET)-based verification and in-house diagnostics
- Benchmarks datasets:
  - Operational GFS forecasts
  - Observations (precip, cloud & radiation, soundings, etc.)

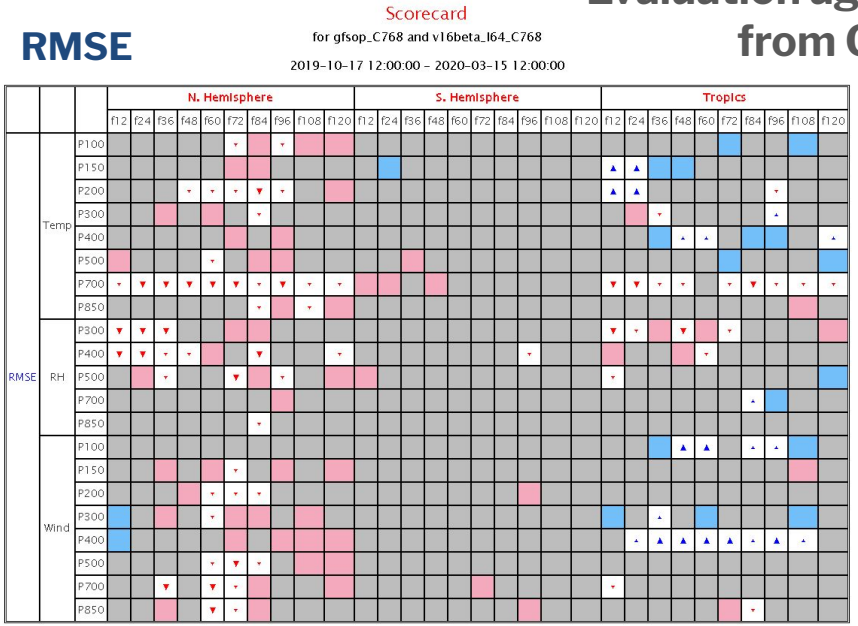
Evaluation items for this T&E activity

1. Performance-oriented evaluation (scorecards, ACC, tropical cyclone verifications, and other skill scores)
2. Systematic errors and biases in vertical profiles of temperature, relatively humidity, and wind speed
3. Boundary layer parameters
   a. Global regions
   b. Contiguous United States (CONUS)
   c. Example of skew-T log-P diagrams (CONUS)
4. Radiation and cloud processes
   a. Seasonal and zonal means
   c. Cloud properties
   d. Surface-based diagnostics (Forecasts vs SURFRAD)
5. Precipitation and moist processes

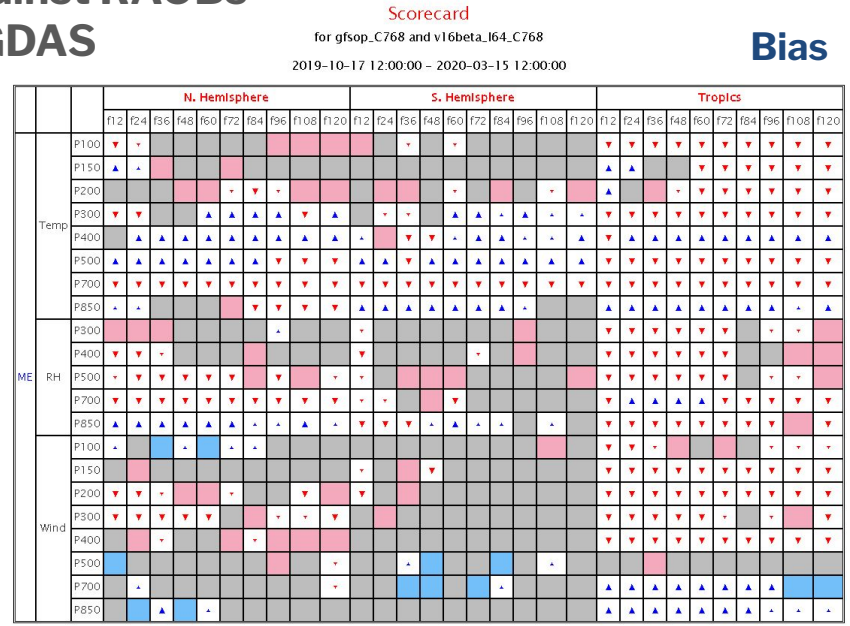
MET-generated scorecard examples (RMSE and Bias - Pressure-level Large-scale Environmental Variables)

Evaluation against RAOBs from GDAS

RMSE



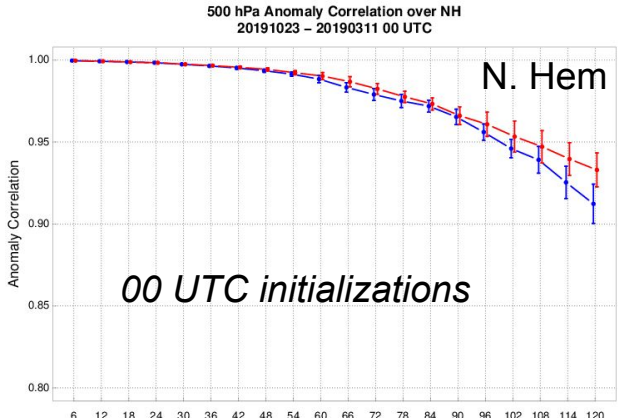
Bias

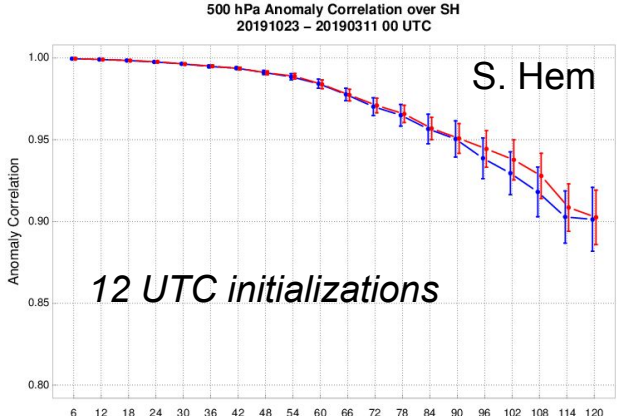


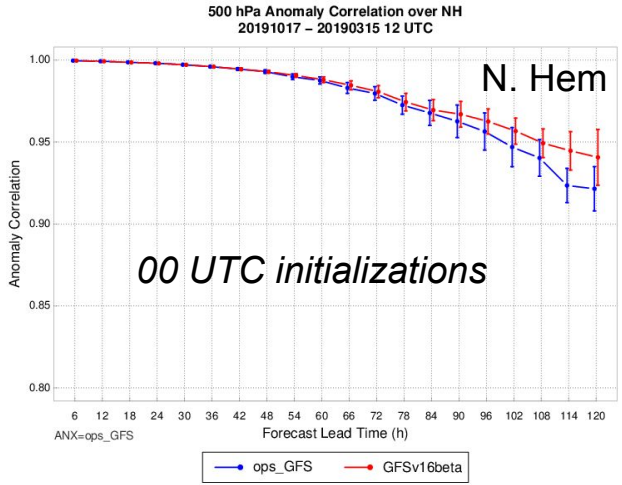
▲ ops_GFS is better than GFSv16beta at the 99.9% significance level
▵ ops_GFS is better than GFSv16beta at the 99% significance level
ops_GFS is better than GFSv16beta at the 95% significance level
No statistically significant difference between ops_GFS and GFSv16beta
ops_GFS is worse than GFSv16beta at the 95% significance level
▿ ops_GFS is worse than GFSv16beta at the 99% significance level
▼ ops_GFS is worse than GFSv16beta at the 99.9% significance level
Not statistically relevant
Statistic for symbols: DIFF_SIG

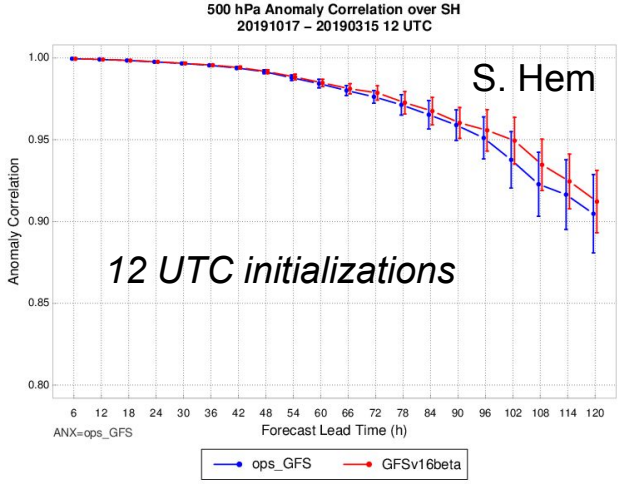
Pink denotes that GFSv16beta is **better** than ops_GFS

500 hPa Anomaly Correlation (ACC)









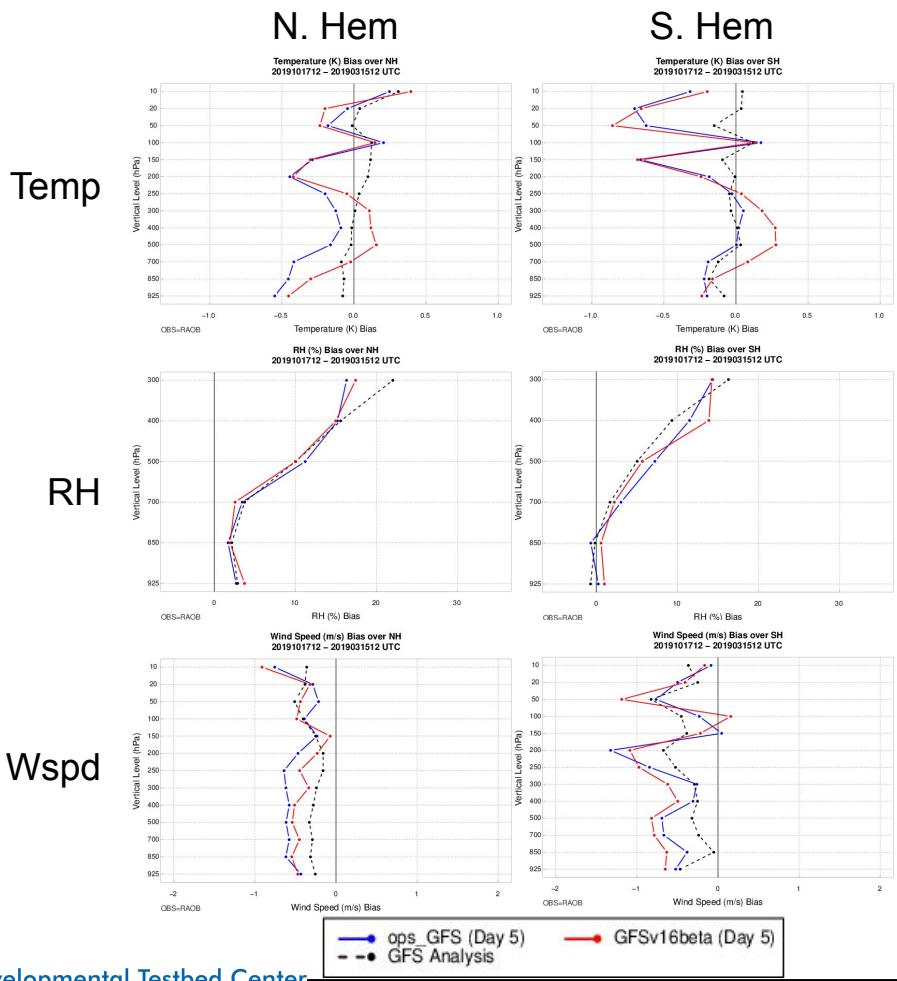
Summary: RMSE of GFSv16beta (Days 1-5) vs ops_GFS

Skill (RMSE)	200-hPa Winds	850-hPa Temp	850-hPa RH	850-hPa Winds	10-m Winds	2-m Temp	2-m Dew point Temp	500-hPa GPH (ACC)
N. Hem	Improved	Improved	Somewhat Improved	Improved	N/A	N/A	N/A	Improved
S. Hem	Neutral	Neutral	Neutral	Neutral	N/A	N/A	N/A	Improved
Tropics	Neutral	Neutral	Neutral	Improved	N/A	N/A	N/A	
CONUS	Improved	Improved	Somewhat Improved	Somewhat Improved	Improved	Neutral	Somewhat Improved	
E. CONUS	Somewhat Improved	Improved	Neutral	Somewhat Improved	Improved	Neutral	Neutral	
W. CONUS	Improved	Improved	Somewhat Improved	Somewhat Improved	Improved	Improved	Improved	

Summary: Bias of GFSv16beta (Days 1-5) vs ops_GFS

Skill (Bias)	200-hPa Winds	850-hPa Temp	850-hPa RH	850-hPa Winds	10-m Winds	2-m Temp	2-m Dew point Temp
N. Hem	Improved	Improved	Degraded	Degraded	N/A	N/A	N/A
S. Hem	Improved	Degraded	Somewhat Degraded	Neutral	N/A	N/A	N/A
Tropics	Improved	Degraded	Improved	Degraded	N/A	N/A	N/A
CONUS	Somewhat Improved	Neutral	Degraded	Neutral	Improved	Neutral	Somewhat Improved
E. CONUS	Somewhat Improved	Neutral	Degraded	Neutral	Improved	Somewhat Degraded	Improved
W. CONUS	Somewhat Improved	Improved	Degraded	Degraded	Improved	Improved	Improved

Vertical profiles: GFS vs Obs (GDAS RAOB)



Black - GFS anal
Blue - ops_GFS (Day5)
Red - GFSv16beta (Day5)

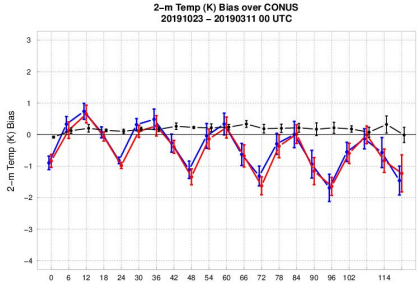
- **GFS anal**: systematic errors of being cold (lower levels or ABL), moist, and less windy
- **GFS fcst** (v16beta & ops_GFS): cold (ABL), moist and negative wind bias
- **V16beta vs ops_GFS:** milder cold bias (even a warm bias above ABL) but more humid, better w.r.t weaker winds
- **CONUS** shared similar systematic errors and biases (not shown) but **will focus on CONUS next**

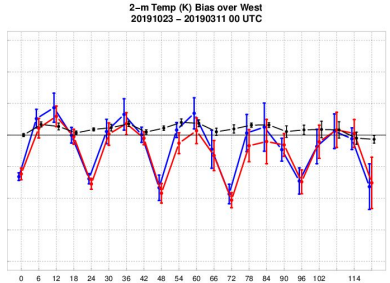
00Z cycle

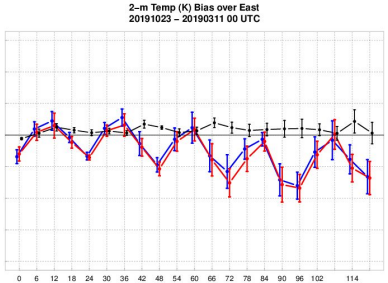
CONUS	W. CONUS	E. CONUS

Evaluate against NDAS METAR
Blue - ops_GFS
Red - v16beta
Black - ops_GFS minus v16beta

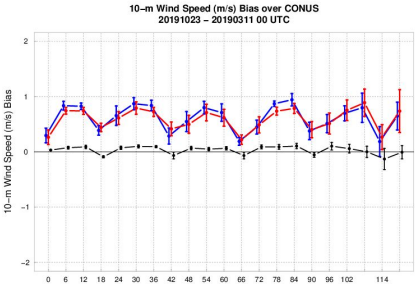
T2m

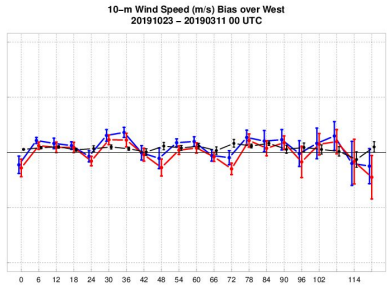


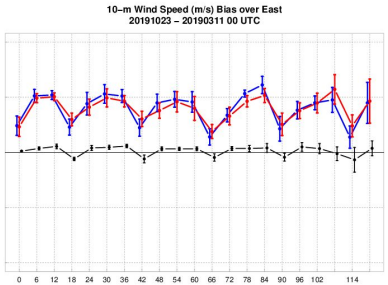




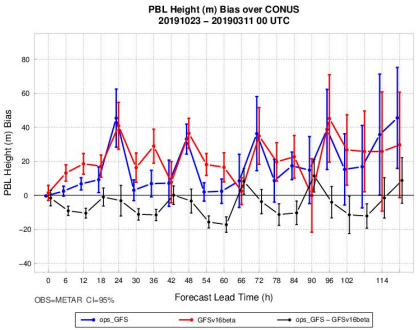
Wspd10m

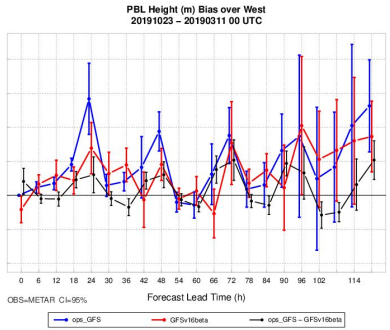


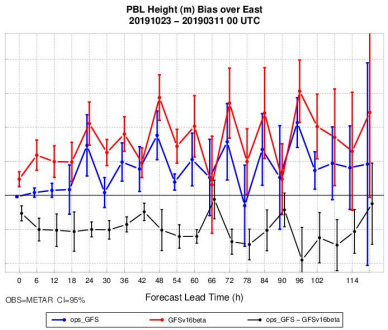




PBLH





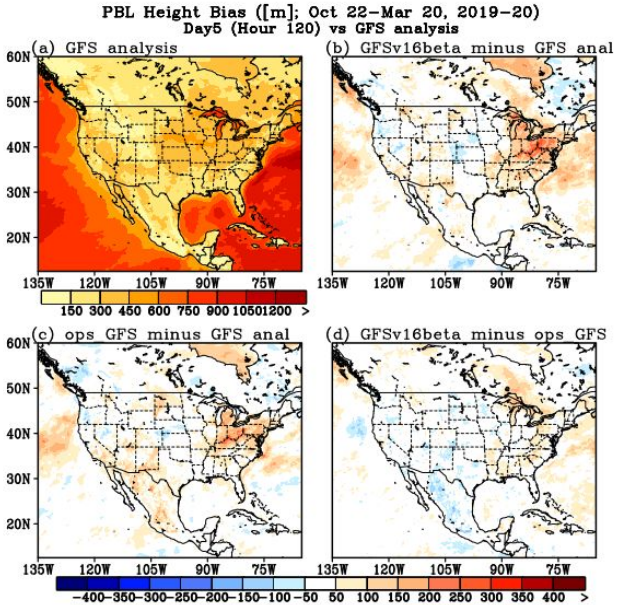


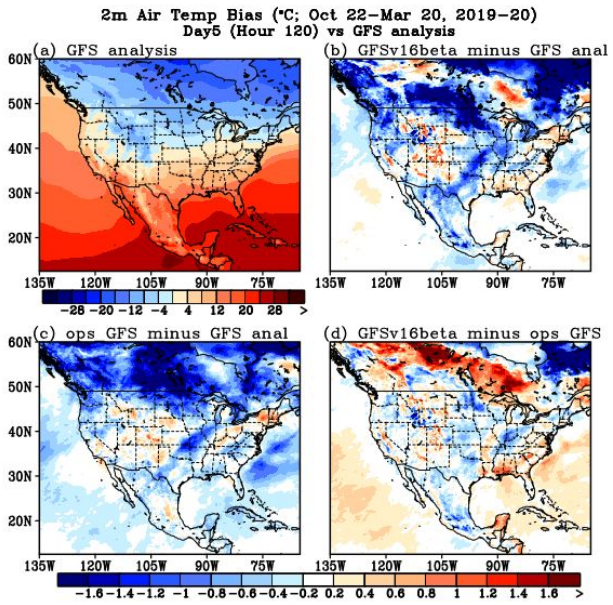
- Overall cold bias with diurnal variations
- Problematic Eastern CONUS: 1) cold bias increases with lead times; 2) overly stronger near-sfc winds; 3) deeper PBL; 4) larger nighttime Temp and wind bias
- Less problematic W. CONUS.
- v16beta: slightly colder, less windy, and PBL might be deeper than ops_GFS in E CONUS

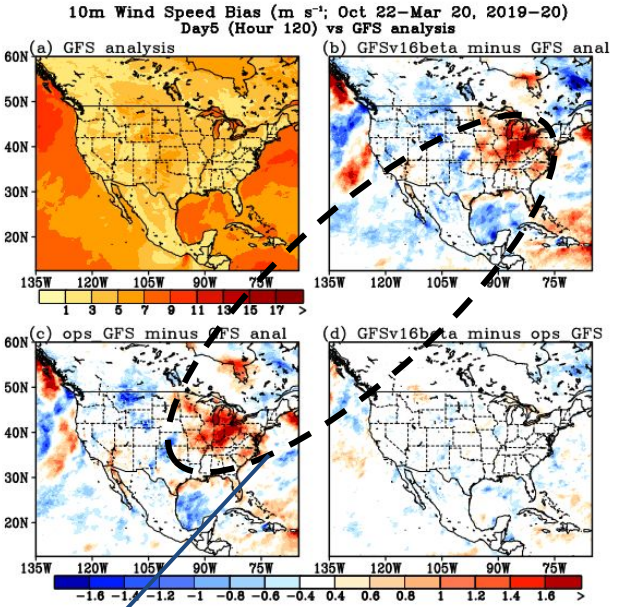
Deeper PBL in N.E. CONUS

Cold bias

Windy Great Lakes



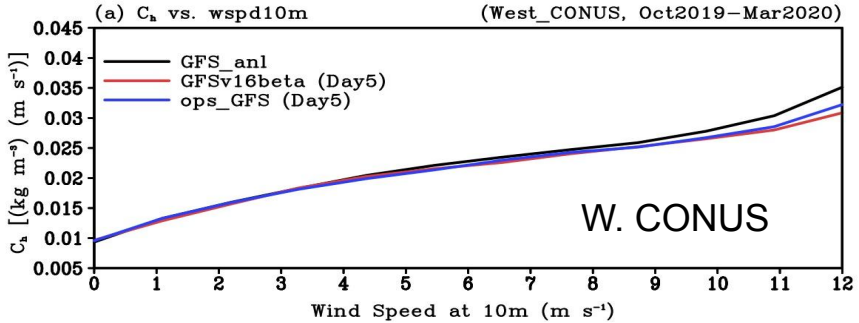


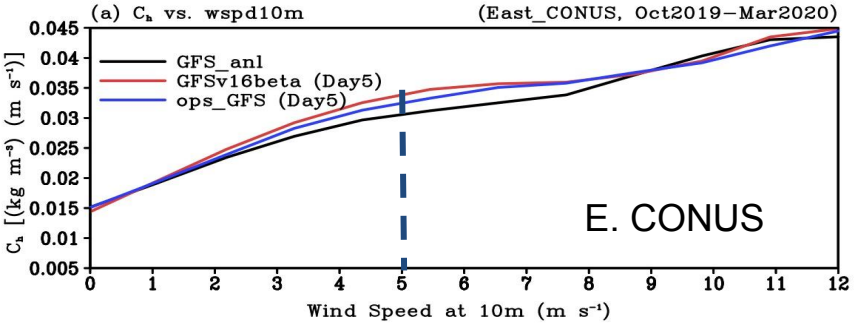


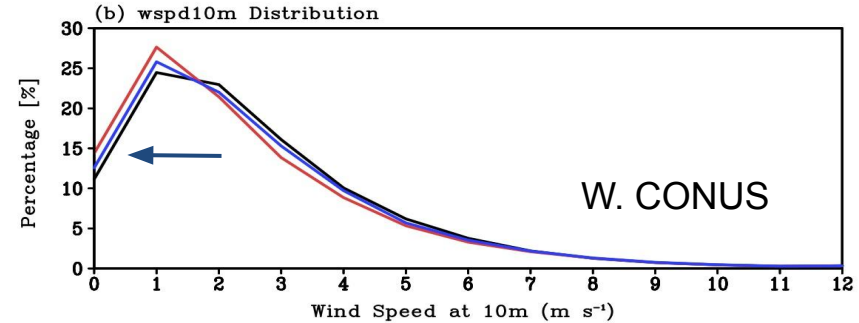
Overestimated PBL height in N.E. CONUS (including Great Lakes):

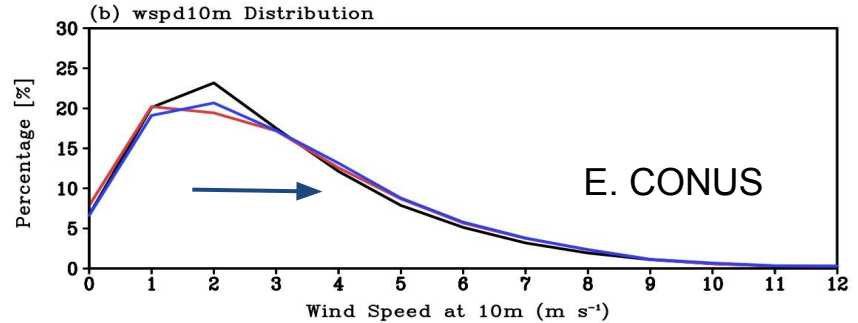
- Collocates with near-sfc wind speed (overmixing)
- Hypotheses for stronger wind:
  a. underrepresented surface drag in PBL scheme (also nearly stable wintertime PBL strongly depends on sfc roughness length)
  b. underrepresented orography variability esp. snow or ice covered areas e.g., Great lakes
  c. destabilization of the stable or nearly neutral PBL - problem in the stability func in the surface scheme

Dependence of exchange coefficient (*Ch*) of heat & moisture on 10-m wind speed (Day5 fcst vs GFS anal)



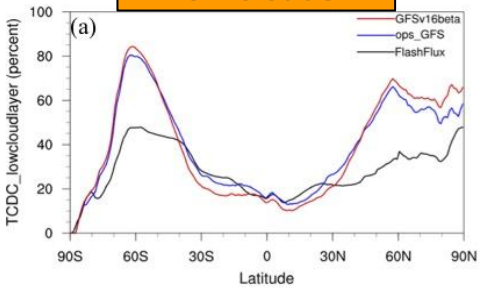




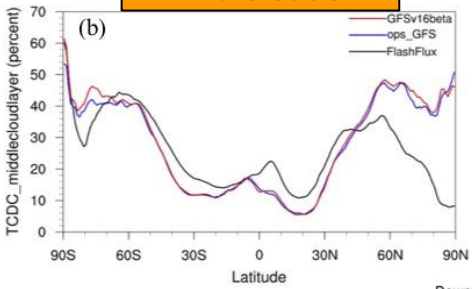


- W. CONUS: overall consistent with GFS anal, although wind speed shifts towards smaller values
- E. CONUS: land-atmospheric exchanges are stronger for a given amount of wspd10m in both v16beta and ops_GFS, and wind speed shifts towards larger values
  - **Ch depends on friction velocity and the stability func in surface scheme** - again, suggesting a possible deficiency in the calculation of Ch and/or stability func in the sfc_diff.f

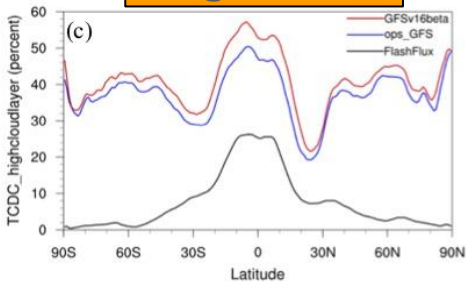
low clouds



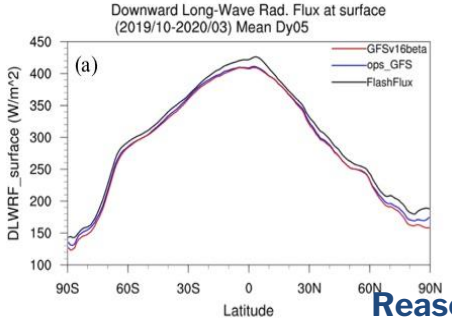
mid clouds

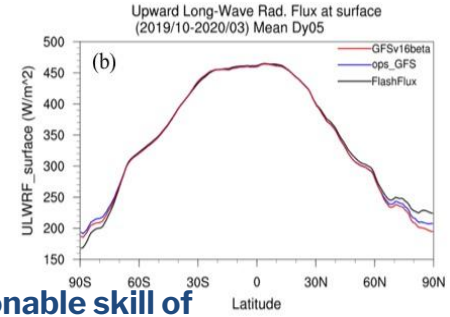


high clouds

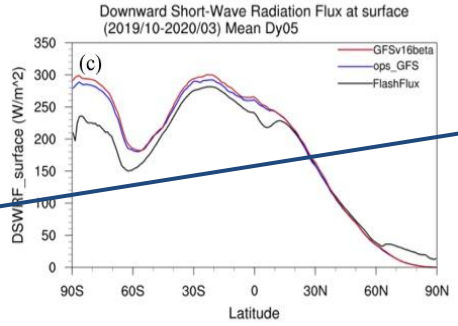


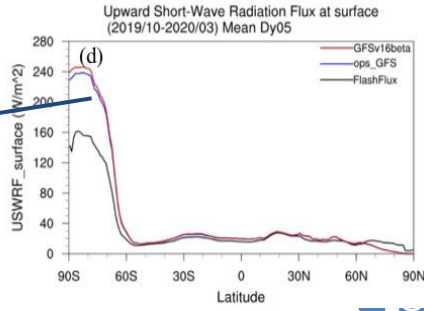
A contradiction between the cloud diagnostic outputs and radiation biases has to be explored further





Reasonable skill of predicting LW radiation





- Excessive clouds fraction contradicts with skillful surface radiation forecast
- Overpredicted downward shortwave radiation may suggest underestimated cloud thickness
- +albedo bias (not shown) at S. pole contributes to overestimated upward shortwave radiation

Summary

This T&E activity focused on the perpetuating GFS forecast errors pertaining to the PBL, land-surface, radiation, and cloud processes.

- GFSv16beta **outperformed** ops_GFS (v15) with respect to RMSEs of large-scale environmental variables and ACC for 500 hPa GPH.
- However, **larger biases** associated with key physical processes were identified in the GFSv16beta forecasts, including that,
  - the near-surface and boundary-layer **cold biases** is partly related to the **systematic negative temperature errors in the GFS analysis**.
  - the overestimated near-surface wind speed particularly at night in the northeastern U.S. implies that the **surface drag may be underrepresented and stability function may play a role**
  - **a contradiction between the cloud diagnostic outputs and radiation biases** has to be explored further.